Exploring the sensitivity of probabilistic surge estimates to forecast errors

WoongHee Jung, Aikaterini P. Kyprioti, Ehsan Adeli and Alexandros A. Taflanidis
University of Notre Dame, Department of Civil and Environmental Engineering and Earth Sciences

Abstract

Statistical predictions of storm surge are critical for guiding evacuation and emergency response/preparedness decisions during landfalling storms. The probabilistic characteristics of these predictions are formulated by utilizing historical forecast errors to quantify relevant uncertainties in the National Hurricane Center advisories. This ultimately leads to the description of probability distributions quantifying the deviation from the nominal advisory for four different storm features: intensity, size, cross-track variability and along-track variability. Propagation of the uncertainty in these four storm features, serving as input to a numerical model for calculating storm surge, leads to the definition of the statistical surge estimates. This work investigates the application of variance-based global sensitivity analysis (GSA), quantified through the estimation of Sobol' indices, to explore the importance of the forecast errors in the peak storm surge predictions. This GSA can assist in better understanding the impact of the different forecast errors for typical storms, and can also offer important insights for a specific storm, regarding the characteristics that influence the probabilistic surge predictions across its different advisories, as the storm comes closer to landfall. An efficient GSA implementation is presented here to address two key challenges of the specific problem: (i) the need to perform the GSA for a multi-dimensional output, corresponding to the surge for multiple locations within the geographic domain of interest that will be affected by a specific storm, and (ii) the restriction to use only a small number of hydrodynamic numerical simulations, since the associated computational burden of such simulations is significant. For addressing these challenges, dimensionality reduction through Principal Component Analysis (PCA) and a probability-based estimation of the variance of conditional expectations are combined to provide the necessary efficiency in the proposed GSA framework. The development of aggregated importance indices across the entire geographic domain is also discussed, incorporating the importance of the surge for each separate location (within this domain) using a variance-based weighting. This formulation is compared with an alternative, computationally efficient, definition of the aggregated importance, based on the readily available PCA information. A demonstration of this framework's utility considering different historical storms (using National Weather Service advisories and forecast errors for past events) is provided, establishing comparisons across them and across multiple advisories for each storm.

Keywords: storm surge predictions, landfalling storms, probabilistic surge estimation, storm forecast errors, variance-based global sensitivity analysis, Sobol' indices.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

1 Introduction

Recent extremely active hurricane seasons (Blake 2018; Klotzbach et al. 2018; Wood et al. 2019; Pasch et al. 2020), as well as concerns related to the future effects of sea level rise, storm intensification and increased hurricane occurrence rate projections on coastal areas (Dangendorf et al. 2017; Lim et al. 2018; Javeline and Kijewski-Correa 2019), stress the importance of predicting storm-related impacts. The estimation of storm surge – the wind-driven rise of water above astronomical tide levels during landfalling events – has been receiving increased attention in this context, as it has been repeatedly proven to be an especially destructive force for coastal areas (Robertson et al. 2007; Marshall et al. 2019). For evacuation and emergency response management decisions during landfalling events, it is important to develop probabilistic predictions of the storm surge that can incorporate potential errors related to storm forecasts (Taylor and Glahn 2008; Kijewski-Correa et al. 2020). Such errors related to the storm characteristics (describing track, size and intensity) (Chen et al. 2019) impact the surge predictions, and should be explicitly considered in order to support well-informed decisions (Hamill et al. 2012). The standard approach (Taylor and Glahn 2008; Gonzalez and Taylor 2018) for accommodating probabilistic estimates of the anticipated surge in this setting, and the one adopted by the National Weather Service (NWS), is to combine the National Hurricane Center's (NHC) official advisory for the storm current/forecasted features, along with historical errors associated with the forecasts. These historical errors are utilized to specify the probability distribution functions for four different storm features, quantifying ultimately the uncertainty in the NHC advisories (Taylor and Glahn 2008): (a) the cross-track variation; (b) the along-track variation (i.e. the storm forward translational speed); (c) the storm size, represented by the radius of maximum winds; and (d) the storm intensity, represented by the maximum velocity of sustained winds. These features are combined with the official advisory to create storm scenarios to serve as input to a hydrodynamical numerical model to predict the storm surge. In this context, the aforementioned storm features may be considered as the input for predicting the storm surge for a specific NHC advisory. Propagation of the uncertainties in their probabilistic description supports finally the derivation of the desired probabilistic surge estimates.

