Quantifying impacts of forecast uncertainties on predicted storm surges

Donald T. Resio¹ · Nancy J. Powell² · Mary A. Cialone³ ·
Himangshu S. Das⁴ · Joannes J. Westerink⁵

Received: 23 January 2017 / Accepted: 11 May 2017 / Published online: 25 May 2017
© The Author(s) 2017. This article is an open access publication

Abstract In this paper, we propose a framework for quantifying risks, including (1) the effects of forecast errors, (2) the ability to resolve critical grid features that are important to accurate site-specific forecasts, and (3) a framework that can move us toward performance-based/cost-based decisions, within an extremely fast execution time. A key element presently lacking in previous studies is the interrelationship between the effects of combined random errors and bias in numerical weather prediction (NWP) models and bias and random errors in surge models. This approach examines the number of degrees of freedom in present forecasts and develops an equation for the quantification of these types of errors within a unified system, given the number of degrees of freedom in the NWP forecasts. It is shown that the methodology can be used to provide information on the forecasts and along with the combined uncertainty due to all of the individual contributions. A potential important benefit from studies using this approach would be the ability to estimate financial and other trade-offs between higher-cost “rapid” evacuation methods and lower-cost “slower” evacuation methods. Analyses here show that uncertainty inherent in these decisions depends strongly on forecast time and geographic location. Methods based on sets of surge maxima do not capture this uncertainty and would be difficult to use for this purpose. In particular, it is shown that surge model bias can play a dominant role in distorting the forecast probabilities.

Keywords Storm surge · Forecasting · Quantifying risk · Uncertainty

✉ Donald T. Resio
don.resio@unf.edu

¹ University of North Florida, Jacksonville, FL, USA

² ARCADIS, Kenner, LA, USA

³ Engineering Research & Development Center, Vicksburg, MS, USA

⁴ Jackson State University, Jackson, MS, USA

⁵ University of Notre Dame, Notre Dame, IN, USA

1 Introduction

The devastation associated with Hurricanes Ivan, Katrina, Rita, Ike, and Isaac, combined with the recent impact of “Superstorm” Sandy, has generated a renewed interest in accurately evaluating hurricane hazards and risks within the coastal areas of the USA. Prior to Hurricane Katrina, studies of hazards and risks primarily relied on methods that utilized only small sets of storms (for example, a limited number of hypothetical storm events or historical storms selected from a local area). Although some methods, such as the empirical simulation technique (EST) used resampling methods to estimate aleatory uncertainty for life-cycle simulations, they neglected overall uncertainties in hazard/risk estimates. Following Katrina, methods for surge risk estimation were re-evaluated; and improved hazard/risk estimation techniques were developed (Resio et al. 2009; Irish et al. 2009; Toro et al. 2010; Niedoroda et al. 2010; Irish and Resio 2010; Irish et al. 2011; Resio et al. 2013). These studies not only emphasized the need to consider the entire range of possible storms that could affect an area but also showed that it was important to recognize and quantify the role of uncertainty in hazard/risk estimates and the impact of this uncertainty on decision-making in that area.

Coastal zones represent some of the most diverse and threatened areas within the USA in terms of potential loss of life, loss of livelihood, and loss of habitat. Hazards regularly impact human safety, private and public property, ecologic health and economic productivity. Attempts to treat these factors additively inevitably lead to heated discussions of the relative importance of each of these factors; however, this does not diminish the need to consider all of these factors. Another complication is that the timescale of needed information spans a range that includes long-term planning, pre-storm preparations, storm evacuations, and post-storm recovery.

Given the increasing vulnerability and complexity of different processes and time scales that must be considered in planning and emergency operations in coastal areas, the focus of many investigations has shifted from coastal hazards and risks to coastal resilience (NRC 2014). In this context, it is extremely important to recognize that many needs for coastal information exist outside of its critical role in community evacuations. This objective of this paper is to develop a framework that can meet the needs of coastal communities in this broader context, which must recognize the existence of many individuals, private industry, local and state government, and even federal government which require information substantially beyond that which is available today. In particular, the coupling between hazards and consequences must be better quantified to provide this information for all its applications over a range of scales and decision-makers. Thus, while the information on expected storm impacts for evacuations may be suitable for most evacuation managers, many operations, such as the closure of floodgates and navigation locks, the shutdown of manufacturing plants, and the timing and need for site preparations by various entities is difficult to understand without some knowledge of the impacts of different decisions.

The framework developed in this paper proceeds as follows:

1. A general formulation of surge probabilities for a forecast event;
2. An analysis of the relative role of random errors and bias in numerical models relative to errors in forecast hurricane parameters;
3. The introduction of a simple means for estimating correct and incorrect decisions, given a specified decision threshold; and
4. The expansion of this forecast method to incorporate consequences in decision-making.

Section 2 examines existing prediction methods used in forecasts. Section 3 provides an overview of the forecasting framework for predicting surges. Section 4 expands this framework to include a theoretical approach for including uncertainty. Section 5 develops a methodology for quantifying errors in forecast surges related to errors in forecast hurricane parameters at landfall. Section 6 develops a method for estimating the Surgesurf Response Functions from a discretized set of surge model executions, in this case using an available set of 152 Southeast Louisiana storms from the Interagency Performance Evaluation Task Force (IPET) study. Section 7 investigates the potential effects of combined random errors and bias in numerical weather prediction (NWP) models and surge models. Sections 7 and 8 present a discussion of results and our conclusions, respectively.

2 Existing methods

Recently, the U.S. National Weather Service has introduced a new product (slosh.nws.noaa.gov/psurge2.0) to address the pre-event forecast aspect of this information need. The system generates a “Potential Storm Surge Flooding Map,” containing the 90th percentile estimates of coastal inundation expected from tropical cyclones. It is intended for decision-making related to National Hurricane Center mission in this area and is an excellent step toward including uncertainty into forecast surge information. However, for consistency with existing National Hurricane Center forecast information, it uses the same model (SLOSH) developed by Jelesnianski (1967) and Jelesnianski et al. (1992) as used in the real-time forecasts, which limits the time available for execution to a relatively small increment that fits within the portion of a forecast “cycle time,” (the total clock time allotted for all necessary predictions in TC forecasts). This limitation significantly affects the spatial resolution and the physical representations used within these computer simulations.

As shown by Leuttich et al. (2013) and Kerr et al. (2013), the SLOSH model has not performed as well as unstructured grid models developed more recently that do not share the run-time constraint imposed on the operational SLOSH model. Unstructured circulation models and SLOSH were executed using the same wind stress and pressure forcing, but despite this and despite using a Gulf scale domain, the SLOSH runs consistently underperformed in comparison with the unstructured grid models. In general, the high-water mark (HWM) mean absolute error was 2–3 times greater for the SLOSH simulations than for the unstructured mesh model simulations and there was significantly more bias in the SLOSH comparisons. The impact of mean absolute errors (bias) will be shown to be significantly different than the impact of random errors in quantifying impacts of uncertainty in this paper.

Although a single-valued estimate of surge over a relatively large domain might be relevant for evacuation, it might not be relevant for many other types of decisions. Today, simple “rules of thumb” and/or expert opinions are often used to supplement forecast information in order to convert single-valued information, such as the mean expected surge, maximum expected surge or 90th percentile surge, into actionable decisions. Given recent test results using SLOSH (Leuttich et al. 2013; Kerr et al. 2013), the accuracy of this model as the basis for operational decision-making other than for evacuation are somewhat questionable. Also, for applicability to planning, such information does not convey information about the magnitude of uncertainty in these estimates nor a quantitative estimate of potential impact of uncertainty on the impacts of these decisions. Improving

resilience will require such quantitative information for all timescales, from long-term planning to pre-storm preparations.

Uncertainty can be very challenging to decision-makers, and reliable, objective methods for incorporating uncertainty into the evaluation of options are often lacking. In most fields, treatments of uncertainty involve the development of an overall cost function, which can then be optimized for different operational scenarios. In the case of mandatory evacuations, this is a difficult topic since it involves possible loss of life. Hence, such a cost function might have to be limited to decisions other than evacuation. They would also have to be geographically specific and would have to consider human, economic and ecologic health separately. Such a generalized tool is beyond the scope of this paper; however, we feel it is important to begin moving toward this goal.