This work investigates the application of global sensitivity analysis (GSA) to explore the sensitivity of the probabilistic surge estimates to the different forecast errors. The objective of the GSA (Saltelli et al. 2008) is to quantify the importance of the different inputs (storm characteristics) with respect to their impact on the output (predicted surge). This quantification can provide valuable insights regarding the anticipated surge variability, and it can be used to guide various tasks, including dimensionality or uncertainty reduction, as has been shown in similar studies (Saltelli 2002; Fu et al. 2012; Vetter and Taflanidis 2014; Hu and Mahadevan 2016). Specifically, a variance-based decomposition utilizing Sobol' indices (Sobol' 1990) is adopted here for the GSA formulation, since it represents undoubtedly one of the most popular

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

GSA approaches (Saltelli et al. 2008). Sobol' indices quantify the importance of input characteristics both at an individual level, considering exclusively each input, and at an interaction level, where the correlation between inputs is also taken into account. In the context of the examined application here, Sobol' indices will reveal which storm input characteristics create the observed variability in the storm surge predictions for each location of interest within the given geographic domain.

A number of recent studies have stressed and examined some form of sensitivity analysis for storm surge predictions. Some studies (Peng et al. 2004; Irish et al. 2008; Kennedy et al. 2012; Zhang and Li 2019; Ramos-Valle et al. 2020) have focused on parametric investigations that aim to examine how various storm input parameters impact the surge output of numerical hydrodynamic models, providing valuable insights in understanding better storm physics. Other advanced studies (Resio et al. 2017; Sochala et al. 2020; Ayyad et al. 2021) have adopted settings similar to the one examined here, and explicitly considered the importance of forecast errors or implemented some type of global sensitivity analysis. Specifically, Resio et al. (2017) focused on the impact of forecast and numerical errors on the predicted surge for landfalling storms. In that study, each type of error was separately examined, and although investigation offered an important understanding of the overall sensitivity of surge predictions to the factors that influence them, it did not attempt to establish a formal quantification of the contributions of individual types of forecast uncertainties to the experienced storm surge. Studies by Ayyad et al. (2021) and Sochala et al. (2020) extended this effort, and formally considered the estimation of Sobol' sensitivity indices to quantify the individual importance of storm inputs and, in the second study, the additional importance of some numerical model characteristics. Both approaches relied on polynomial chaos expansion (PCE) techniques to estimate Sobol' indices, discussed sensitivity results only for first-order indices, and considered an implementation setting that deviates substantially from the established NWS/NHC framework for describing the variability in landfalling storm forecasts. None of the aforementioned studies attempted to offer any further insights by comparing sensitivity results across different storms and different advisories. Contrary to what has been done thus far, this work: (i) adopts directly the NWS/NHC framework for quantifying probabilistic surge quantities, (ii) considers the estimation of both first- and higher-order Sobol' indices, (iii) establishes a data-driven approach for the estimation of these indices that can be integrated within any existing computational workflow for surge predictions, and (iv) examines a wide range of case-study applications to offer in-depth insights for the surge manifestation, as well as to demonstrate the robustness of the developed GSA approach. These four topics constitute the novel contributions of this paper.

The challenges regarding an efficient GSA implementation in this setting pertain to the need to perform the GSA for a multi-dimensional output, corresponding to the surge for multiple locations (in the order of hundreds of thousands) within the geographic domain the storm will impact, and the restriction to use only