In this paper, we shall use examples from the New Orleans area to show the utility of this methodology, but do not have the information to generalize these impacts into a unified mortality, economic and ecologic framework. In the New Orleans area, certain types of operations, such as closing large barriers across an active waterway and drainage canals, are extremely site specific and can have high associated execution costs. Consequently, it is a good example of a location where very local information and the inherent uncertainty are both critical to optimal decision-making.

Historically, forecasts have been presented in terms of ranges of values over relatively large areas, either based on outputs from a single model or synthesized from a suite of models. Such results are difficult to interpret in a probabilistic fashion, since a clear measure of the uncertainty is not included. As will be shown in this paper, this “broad-brush” approach seems appropriate on one hand, since uncertainties in forecasts tend to add a widespread random component to all of the estimated values; however, as was shown in Sandy, vulnerability to inundation in a city such as New York can be extremely localized. Models that do not properly resolve critical small-scale features affecting local inundation levels may yield erroneous results in areas in locations where the vulnerability is critically sensitive to water levels exceeding a particular threshold value (such as tunnel entrances, access routes to subways, and engineered flood protection structures).

As might be expected from the above perspective, there has been considerable debate over the last decade or so concerning the needed accuracy of surge models used in forecasting. Several papers have attempted to establish some level of understanding of the relative errors in differing models (Leutich et al. 2013; Kerr et al. 2013; Resio and Irish 2015); however, as will be shown here, it is important to understand errors in storm surges relative to the errors in forecast hurricane characteristics. It is hoped that this paper will help provide the scientific and engineering community with at least a preliminary estimate of the implications of surge modeling accuracy for forecast decision-making.

3 Overall forecast framework

In a forensics study or in investigations of potential storm surge model performance, detailed wind fields, prepared from a combination of analyses and measurements, are typically used to drive high-resolution numerical simulations (Powell et al. 2010; Dietrich et al. 2010; Bunya et al. 2010). Such wind fields are extremely useful for evaluating the physics and numerical approximations within models used for surge prediction. However, existing prediction methods for hurricanes typically provide forecast information on only a set of hurricane parameters, typically storm intensity, storm size characteristics, forward

speed, angle of track intersection with the coast and landfall location. This creates somewhat of a mismatch between the inputs used for testing storm surge model performance with hindcast wind fields and the information available for typical surge forecasts.

Since numerical storm surge models require detailed wind speeds and directions over the entire simulation region, a method for deriving a realistic spatial distribution of wind speeds and directions over the entire domain from forecast parameters is needed. In the 1970s, parametric “vortex” models assumed that four parameters sufficed for hurricane wind field estimation, the pressure differential (pressure at the periphery of the storm minus the central pressure), storm size (typically in terms of the radial distance from the center of the storm to the maximum wind speed), the forward storm speed, and the angle of storm movement (Chow 1971; Myers 1975). Holland (1980) showed that the “peakedness” of the wind speed distribution along a radial line within hurricanes varied from storm to storm and through time within any given storm. He introduced an additional term into the vortex equation, called the “Holland B” parameter to allow storm wind fields to incorporate this type of variability, increasing the number of parameters used to characterize hurricane winds to five. Planetary boundary layer (PBL) models (Thompson and Cardone 1996; Vickery et al. 2000) have now been developed to estimate wind fields within a hurricane from similar or slightly expanded sets of parameters. Since forecast information on storm size and the Holland B parameter is generally lacking, this information is generally obtained for applications either by assuming persistence or from the regional climatology. It should be noted here that a sixth parameter, proximity of a site of interest to the landfall location, will be added to this set of parameters later in the paper since this is critical to local surge heights.

In recent years, it has been recognized that hurricanes undergo systematic changes as they approach a coast, even before making landfall. In particular, it has been shown that intense hurricanes tend to weaken, with an accompanying increase in storm size (IPET 2008; Levinson et al. 2010; Rappaport et al. 2010), while weaker storms can maintain their strength or even intensify. At this point in time, such subscale effects are typically not well predicted by existing forecast models and must be estimated parametrically from climatological analyses. A simple “fix” to this problem might be to assume that storms retain their offshore intensity up to landfall. This would be a conservative assumption for strong storms and might be reasonable for purposes of evacuation planning, but in general, this would result in significant biases in forecast surges for intense storms.

From this discussion, it can be seen that forecast wind fields used today for forecasting storm surges remain relatively rudimentary; however, the focus of this paper is not to improve deterministic forecast capabilities, but rather to examine the inherent uncertainty in forecast surges. Adding the distance between a point of interest as one more parameter, which brings the total number of hurricane parameters to six, we will assume that values for all of these parameters (central pressure, storm size, forward speed of the storm, angle of track with respect to landfall, landfall location and the value of the Holland B parameter) will be available at each forecast time, even if its forecast value represents only a climatological average or an estimate from an independent analysis. The exact values are used in such parameterizations depends on which parametric or PBL model is used to estimate the wind fields.

Given the obvious differences in the wind and pressure forcing, it is clear that surge simulations driven with detailed, post-storm wind fields may not exhibit error characteristics similar to those generated in simulations using parametric wind fields of the type available for forecasts. In spite of this, forecast surge levels along the coast are in many

respects very similar to hindcast surge levels, and, similar to these model applications, to ensure accurate, unbiased surge estimates, it is critical for four primary criteria to be met:

1. The bathymetry and topography must be accurately resolved and represented within the surge model;
2. All forcing mechanisms (winds, waves, pressures, etc.) must be accurately specified and accurately captured within the surge model;
3. The numerical methods embodied within the model must properly reflect the actual physics of all significant surge generation processes; and
4. The model error characteristics must be quantified for areas in which they are used.

4 Theoretical framework for forecast surge uncertainty

In general, the error/uncertainty in forecast hurricane surges comes from two main sources: (1) inherent errors in the surge modeling system (i.e., the errors which would persist even given a “perfect” forecast), and (2) uncertainties in the forecast hurricane tracks/characteristics. The first source is related to the predictive accuracies of existing surge models (when forced with parametric wind fields) and is independent of forecast time. The second source is related to the accuracies of existing NWP models utilized for predicting hurricane tracks and other characteristics, which is very dependent on forecast time. Random errors within the Advanced Circulation (ADCIRC) model (Westerink et al. 1992) were evaluated during the post-Katrina study conducted by IPET (2010); and here we shall simply adopt their estimates to characterize random errors inherent in the surge modeling system. As will be shown here, the role of potential surge model bias is also very important to forecast accuracy. The focus of this paper will be to consider the combined effects of all models related (NWP models and surge models) to a unified framework.

Let us assume that a forecast has been issued at time t_0 , for some time in the future $t_0 + \Delta t$, where Δt is the increment of time between the time of the forecast and the time for which the forecast is valid. Each NWP model produces a single deterministic outcome for what will happen at time $t_0 + \Delta t$; however, as noted previously, each forecast will contain errors that cause the forecast characteristics to deviate from the actual characteristics which occur in the future. Given a deterministic prediction of a hurricane parameter set, suitable for driving a PBL model for forcing storm surges, and a specified error distribution for each of these parameters, we can write the value of the parameter at a future time as

$$x_i = (\hat{x}_i - \beta_{w,i}) + \varepsilon_{w,i} \quad (1)$$

where x_i = forecast value of the i th parameter; $\beta_{w,i}$ = persistent deviation between the forecast parameter and actual future value (bias); $\hat{x}_i$ = deterministic prediction of the i th parameter; and $\varepsilon_{w,i}$ = random deviation between the actual and forecast i th wind field parameter value.

In Eq. 1, it should be understood that the random value that occurs at time $\delta t = t_0 + \Delta t$ is conditionally dependent on the forecast value, i.e. $p(x_i) \equiv p(x_i|\hat{x}_i)$. For simplicity, we also assume that we have available data which allows us to estimate a continuous “deterministic” multidimensional Surge3 Response Function (SRF), represented here as $\Psi(x_1, \dots, x_n)$, which relates a set of hurricane parameter values to an estimated surge at a point based on a simulated set of storms for a comprehensive set of storm parameters. Meteorological forecast models have been tested in a consistent manner with annually

published performance evaluations and have been calibrated to achieve near-zero bias, so we will assume on an ad hoc basis here that $\beta_{w,i} \approx 0$ for all parameters. We also recognize that, at present, there is insufficient surge data to stratify our sample into dependences on variations in wind-forcing parameters and will use an overall average bias term for surge model errors here. These assumptions allow us to estimate cumulative distribution of surges directly from the set of parameters available from the forecasts,

$$\hat{F}(\eta) = \int \dots \int p(x_1, \dots, x_n, \varepsilon_{w,i}, \varepsilon_m | T_f, I_f, R_{\max,f}) H[\Psi(x_1, \dots, x_n) + \beta_m + \varepsilon_{w,i} + \varepsilon_m - \eta] dx_1 \dots dx_n d\varepsilon_{w,i} d\varepsilon_m \quad (2)$$

where $\hat{F}(\eta)$ is the cumulative conditional distribution function for surge, η , given T_f , I_f , and $R_{\max,f}$; T_f is the forecast track at time t_0 ; I_f is the forecast intensity at time t_0 ; $R_{\max,f}$ is the forecast storm size parameter at time t_0 ; $\varepsilon_{w,i}$ is the random error in i th NWP predicted parameter; ε_m is the random error in the numerical surge model; β_m is the bias in the numerical surge model; $H(\cdot)$ is the Heaviside function, defined as

$$\begin{aligned} H(\cdot) &= 1, \quad \text{if } (\cdot) \geq 0; \\ H(\cdot) &= 0, \quad \text{otherwise; and} \end{aligned}$$

Ψ is a function which relates the parameter set and model error characteristics to storm surge.

Since no detailed studies exist for quantifying the covariance among forecast errors for different parameters, we assume here that forecast errors are approximately independent. In this case, the joint probability function in Eq. 2 is simplified to the form $p(x_1, \dots, x_n) \approx p(x_1) \dots p(x_n)$, reducing the probability function for forecast errors to a multi-normal distribution. From Eq. 2, we see that the uncertainty in the forecast errors are expected to be a function of three primary factors: existing bias in modeling errors, random errors in surge modeling systems and random errors in forecast hurricane parameters.

4.1 Forecast decisions for an all possible storms

Equation 2 is generalized to examine the potential effects of storms over all possible tracks instead of a single track via an additional integration over all tracks within the model domain, i.e.

$$\begin{aligned} F(\eta) &= \sum_{T_{f,i}} \int \dots \int p(x_1, \dots, x_n, T, \varepsilon_{w,i}, \varepsilon_m | T_{f,i}, I_f, R_f) H[\Psi(x_1, \dots, x_n) \\ &\quad + \beta_m + \varepsilon_{w,i} + \varepsilon_m - \eta] dx_1 \dots dx_n d\varepsilon_{w,i} d\varepsilon_m \Delta s(T_i) \end{aligned} \quad (3)$$

where the subscript “f” is used to denote the forecast value; $T_{f,i}$ is the i th subincremented track in the matrix of archived tracks; $F(\eta)$ is the cumulative distribution function for surge for all tracks, η ; and $\Delta s(T_i)$ is the incremental distance between tracks for i th track in the simulation domain.

Since track location is already included in Eq. 3, this new integration represents a convolution of the forecast track and the error function. If the discretization of the tracks is constant along the coast and the frequency of tropical cyclones (TCs) is constant along this domain, the incremental frequency can be moved outside the summation in Eq. 3.

5 Estimation of surge errors related to errors in forecast parameters

Surge forecasts are typically derived from deterministic computer simulations using forecast hurricane information from a single model or an ensemble of models. It has been shown by several investigators that “best-track” estimates from an ensemble of model runs provide a better estimate of forecast storm tracks than the use of a single simulation model (Zhang and Krishnamurti 1997, 1999; Hamill et al. 2012). It is often assumed that the spatial deviations among the various forecast tracks provide a reasonable estimate of the uncertainty in the forecast. Such variations provide a good measure of what is different among the different NWP models, and they represent some of the randomness inherent in these predictions. However, such variations, in general, may not represent all of the randomness/uncertainties in the forecast tracks since they contain elements which are mutually correlated in all of the models, such as physical assumptions, numerical methods, and initial conditions. As an example of this, the actual track of Hurricane Rita in 2005 fell significantly outside of all of the ensemble tracks predicted 2–4 days before landfall. The ensemble tracks were all relatively tightly clustered, suggesting a high confidence in the track; however, the set of forecast tracks was significantly displaced westward from the actual landfall location. The numerical guidance indicated a significant threat of landfall for the Houston, Texas, region on 24 September. The largest mass evacuation in the US history was ordered, with the evacuation resulting in more fatalities than Hurricane Rita itself (Galarneau and Hamill 2015). For this case, variations within the ensemble tracks dramatically underestimated the uncertainty in the forecast track. In many other cases, the opposite is true with the spread in landfall location being much greater than the deviations expected from the official error statistics, such as happened for Hurricane Katrina that same year. Such inconsistency makes the use of ensemble tracks somewhat ineffective as an indicator of expected uncertainty in forecasts. For this reason, we utilize the official error statistics as a stable measure of potential deviations in this paper rather than using ensemble characteristics; however, this choice should be revisited in subsequent efforts.

Storm surges in coastal areas have been shown to depend primarily on the proximity of a geographic point to the location of maximum onshore winds, storm intensity, and storm size (Irish et al. 2009). To a lesser extent, they are also dependent on the forward speed of a storm and the angle of intersection between the storm track and the coast (Niedoroda et al. 2010). If we were focusing on the timing of the storm surge, it would definitely be critical to include errors in storm speed; however, for the purposes of this paper, we shall concentrate only on errors in the magnitude of storm surge. This allows us to focus on errors in storm landfall location, storm intensity and size. Inclusion of additional uncertainties in storm speed and track angle would be expected to slightly increase the uncertainty in surge magnitude over the values reported here.

Over the last decade or so, significant improvement has been made in forecasting hurricane track. Typically, this error is characterized in terms of the radial distance between the forecast location at time $t + \Delta t$ and the actual location at that time, neglecting the angle of the deviation. In the context of the forecast radial errors, it is understood that cross-track errors in storm position will produce deviations in the magnitude of the surge, while along-track errors in storm position will produce deviations in the timing of the surge.

Figure 1 shows the official forecast errors from the National Hurricane Center (NHC) over a recent 20-year period. In this figure, the measure of radial error used is the root-mean-square (rms) error defined as

Fig. 1 Variation in root-mean-square error in track forecasts: 1970–2010 (available from <http://www.nhc.noaa.gov/verification>)

$$\varepsilon_r(\delta t) = \frac{1}{n(\delta t)} \left\{ \sum_{i=1}^n [x_i(\delta t) - \hat{x}(\delta t)]^2 + [y_i(\delta t) - \hat{y}(\delta t)]^2 \right\}^{1/2} \quad (4)$$

where $\varepsilon_r(\delta t)$ is the root-mean-square radial error for a forecast at $t_0 + \delta t$, $x(\delta t)$, $y(\delta t)$ are the observed spatial coordinates at time $t_0 + \delta t$, $\hat{x}(\delta t)$, $\hat{y}(\delta t)$ are the observed spatial coordinates at time $t_0 + \delta t$, and n is the number of samples considered for comparisons at $t_0 + \delta t$.

Clearly, the radial error in forecast position has been reduced by substantially during the last two decades, from about 185 km in 1988 to about 93 km in 2010. Studies at NHC have shown that the cross-track error is slightly smaller than the along-track errors; consequently, the standard deviation of cross-track deviations is actually only about 85% of the total rms shown in this figure.