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

a small number of numerical storm surge simulations, since the computational burden of these simulations is typically significant. Furthermore, these simulations cannot be selected specifically for the purpose of accommodating the GSA, meaning that no restrictions can be placed on them having special characteristics (as needed for the numerical GSA estimation). The latter aspect is necessary for accommodating an easier integration with existing computational workflows, like the one currently employed by NWS through the Probabilistic tropical storm Surge (P-Surge) model (Taylor and Glahn 2008; Gonzalez and Taylor 2018). To address these challenges, the estimation of the Sobol' indices is accomplished by combining recent work based on dimensionality reduction through Principal Component Analysis (PCA) (Li et al. 2020) and a probability-based estimation of the variance of conditional expectations (Hu and Mahadevan 2019). PCA restricts the estimation of statistics to a small number of latent outputs, and the probability model-based GSA of Hu and Mahadevan (2019) is then extended to efficiently estimate the necessary statistics for the latent outputs, supporting ultimately the Sobol' index calculation for the original output. The development of aggregated importance indices is also considered, established by combining the results for individual outputs (locations) through a variance-based weighting. This is compared to the indices obtained for the individual latent outputs from the PCA. Different case studies – for different past storm events – are considered and through them the insights and benefits offered by GSA are extensively discussed. The validation of the efficient GSA is also examined by comparing the predicted indices to reference results obtained through a double-loop Monte Carlo integration.

The remainder of the paper is organized as follows. In Section 2 the probabilistic storm surge estimation framework for landfalling events is reviewed. Section 3 discusses the GSA formulation and presents the proposed data-driven approach for the Sobol' index estimation. In Section 4 the case-study examples are presented in detail, and in the final section (Section 5) the case study results are reviewed.

2 Probabilistic surge characterization for landfalling storms

2.1 Storm parametrization

The uncertainty characterization for the definition of probabilistic surge estimates follows the current NWS framework regarding the description of the storm forecasts and the associated errors. This is actually the framework currently deployed by the NWS to provide guidance on evacuation and emergency response/preparedness decisions. A detailed description of the framework is included in (Kyprioti et al. 2021a). Here this description is briefly reviewed. A detailed presentation of the P-Surge model, the software developed by the NWS to support the implementation of this framework for hurricane risk predictions, can be found in (Taylor and Glahn 2008; Gonzalez and Taylor 2018).

The framework for the uncertainty characterization utilizes the NHC advisories during landfalling events, providing information for the past and forecasted storm track, size and intensity characteristics. The

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

track is typically described by the latitude, s_{lat}, and longitude, s_{lon}, of the storm center, the storm size is quantified by the radius of maximum winds, R_{mw}, and the storm intensity by the maximum sustained wind speed, v_w, and/or the pressure difference between the center of the storm and the ambient pressure, DP. Note that functional approximations exist (Jelesnianski and Taylor 1973) that relate R_{mw}, v_w and DP, using, additionally, storm track information. As such, one of these parameters can be determined if the other two characteristics (and the storm track) are known. Part (a) of Figure 1 visualizes this information for advisory 18 of hurricane Katrina, one of the historical storms that will be used later on in the case studies examined in this paper, showing specifically the storm track (s_{lat}, s_{lon}), the 1-min sustained 10-m wind speed v_w, and the pressure difference DP. This advisory information can be subsequently used by a parametric model to provide wind and pressure fields, to then serve as forcing for a hydrodynamical numerical model to simulate the anticipated surge. Examples of such models include the Sea, Lake and Overland Surges from Hurricanes (SLOSH) model (Jelesnianski et al. 1992; Glahn et al. 2009) or the ADvanced CIRCulation (ADCIRC) model (Luettich et al. 1992).