Unfortunately, as can be seen in Fig. 2, there has been little or no improvement in the intensity forecasts over the last decade. Studies conducted by the NHC use maximum storm winds as their predictand rather than central pressure. To convert this to a measure of error in central pressure (or equivalently the pressure differential) as utilized in the archived data set available to this study, we must determine the overall relationship between maximum storm speed in m/s and central pressure in millibars. Based on the Saffir and Simpson pressure–speed equivalences, we find that this conversion can be approximated as (using mb to abbreviate millibars and m/s to abbreviate meters per second),

$$\sigma_{mb} \approx \sigma_{m/s} \frac{\partial c_p(\text{mb})}{\partial V_{\max}(\text{m/s})} \approx 0.54 \sigma_{m/s} \quad (5)$$

where σ_{mb} is the standard deviation in central pressure predictions (in mb); σ_{kts} is the standard deviation in predictions of maximum wind speed (in m/s); c_p is the central pressure in mb; and $V_{\max}$ is the estimated maximum hurricane wind speed (in m/s).

Fig. 2 Variation in intensity forecasts (in terms of maximum sustained surface wind speed): 1970–2010 (available from <http://www.nhc.noaa.gov/verification>)

There have been limited investigations into the errors in predicted storm size. In fact, most of the predictions of storm size use relatively qualitative terms such as large or small. For the purpose of this paper, we shall adopt some very rough values taken based on the standard deviation in the change of storm size from its initial value over the forecast interval. Since storm size is typically held relatively constant in forecasts, this should provide at least a first cut estimate of the expected deviations.

Table 1 presents the values for the standard deviations developed here in: (1) storm landfall location along the coast (km), (2) storm intensity (central pressure in mb), and (3) storm size (km). As can be seen in this table, the estimates are sufficiently large that they are expected to impact the uncertainty in forecast surges significantly. For simplicity, since errors in track angle do not have as large of an impact on surges as these three terms, we will neglect errors related to errors in the track heading in the forecast; however, this will be considered within the estimation function used in the context of a deterministic forecast.

Table 1 Estimated standard deviations for forecast errors as a function of time for storm along-coast error in track location at landfall (δx), storm size ($R_{\max}$), and central pressure (c_p)

	12 h	24 h	36 h	48 h	72 h
Track error at landfall (km)	15	30.37	44.26	60.37	84.45
Radius to maximum winds (km)	15.55	22.41	25.18	26.11	27.78
Central pressure (mb)	11.8	13.5	16.1	20.5	22.8

The time reference in hours is the time between the forecast and the predicted landfall time, for example, the 24-h errors are specific to a forecast made 24 h before the expected landfall

6 The estimation of a surge response function, Ψ , from a discretized storm set

The forms of Eqs. 3 and 4 imply that the surge at a location can be defined as a continuous function of some set of parameters, i.e., $\Psi(x_1, \dots, x_n)$; however, the information used to generate such a function comes from a set of computer simulations utilizing discrete values of these parameters. As the number of parameters used to characterize the hurricane wind field increases and the number of simulated values for each parameter also increases, the number of simulations required can become prohibitively high, particularly for high-resolution grids that are required for accurate hazard/risk characterization (Resio and Westerink 2008; Bunya et al. 2010; Dietrich et al. 2010). The initial 152 storm set of computer simulations from the IPET study represented an attempt to obtain a balance between the robustness of the total set of runs and the number of executions necessary. Based on a number of considerations, the final set of parameters chosen for inclusion in the wind field model used in these simulations were: (1) landfall location, (2) storm intensity, (3) storm size, (4) storm heading angle, (5) storm forward speed, and (6) a climatological average value of the Holland B parameter. Additionally, the storm intensity, size and the Holland B parameters were all allowed to vary systematically as storms approached the coast (IPET 2009). The rest of this section develops a method for estimating continuous functional interrelationships from the discretized set of IPET simulations.

As shown by Niedoroda et al. (2010), landfall position along the coast (δx), storm intensity (parameterized by the pressure differential, ΔP), and storm size (parameterized in terms of radius to maximum wind speed, $R_{\max}$) are the primary parameters affecting surge response in most areas, so we will first develop the Surge Response Function for these three parameters. Simple analytical arguments and numerical experiments (Resio and Westerink 2008; Irish and Resio 2010) have shown that the surge varies approximately linearly with pressure differential. Thus, interpolations/extrapolations of surges between different pressure differential values for fixed values of $R_{\max}$ and location of landfall are straightforward. Irish et al. (2008) showed that surge height increased monotonically with $R_{\max}$, so interpolations/extrapolations of surge heights for fixed values of storm intensity and storm size is also straightforward. However, in the 152-storm set developed for the IPET study, the along-coast track distribution for the central angle of storm incidence was represented by the five tracks shown in Fig. 3. The spacing between tracks is approximately 39 nautical miles, which is relatively large, so it is important to estimate the along-coast variation in surge height, given constant values of ΔP and $R_{\max}$, in a fashion that preserves the peak surge levels. A simple estimator for this can be written as

$$\begin{aligned}\eta_{\max} &= \eta_j + \frac{|\delta\eta_j - \delta\eta_{j+1}|}{4} \\ \delta\eta_j &= \eta_j - \eta_{j-1} \\ \delta\eta_{j+1} &= \eta_j - \eta_{j+1}\end{aligned}\quad (6)$$

where η_j is the maximum surge value, which occurs in the j th track; $\eta_{\max}$ is the estimated maximum surge; and

$$x_{\max} = j - \frac{|\delta\eta_j - \delta\eta_{j+1}|}{2\delta_{\text{tot}}}$$

where $\delta_{\text{tot}} = \delta_j + \delta_{j+1}$ and $x_{\max}$ is the location of the maximum surge.

Fig. 3 Map showing spacing of the five primary tracks used in the IPET surge simulations

Fig. 4 Surges produced by storms along Tracks 1–5 with a maximum value estimated via Eq. 6, between adjacent points with highest surges shown by *symbols*

Figure 4 shows a plot of the along-track distribution of surges with the maximum surge value, estimated from Eq. 6, included with the values at each of the 5 tracks, for three locations in the vicinity of closure structures in the New Orleans Hurricane and Storm Damage Risk Reduction System (HSDRRS) shown in Fig. 5. As can be seen in Fig. 4, the peak values between the tracks can be 0.3–0.6 m higher than the highest value at any of the tracks. The method described here has been shown to help remove spurious crenulations in the along-coast distribution of risks (Niedoroda et al. 2010) related to the relatively large spacing between adjacent tracks utilized in the original Joint Probability Method (JPM) study for southeast Louisiana. Utilizing this approach to establish a set of values of $R_{\max}$, at a fixed value of ΔP , estimates for arbitrary storm sizes can be obtained via linear interpolation/extrapolation, which yields a full two-dimensional estimator of surges for any value of along-coast landfall location and storm size $\Psi_{2D}(\delta x, R_{\max})$, where the subscript “2D” is added to explicitly note that the Surge Response Function at this point only includes two parameters (i.e., two dimensions in the response function). Since the dependence of surge levels on pressure differential is quite linear over a substantial range, the three-dimension Surge Response Function, $\Psi_{3D}(\delta x, R_{\max}, \Delta P)$, can also be estimated via linear interpolation. This methodology can be used to form a discretized matrix with arbitrarily small intervals in all three dimensions for subsequent numerical integration. For this paper, we shall use the following step sizes and ranges for numerical integrations based on these three parameters:

$\delta x_i = 1\text{--}51$ (interpolated tracks), spacing 1/10 of the major track spacing (or 7.2 km);
 $(R_{\max})_j = 59.3\text{--}111$ (km), in increments of 1.85 km; and
 $\Delta P_k = 30\text{--}150$ mb, in increments of 1 mb.

Fig. 5 Map showing locations points 305, 557 and 345

This discretization allows detailed resolution of the integrand in Eq. 3.

The effect of landfall angle, relative to the central angle in the IPET storm set, can be estimated in a fashion similar to that recommended in Resio et al. (2009), i.e.

$$\Psi_{4D}(\delta x, R_{\max}, \Delta P, \theta_f) = \Psi_{3D}(\delta x, R_{\max}, \Delta P) + \left(\frac{\partial \eta}{\partial \theta_f} \Delta \theta_f \right)_{\delta x} \quad (7)$$

where θ_f is the angle of the track with respect to the central angle at landfall; $\Delta \theta_f$ is the increment between the two angles, and the subscript “ δx ” shows that the derivative is taken for a fixed value of δx .

Typically, the effect of angle variation is not too large for the area studied here. Figure 6 shows the effect of a variation of 45 degrees in the track angle for Point 557 on the south shore of Lake Pontchartrain, which exhibited the highest variation of all of the points examined. Even at this location, the maximum effect of angle variation on surges is only about 15% and is less than 5% in the region of maximum surge response.

The effect of forward speed on surge levels at a point is also rather small and can be written in a similar form to Eq. 7,

$$\Psi_{5D}(\delta x, R_{\max}, \Delta P, \theta_f, v_f) = \Psi_{4D}(\delta x, R_{\max}, \Delta P, \theta_f) + \left(\frac{\partial \eta}{\partial v_f} \Delta v_f \right)_{\delta x} \quad (8)$$

where v_f is the forward storm speed and Δv_f is the increment between the two forward speeds, and the subscript “ δx ” shows that the derivative is taken for a fixed value of δx .

Similar to Niedoroda et al. (2010), we find that the differences in surges due to forward speeds are fairly small, with less than 15% deviation in all cases between the storms moving at 5.5 m/s and those moving at 3 and 8.5 m/s. A typical comparison in which only storm speed is varied is given in Fig. 7a, b for point 305, in proximity of the Inner Harbor Navigation Canal-Lake Borgne Surge Barrier.

To this point, we have examined contributions to wind field uncertainty related to errors in storm landfall position, storm intensity, and storm size. Errors in forecast storm speed

Fig. 6 Surge variation with angle of approach at Point 305 for landfalling storms along track slightly to the west of this point

Fig. 7 Surge variation with forward storm speed at Point 305 for specific tracks and intensities annotated in charts

primarily produce errors in the timing of surges and errors in track angle at the coast contribute less to the overall surge response than the three parameters considered here. Additional error terms will increase the uncertainty reported here and can be incorporated in subsequent efforts.

Once a set of storms has been simulated, the information contained within the archived results is generally used primarily for hazard/risk assessment; however, recent studies have shown that this same information can be utilized effectively for a forecast storm. Based on the storm track and other storm parameters (intensity, size, forward speed and track angle), one can select a storm from the archived set with the closest resemblance to the forecast storm (Das et al. 2010). Or if one uses a functional interpolation method as defined in Sect. 5, a forecast can be made from an essentially continuous set of parameters. The latter approach, using a structured sampling approach within an assumed continuous multidimensional parameter domain, makes it much easier to include the uncertainty in a simple direct fashion as opposed to the methods based on probability masses, such as the method used in the Das et al. study or in the Bayesian Quadrature methods (Toro et al. 2010; Nedoroda et al. 2010).

As an example of the need to interpolate from a sparsely sampled set, such as sets available from Federal Emergency Management Agency (FEMA) studies, a storm with a central pressure of 945 mb might fall midway between Tracks 2 and 3 with a radius to maximum winds of 21 nautical miles, we can obtain the estimates for surge (in meters) shown in Table 2. As can be seen here, the interpolated values can deviate substantially from the values at adjacent discrete parameter values. Using estimates at discrete points with their associated probability masses would provide a very crude approximation to the actual error integral.

On the one hand, the use of archived information might appear to lose accuracy since its wind fields are only parametric with a limited number of degrees of freedom, *dof*, but it should be recalled that the forecast information embodies even a smaller number of *dof*

Table 2 Forecast surge levels at Points 305, 557, and 754 for a selected storm

	960 mb 38.9 km Track 2 (m)	960 mb 38.9 km Track 3 (m)	930 mb 32.7 km Track 2 (m)	930 mb 32.7 km Track 3 (m)	Interpolated value (m)
Point 305	3.23	3.81	4.57	5.64	3.29
Point 557	1.58	1.58	2.07	2.35	1.98
Point 754	1.86	0.88	2.80	0.94	1.46

Columns 2–5 contain surge levels estimated on the basis of storms with the closest match to the actual storm parameters, while column 6 contains the interpolated value obtained from the application of the continuous surge response functions developed in Eqs. 6–8 to the surges at each given location

than the number of parameters used to force these hindcasts. Given the current state of the art, the accuracy of interpolations is primarily related to the overall resolution used to discretize these parameters (i.e., track spacing, number of intensities, number of sizes, number of angles of approach). If this is sufficiently fine, interpolations of surge levels between points in this parametric space can essentially be equivalent to using an analog approach (i.e., using continuous parameters in computer simulations).

The application of archived data to forecasts offers several very attractive features. First, grid resolutions used in hindcast simulations are typically much finer than those used in typical forecast simulations. This improved resolution can provide superior surge estimates in many complex coastal areas. Second, a modeling system used for hindcasts is typically better validated for historical storm surges than forecast modeling systems, due to the importance of high accuracy in hazard/risk specifications. Third, since the hindcast model simulations are not executed in “real-time,” more details in processes can be accommodated within the modeling system. This allows coupled wind and wave forcing to be incorporated within the hindcast surge modeling system rather than relying on wind forcing alone, as is presently done in many forecast systems. Fourth, since the detailed computations are completed before a hurricane occurs, the computer time for performing a forecast is extremely small compared to that required to model surges in a typical forecast mode. In fact, using an archived data set, a personal computer (PC) can provide a wide range of products within minutes. Fifth, using the approach described in this paper, it is relatively easy to transition from a deterministic forecast to a probabilistic forecast.

7 Estimation of surge uncertainty in forecasts due to surge model errors characteristics

The forms of Eqs. 2 and 3 are such that the effects of the error terms in NWP models and surge models can be rewritten as a combination of impacts of independently distributed forecast errors (as a function of time) and modeling errors,

$$\left[\left(\sum_i \varepsilon_{w,i} \right) + \varepsilon_m + \beta_m \right] = \varepsilon_{\text{tot}} + \varepsilon_m + \beta_m = \varepsilon_{\text{tot}} \left(1 + \frac{\varepsilon_m}{\varepsilon_{\text{tot}}} \right) + \beta_m, \quad (9)$$

where $\varepsilon_{\text{tot}} = \mathcal{A}(0, \sigma_{\text{tot}})$ with $\sigma_{\text{tot}} = \sqrt{\sum_i \sigma_i^2}$, with $\mathcal{A}(0, \sigma_{\text{tot}})$ representing a Gaussian distribution for a mean value of zero error (unbiased forecast and surge models) and a

standard deviation of total error given by $\varepsilon_{\text{tot}} \left(1 + \frac{\varepsilon_m}{\varepsilon_{\text{tot}}} \right)$. This shows the role of random modeling errors within the same context as random errors related to deviations in forecast storm values as a function of time until landfall. As long as the random error in the forecast surge is much larger than the random error in the forecast, the second term in the last part of Eq. 9 can be neglected and the impact of surge model errors will be relatively small.

Although the importance of the random component of error in the surge model must be considered relative to the random errors expected from deviations in the NWP forecasts, the bias term in Eq. 9 does not have a counterpart in the NWP forecasts, given the assumption that they have been carefully calibrated for many years of testing against observations. Given the importance of bias to the estimation of risk in coastal areas, it is critical that models be carefully calibrated not only for the maximum surge in a storm, but for a wide range of surges generated in different places by different storms. The ADCIRC model has been shown to provide a relatively unbiased estimate of surge levels not only in the region of maximum surges but also in areas of lower surges (Westerink et al. 2008). The SLOSH model, currently used by the NHC has not been as thoroughly tested over as wide of a range of conditions. Currently, there are no published comparisons of SLOSH results which support its capabilities over a wide range of conditions, so it is difficult to make any definitive statement concerning its capabilities in this regard.