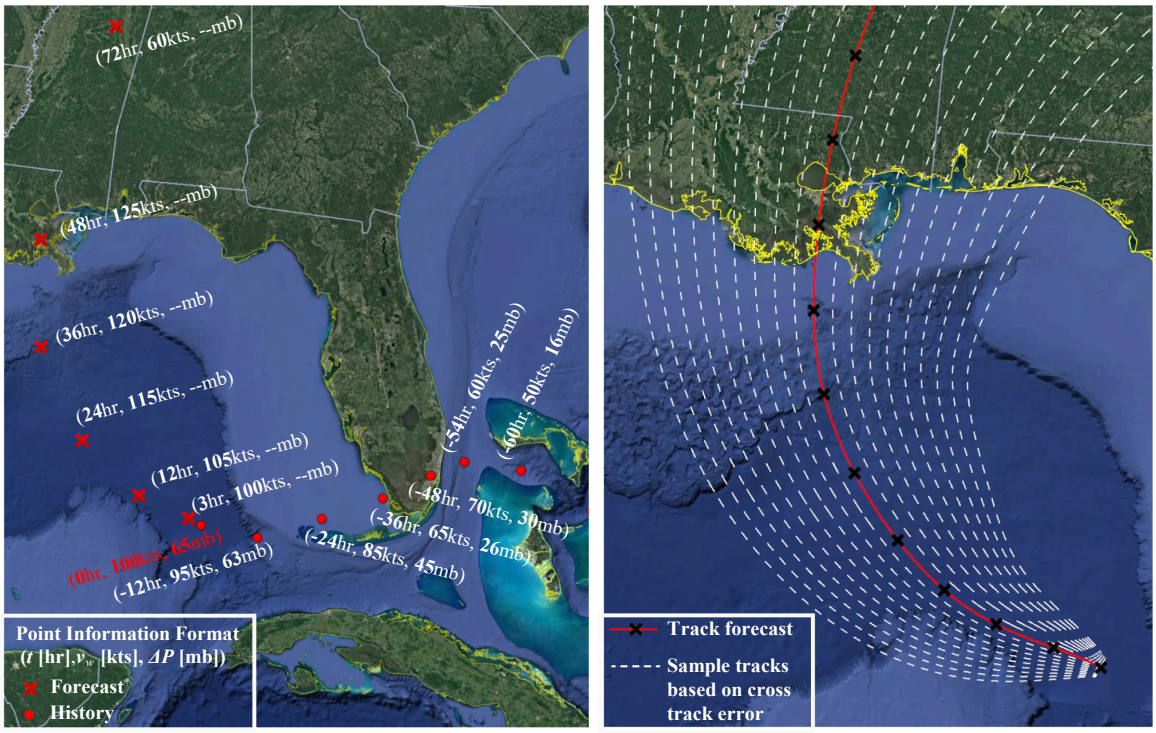
To formalize this setup, let $\mathbf{q}$ denote the four-dimensional vector of features used to describe each storm:

$$\mathbf{q} = \begin{bmatrix} R_{mw} \\ s_{lat} \\ s_{lon} \\ v_w \end{bmatrix} \tag{1}$$

The variation with time of these characteristics will be described using notation $\mathbf{q}(t)$. The NHC forecast provides the nominal (median) predictions for the storm features, $\{\tilde{\mathbf{q}}(t); t > 0\}$, with notation ˜ used above the vector $\mathbf{q}$ to denote the nominal storm characteristics. Since NHC forecasts do not provide information for the future evolution of R_{mw}, a standard practice is to use the current estimate of R_{mw} as the future forecast (Taylor and Glahn 2008). If R_{mw} is not readily available, then current information for v_w and DP can be leveraged (along with the storm track information) to infer this value based on the aforementioned functional approximations that relate these three storm characteristics.

As discussed in the introduction, probabilistic surge predictions are established by accounting for the errors in the nominal storm forecasts. Four different types of errors are distinguished: the size, the intensity and the along and cross position of the storm center relative to the nominal track (Taylor and Glahn 2008). The along track variability impacts the translational velocity of the storm, while the cross track variability impacts the storm track itself, including storm bearing. Part (b) of Figure 1 shows different storm tracks that originate from the cross track variability of the nominal track advisory shown in part (a) of Figure 1. The different storm tracks shown in this figure are obtained through the sampling process currently employed in P-Surge for representing probabilistic storm ensembles (Taylor and Glahn 2008).

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)



(a) Hurricane Katrina (2005)-NHC advisory 18 and prior storm history

(b) Hurricane Katrina (2005)-cross track variability for advisory 18

Figure 1. (a) NHC advisory 18 for hurricane Katrina, showing past (indicated with $t<0$) and forecast (indicated with $t>0$) information for track and intensity; (b) Storm tracks obtained considering cross track variability of the nominal track forecast shown in part (a).