To provide some perspective on this subject, the comparisons between ADCIRC and SLOSH results contained in the study of the New York City area by Lin et al. (2010) are shown in Fig. 8. Here we see that the two models converge at a surge value equal to approximately 4.2 m. A best-fit regression line between the two models suggests that the SLOSH model predicts higher values than ADCIRC for surges less than about 4.2 m and predicts lower values than ADCIRC for surges above 4.2 m. The SLOSH model has not been tested as critically as the ADCIRC model in this area, and as noted previously, limitations on the SLOSH model, imposed by run-time constraints in real-time forecasts, can create significant differences in model performance between SLOSH and a set of less

Fig. 8 Comparison of the primary surge at the Battery estimated from the SLOSH and ADCIRC models for the nine highest surge generating storms hindcast by Lin et al. (2010)

constrained models (Leutttich et al. 2013; Kerr et al. 2013). As pointed out in Resio and Irish (2015), the winds in SLOSH used in Fig. 8 may not represent a true estimate of the bias of this model, since other studies, using winds specifically intended for SLOSH applications, have indicated that the SLOSH model does not produce a bias of this magnitude in general. In this paper, we will use these results to quantify the potential role of bias in optimizing decisions based on forecast surges. Since the SLOSH model has not conducted a similar study for official applications of SLOSH in this area, Fig. 8 this should not be considered as an official assessment of the SLOSH model, but rather just as a surrogate to evaluate the potential impact of such a bias.

To demonstrate the importance of variability in location within a storm, surge probabilities at sites 306, 557, and 754 in the New Orleans area are evaluated for a storm predicted to make landfall with a central pressure of 930 mb and radius to maximum winds of 46.3 km, heading due north at 5.5 m/s along a longitude of 88.7. The three points examined here are those shown in Fig. 5): 754 (West Bank), 305 (convergence of St. Bernard eastern levees with the GIWW), and 557 (Lake Pontchartrain). Figure 9 shows the

Fig. 9 Cumulative probabilities of surges for a 96-h forecast at sites: **a** 754 (West Bank), **b** 305 (convergence of St. Bernard eastern levees with the GIWW), and **c** 557 (Lake Pontchartrain)

estimated probabilistic results for surges generated by a storm forecast to strike the coast with the associated parameters given above, made 96-h prior to landfall, while Fig. 10 shows the estimated probabilistic results for a forecast made only 24 h before landfall that is predicted to make landfall at the same location with the same associated parameters at landfall. All plots contain five different estimates:

1. Cumulative probability function of surge with NWP forecast error, but no surge model error;
2. Cumulative probability function of surge with documented NWP forecast error plus surge model error;
3. Cumulative probability function of surge with two times the sum of the documented NWP forecast error plus surge model error.
4. Cumulative probability function of surge given documented NWP forecast error plus surge model error, plus bias based on Fig. 8 data;
5. Documented cumulative probability function of surge for two times the sum of the documented NWP forecast error plus surge model error, plus bias based on Fig. 8 (not doubled).

Fig. 10 Cumulative probabilities of surges for a 24-h forecast at sites: **a** 754 (West Bank), **b** 305 (convergence of St. Bernard eastern levees with the GIWW), and **c** 557 (Lake Pontchartrain)

It is obvious in these two figures that forecast time plays a critical role in the probabilistic interpretation of surges at these sites. In this case, the storm is forecast to move along a track slightly to the east of that followed by Hurricane Katrina in 2005, with a slightly higher central pressure and a slightly larger size.

To contribute to our analyses of the potential impact of uncertainty in surge forecasts, we examine a situation in which we have assumed threshold values for necessary operations at a specific point of interest, located at the confluence of two levee systems along the Gulf Intracoastal Waterway (GIWW), Point 305. If we were to act consistently on the basis of only the deterministic surge forecast, a pertinent question would be what percentage of results might end up in the categories defined as (1) true positive forecasts, (2) false positive forecasts, (3) true negative forecasts, and (4) false negative forecasts. To help answer this question, we examine surges at Point 305 for the set of six storms used to generate the information shown in Table 3, which shows the deterministic forecast results (in meters) at Point 305 for these storms.

For a single track and specified storm parameters, the proportion of surges expected above the threshold of 3 m, η_d , can be estimated by the full integral form given in Eq. 2

$$\int_{\eta_d}^{\infty} p(\eta) d\eta = 1 - F(\eta_d) = \text{AEP}(\eta_d), \quad (10)$$

while the proportion of surges below the threshold is given by $1 - \text{AEP}(\eta_d)$. For Track 1 with a central pressure of 920 mb and for Track 3 with central pressures of 920 and 940 mb), the deterministic forecast is above the threshold. For these cases, the ratio of false positives to true positives is given by $R_{\text{incorrect}}^+ = [1 - \text{AEP}(\eta_d)] / \text{AEP}(\eta_d)$. For the other three storms in this set, the deterministic estimate is less than the threshold. In this case, the ratio of false negatives to true negatives is given by $R_{\text{incorrect}}^- = \text{AEP}(\eta_d) / [1 - \text{AEP}(\eta_d)]$.

Figure 11a shows the ratio of incorrect decisions to correct decisions for deterministic predictions along Track 1 (91.5°W), while Fig. 11b shows the ratio of incorrect decisions to correct decisions for predictions along Track 3 (89.5°W), respectively. Positive values denote cases for which the deterministic forecast exceeded the threshold for action and negative values denote cases for which the deterministic forecast was below the threshold for action. In the case of the negative values, the ratio of incorrect to correct decisions should be interpreted as the absolute value of those values shown in these figures.

Two interesting points are worth noting in Fig. 11a, b. First, the ratio of false positives to correct decisions decreases as forecast time decreases, whereas the ratio of false negatives to true negatives increases as forecast time decreases. Hypothetically, as the uncertainty diminishes, one would expect the value of the forecast to converge toward the deterministic number; however, the error ratio would be expected to converge to 0/0 since

Table 3 Deterministic estimates of storm surges at Point 305 for storms making landfall at 91.5°W (Track a), 90.5°W (Track b), 89.5°W (Track c), and 88.5°W (Track d) for hurricanes with central pressures of 975, 960, and 930 mb before any decay during approach to land, with a radius to maximum winds of 32.78 km

	Track a (m)	Track b (m)	Track c (m)	Track d (m)
975 mb	1.52	2.83	1.43	1.16
960 mb	2.13	3.96	1.98	1.62
930 mb	3.57	5.58	2.93	2.16

Fig. 11 Ratio of incorrect to correct decisions (as defined in text) for storms along Tracks 1 and 3. These tracks approach nearly perpendicularly to the coast and make landfall at **a** 91.5°W, and **b** 89.5°W. *Negative values* denote the cases for which the deterministic surge value is lower than the threshold. *Positive values* denote the cases for which the deterministic value is higher than the threshold

uncertainties affect both the numerator and denominator in equal proportion. However, the one uncertainty that does not diminish is the model uncertainty, since this is independent of forecast time. This source of uncertainty can be shown to shift the expected toward higher values. This shift reduces the ratio of overpredictions and decreases the ratio of underpredictions as observed in these figures. Since the cost of a false negative can be much, much higher in terms of property losses and potential loss of life than the cost of a false positive, this is certainly an important point to consider in the terms of the relative importance of surge model error to forecast model error as forecast time before landfall becomes small.

On the other hand, attempts to introduce extreme conservatism into forecasts may exacerbate problems with false alarms and diminish the emergency response when the forecast is correct. Additionally, the ratio of false positives to correct decisions can have a very significant impact on operational costs not associated with evacuations, so it may not be cost effective for many operations to operate under large amounts of conservatism. One thing that is clear, however, is that, since overprediction typically decreases with time, any system that is more time efficient and/or can be conducted under more adverse conditions will likely save a substantial amount of money over systems which are slow and/or must be operated under very favorable conditions.

To investigate the potential impact of potential bias on the true–false ratios, the best-fit regression equation from Fig. 8 was used to alter values from the surge response function to the equivalent values that the alternative model might have produced. Figure 12 shows the effects that this alteration would have on the ratio of incorrect to correct decisions on an action with a 10-foot threshold at Point 305. Only three lines are shown in this figure, since none of the values of the 920 mb storm along Track 1 or the 920 and 940 mb storms along Track 3 fell below 100 and most were well above 1000. This somewhat obscure form of conservatism definitely shows that models should be validated across the entire range of surge within an area and not just for a set of peaks.

8 Discussion

This paper has investigated the relative impacts of uncertainties in meteorological forecasts and surge models on forecast surges. As stated in the introduction, the eventual goal of this approach is to initiate the development of improved methods for estimating risk, by including the consequences of an event, rather than retaining a focus only on its hazard probabilities. For forecast applications, an extension of Eq. 3 to include even a simple measure of consequences would represent a significant step in this direction. Recalling that the estimated probabilities are calculated for each individual point in a grid, one example might be

$$F(D_{\text{Tot}}) = \iint_{X,Y} \sum_{T_{i,j}} \int \dots \int p(\vec{x}, T, \varepsilon_{w,i}, \varepsilon_m | T_{i,j}, I_f, R_i) D[x_s, y_s, \eta(\Psi) - \eta_*] H[(\eta(\Psi) + \varepsilon_m + \beta_m - \eta_*)] dx_1 \dots dx_n d\varepsilon_{w,i} d\varepsilon_m \Delta T dx_s dy_s \quad (11)$$

where $D[x_s, y_s, \eta(\Psi) - \eta_*]$ is the damage estimate (monetary) for a given event at this x_s, y_s (grid cell), and D_{Tot} is the total damage over the entire area.

In the form in which it is written, Eq. 11 can be used to evaluate planning options as well as to evaluate pre-storm actions in forecasts, since it integrates over all tracks, rather

Fig. 12 Ratio of incorrect to correct decisions (as defined in text) for storms along Tracks 1 and 3, with the bias function from Fig. 8 included. These tracks approach nearly perpendicularly to the coast and make landfall at **a** 91.5°W, and **b** 89.5°W. *Negative values* denote the cases for which the deterministic surge value is lower than the threshold. *Positive values* denote the cases for which the deterministic value is higher than the threshold. Lines for central pressures 920 along Track 1 and 920 and 940 along Track 2 are not shown, since for these cases the values are positive with ratios of greater than 100, i.e., for every time an action is taken there is less than a 1% chance that will be needed

than just variability around a single forecast; however, for planning purposes, this form can be simplified by removing the conditional dependence on forecast storm tracks, i.e.

$$F(D_{\text{tot}}) = \iint_{X,Y} \int \dots \int p(\vec{x}, \varepsilon_m, \beta_m) D[x_s, y_s, \eta(\Psi) - \eta_*] H[(\eta(\Psi) + \varepsilon_m + \beta_m - \eta_*)] dx_1 \dots dx_n d\varepsilon_m dx_s dy_s \quad (12)$$

This methodology could also be modified to include errors in storm speed and angle and could switch to ensemble based uncertainty estimates, but the gist of the results should remain the same.

In retrospect, Figs. 9 and 10 clearly demonstrate three specific aspects of the comparisons made here. First, in most situations, the effect of surge model random error only slightly increases the forecast surge probabilities over results that include the estimated random errors in NWP models; however, there are situations in which these deviations could be important to specific decisions. Second, the effects of bias appear to distort the probabilities considerably more than the effects of random errors. Third, deviations in the probabilities from those considering NWP model errors alone are site specific, since there does not seem to be an obvious universal pattern in the deviations at the different sites considered here.

Equations 11 and 12 provide an idea of how these probabilities might be combined with consequences of “right” and “wrong” decisions to develop an objective concept for

incorporating uncertainty into definitive decisions, such as opening or closing a flood gate or evacuating a particular area of a city. An interesting perspective here is that the decision might not be as categorical as the present evacuation orders, since more information is available on the spatial variation of surge probabilities.

A definite advantage of using pre-calculated surge response functions in the forecast mode is that it allows improved physics and resolution to be captured in the surge model that is presently not possible due to run-time constraints on current surge models used for forecasts. In this mode, the set of runs used to formulate the response surfaces could be executed months before hurricane season, utilizing the best physics and resolution available. An extra benefit to this approach is the possibility of a coordinated approach within several government agencies which need this type of information. For example, FEMA and U.S. Army Corps of Engineers (USACE) both need information on long-term surge probabilities, which could help form the basis for NHC's forecasts. In turn, these agencies could shift more of their resources toward the solution of the consequence side of the risk equation.

An interesting aspect of the application of archived files is that it is equivalent to using a digital (discretized) representation of the surge response surface rather than an analog representation (continuous functions). In fact, this concept is quite similar to the use of digital methods in many other applications today, such as photography and audio recordings. As long as the surge behavior can be specified in terms of piecewise functions, the digital estimates can be made to match analog results quite well. Similar to methods in digital photography, this requires sufficient resolution, which can be augmented by appropriate functional interpolation, to achieve a high level of fidelity. In typical digital photography, this is a two-dimensional problem. In the case of storm surge prediction, the number of degrees of freedom depends on the number of parameters that the NWP models are capable of producing, which presently include the storm track, the storm intensity, information on storm size, storm forward speed, and possibly some characterization of the overall shape of the wind field (peakedness and asymmetry—primarily deduced from persistence). All of these parameters can be included within a pre-calculated set, along with site-specific interpolation functions that should be capable of matching the forecast analog surges quite well. The ability to add uncertainty directly into the estimates provides additional information which can be critical for planning and decision-making.

9 Conclusions

This paper has demonstrated an objective method for estimating site-specific, forecast-time-specific uncertainty in hurricane surge forecasts, using archived storm surge information. It is important to note that the number of *dof* in the archived data in these estimates must be at least equivalent to the number of *dof* available in forecasts. Since the long-term need for accurate information and the need for accurate forecast information both require the highest accuracy possible and quantifiable uncertainty to transition to risk-based decision-making, the incorporation of uncertainty into surge estimates is critical to improving resilience within coastal areas of the USA. Based on findings in this paper, it is clear that information from an unbiased modeling system can offer potentially important advantages for improving objective information for decision-making; however, it will be important to avoid using the same degree of conservatism in information for many users who do not require it, while potentially retaining a reasonable level of conservatism for

evacuation and life-threatening situations. Thus, PSURGE, which focuses on a fixed probability of exceedance for a single storm may be quite appropriate for evacuation planning, although it will inevitably lead to many false positives in its predictions, it is cannot be related to the optimal action for all classes of operations.

Additional conclusions are listed below.

1. Hindcasts of complete existing JPM storm sets undergo rigorous detailed calibration and validation in every geographic area in which they are executed. Their primary goal is to develop information for planning decisions. They utilize state-of-the-art surge models which are not constrained by operational forecasting time limits. Forecast models have consistently been shown in comparisons to produce bias when compared to both measurements and comparisons to these improved models. Whereas forecasters and emergency planners may not be too affected by such conservatism, biases in either direction can significantly affect planning decisions of many other kinds and preclude any serious attempts to objectively optimize decisions.
2. Bathymetry and topography resolution (as is essential for accurately quantifying detailed inundation impacts, particularly in urban areas) cannot be accommodated within available computational resources available for real-time forecasts, but are included in available information from most flood-mapping studies.
3. The approach described here allows uncertainty to be objectively included into forecast decisions. This could play an important role in enabling the investigation of included quantitative cost functions (in terms of gains and losses) related to the impacts of different decisions under differing situations.
4. The same system can be used for estimates of long-term (planning) risks and forecast decision-making needed for (1) evaluating coastal construction projects, (2) coastal land use planning, (3) hazard/risk mapping, and (4) forecasting surges. This could not only help unify two often disjointed efforts, but could allow more resources to be devoted to understanding and quantifying the economic, ecologic, and social aspects of consequences. Given the high cost of preparing and maintaining current bathymetric–topographic data accuracy, such coordination would help reduce redundancy of effort and disagreement in content among official sources of data.
5. An additional potential important benefit from such studies would be the ability to estimate financial and other trade-offs between higher-cost “rapid” evacuation methods and lower-cost “slower” evacuation methods. Analyses here show that uncertainty inherent in these decisions depends strongly on forecast time and geographic location. Methods based on sets of surge maxima do not capture this uncertainty and would be difficult to use for this purpose.

We recognize that this paper is only one step toward being able to accomplish the goals laid out above; however, it is hoped that the method developed here can be combined with improved understanding/quantification of “cost” and social factors related to different decisions in future studies and that this work can serve to motivate additional research toward this end.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bunya S, Dietrich JC, Westerink JJ, Ebersole BA, Smith JM, Atkinson JH, Jensen R, Resio DT, Luettich RA, Dawson C, Cardone VJ, Cox AT, Powell MD, Westerink HJ, Roberts HJ (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* 138:345–377
- Chow S (1971) A study of the wind field in the planetary boundary layer of a moving tropical cyclone. MS thesis, New York University
- Das HS, Jung H, Ebersole B, Wamsley T, Whalin RW (2010) An efficient storm surge forecasting tool for coastal Mississippi. *Coast Eng Proc* 32(1):21
- Dietrich JC, Bunya S, Westerink JJ, Ebersole BA, Smith JM, Atkinson JH, Jensen R, Resio DT, Luettich RA, Dawson C, Cardone VJ, Cox AT, Powell MD, Westerink HJ, Roberts HJ (2010) A high resolution coupled riverine flow, tide, wind, wind wave and storm surge model for Southern Louisiana and Mississippi: part II—synoptic description and analyses of Hurricanes Katrina and Rita. *Mon Weather Rev* 138:378–404
- Galarneau TJ, Hamill TM (2015) Diagnosis of track forecast errors for Tropical Cyclone Rita (2005) using GEFS reforecasts. *Weather Forecast* 30:1334–1354. doi:[10.1175/WAF-D-15-0036](https://doi.org/10.1175/WAF-D-15-0036)
- Hamill TM, Brennan MJ, Brown B, DeMaria M, Rappaport EN, Toth Z (2012) NOAA's future ensemble-based hurricane forecast products. *Bull Am Meteorol Soc* 93(2):209–220
- Holland GJ (1980) An analytic model of the wind and pressure profiles in hurricanes. *Mon Weather Rev* 108:1212–1218
- IPET (2008) Interagency performance evaluation task force. The storm, vol 4. <https://IPET.wes.army.mil>
- IPET (2009) Performance evaluation of the New Orleans and Southeast Louisiana hurricane protection system: final Report. <https://ipet.wes.army.mil/>
- IPET (2010) Interagency performance evaluation task force. Engineering and operational risk and reliability analysis, vol 8, June 2009. <https://IPET.wes.army.mil>
- Irish JL, Resio DT (2010) A hydrodynamics-based surge scale for hurricanes. *Ocean Eng* 37(1):69–81
- Irish JL, Resio DT, Ratcliff JJ (2008) The influence of storm size on hurricane surge. *J Phys Oceanogr* 38(9):2003–2013
- Irish JL, Resio DT, Cialone MA (2009) A surge response function approach to coastal hazards assessment. Part 2: quantification of spatial attributes of response functions. *Nat Hazards* 51:183–205
- Irish JL, Resio DT, Divoky D (2011) Statistical properties of hurricane surge along a coast. *J Geophys Res* 116:C10007
- Jelesnianski CP (1967) Numerical computation of storm surges with bottom stress. *Mon Wea Rev* 95:740–756
- Jelesnianski CP, Chen J, Shaffer WA (1992) SLOSH: sea, lake, and overland surges from hurricanes. NOAA Technical Report, NWS 48. Washington, DC
- Kerr PC, Donahue AS, Westerink JJ, Luettich RA Jr., Zheng LY, Weisberg RH, Huang Y, Wang HV, Teng Y, Forrest DR, Roland A, Haase AT, Kramer AW, Taylor AA, Rhome JR, Feyen JC, Signell RP, Hanson JL, Hope ME, Estes RM, Dominguez RA, Dunbar RP, Semeraro LN, Westerink HJ, Kennedy AB, Smith JM, Powell MD, Cardone VJ, Cox AT (2013) U.S. IOOS coastal and ocean modeling testbed: inter-model evaluation of tides, waves, and hurricane surge in the Gulf of Mexico. *J Geophys Res* 118:5129–5172
- Luettich RA, Wright LD, Signell R, Friedrichs C, Friedrichs M, Harding J, Fennel K, Howlett E, Graves S, Smith E, Crane G, Baltes R (2013) Introduction to special section on the US IOOS coastal and ocean modeling testbed. *J Geophys Res* 118:6319–6328
- Levinson DH, Vickery PJ, Resio DT (2010) A review of the climatological characteristics of land falling Gulf hurricanes for wind, wave, and surge hazard estimation. *J Ocean Eng* 37(1):13–25
- Lin N, Emanuel KA, Smith JA, Vanmarcke E (2010) Risk assessment of hurricane storm surge for New York City. *J Geophys Res* 115:D18121
- Myers VA (1975) Storm tide frequencies on the South Carolina Coast. NOAA Tech. Rep. NWS-16, p 79
- Niedoroda AW, Resio DT, Toro GR, Divoky D, Das HS, Reed CW (2010) Analysis of the coastal Mississippi storm surge hazard. *Ocean Eng* 37(1):82–90
- NRC (National Research Council), (2014) Reducing coastal risk on the east and gulf coasts. National Academies Press, Washington, DC
- Powell MD, Murillo S, Dodge P, Uhlhorn E, Cardone V, Cox A, Otero S, Carrasco N, Annone B, St. Fleur R (2010) Reconstruction of Hurricane Katrina's wind fields for storm surge and wave hindcasting. *Ocean Eng* 37(1):26–36
- Rappaport EN, Franklin JL, Schumacher AB, DeMaria M, Shay LK, Gibney EJ (2010) Tropical cyclone intensity change before US gulf coast landfall. *J Am Meteorol Soc.* doi:[10.1175/2010WAF2222369.1](https://doi.org/10.1175/2010WAF2222369.1)

- Resio DT, Irish JL (2015) Tropical cyclone storm surge risk. *Curr Clim Change Rep* 1:74–84
- Resio DT, Westerink JJ (2008) Modeling the physics of storm surges. *Phys Today* 61(9):33–38
- Resio DT, Irish JL, Cialone MA (2009) A surge response function approach to coastal hazards assessment—part 1: basic concepts. *Nat Hazards* 51:163–182
- Resio DT, Irish JL, Westerink JJ, Powell NJ (2013) The effect of uncertainty on estimates of hurricane surge hazards. *Nat Hazards*. doi:[10.1007/s11069-012-0315-1](https://doi.org/10.1007/s11069-012-0315-1)
- Thompson EF, Cardone VJ (1996) Practical modeling of hurricane surface wind fields. *ASCE J Waterw Port Coast Ocean Eng* 122(4):195–205
- Toro GR, Niedoroda AW, Reed CW, Divoky D (2010) Quadrature-based approach for the efficient evaluation of surge hazard. *Ocean Eng* 37(1):114–124
- Vickery PJ, Skerjil PF, Twisdale LA (2000) Simulation of hurricane risk in the US using empirical track model. *J Struct Eng* 126:1222–1237
- Westerink J, Luettich R, Baptists A, Scheffner N, Farrar P (1992) Tide and storm surge predictions using finite element model. *J Hydraul Eng*. doi:[10.1061/\(ASCE\)0733-9429\(1992\)118:10\(1373\)](https://doi.org/10.1061/(ASCE)0733-9429(1992)118:10(1373))
- Westerink JJ, Luettich RA, Feyen JC, Atkinson JH, Dawson C, Roberts HJ, Powell MD, Dunion JP, Kubatko EJ, Pourtaheri H (2008) A basin-to channel-scale unstructured grid hurricane storm surge-model applied to southern Louisiana. *Mon Weather Rev* 136(3):833–864
- Zhang Z, Krishnamurti TN (1997) Ensemble forecasting of hurricane tracks. *Bull Am Meteorol Soc* 78(12):2785–2795
- Zhang Z, Krishnamurti TN (1999) A perturbation method for hurricane ensemble predictions. *Mon Weather Rev* 127:447–451