Let $\Delta R_{mw}(t)$ and $\Delta v_w(t)$ denote, respectively, the variability (error) in the size and intensity of the storm parameters and $\Delta s_{along}(t)$ and $\Delta s_{cross}(t)$ denote, respectively, the along and cross track variability of the storm. The vector characterizing the variability of the storm features is then defined as:

$$\Delta\mathbf{q}(t) = \begin{bmatrix} \Delta R_{mw}(t) \\ \Delta s_{cross}(t) \\ \Delta s_{along}(t) \\ \Delta v_w(t) \end{bmatrix} \tag{2}$$

The combination of $\tilde{\mathbf{q}}(t)$ and $\Delta\mathbf{q}(t)$ leads to the definition of $\mathbf{q}(t)$ for the storm having forecast error $\Delta\mathbf{q}(t)$ with respect to the nominal advisory $\tilde{\mathbf{q}}(t)$. Size and intensity are defined directly through their respective variabilities, $R_{mw}(t) = \tilde{R}_{mw}(t) + \Delta R_{mw}(t)$, $v_w(t) = \tilde{v}_w(t) + \Delta v_w(t)$, while for the storm track, the combination of $\Delta s_{along}(t)$ and $\Delta s_{cross}(t)$ jointly influence the updated storm track $s_{lat}(t)$ and $s_{lon}(t)$ in the following way: the nominal track $\{\tilde{s}_{lat}(t),\ \tilde{s}_{lon}(t)\}$ is varied by $\Delta s_{cross}(t)$ perpendicular to its bearing (at each t) to obtain a cross-path modified track; then each point of that modified track is varied along the track by $\Delta s_{along}(t)$ to obtain the final track.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

2.2 Uncertainty description and probabilistic surge estimates

The probabilistic description of $\Delta\mathbf{q}(t)$ is based on a statistical analysis of past forecast errors performed by the NWS. The suggested formulation, and the one adopted in this study, is to treat these errors as statistically independent variables since they represent forecast uncertainties for fundamentally different storm characteristics (Taylor and Glahn 2008; Gonzalez and Taylor 2018). This leads, equivalently to independence in the probability models for $\Delta\mathbf{q}(t)$ (Taylor and Glahn 2008; Gonzalez and Taylor 2018). Note that this does not mean of course that the storm features are independent, simply that the forecast errors around the nominal values are independent. Dependences can certainly exist for the nominal storm features. Furthermore, for each storm feature, a perfect correlation between different times is assumed (Taylor and Glahn 2008). This perfect correlation means that if, for example, the $v_w(t)$ value is larger by one standard deviation than its nominal value at t=24 hr, it will still be larger by one standard deviation than its median value at t=48 hr. Of course, the latter standard deviation, describing the magnitude of the variability, will change for different t, and become larger (larger forecast error) as t increases. These assumptions simplify the uncertainty description, requiring a four-dimensional vector $\mathbf{x}$ to represent the random variables. The exact definition of $\mathbf{x}$ is dependent upon the probability model assumed for the forecast errors, $p(\mathbf{x})$, which is discussed next.

Following (Taylor and Glahn 2008), a Gaussian probability distribution is assumed for $\Delta s_{along}(t)$, $\Delta s_{cross}(t)$ and $\Delta v_w(t)$. Under these assumptions we can define:

$$\Delta q_i(t) = \sigma_i(t) x_i;\ i=2,...,4 \tag{3}$$

where x_i are independent standard Gaussian random variables and $\sigma_i(t)$ is their respective standard deviation, representing the scaling parameters that dictate the size of the forecast errors at different times. The latter ones are selected based on the 5-year mean absolute error statistics $e_i(t)$ as:

$$\sigma_i(t) = e_i(t) / a;\ i=2,...,4 \tag{4}$$

with a taken as 0.7979 (Gonzalez and Taylor 2018). Parts (b)-(d) of Figure 2 show examples for $e_i(t)$ based on the 2005, 2012 and 2020 NHC forecast errors for hurricanes. These years correspond to some of the events that will be utilized later in the case studies. It is evident that as t increases the errors associated with the NHC forecast become larger and that accuracy of the forecasts has improved over the years since average errors in 2020 are substantially smaller than the average errors of 2005.

The last parameter ΔR_{mw}, on the other hand, is treated as a discrete random variable with three possible values representing small, medium and large size storms, taken to correspond to the 15th, 50th and 85th percentiles for the storm size error (Taylor and Glahn 2008). Five different error distributions are defined based on the nominal storm size $\tilde{R}_{mw}(t)$, that are also shown in part (a) of Figure 2. These ultimately define three possible values $\{\Delta R_{mw}^{(r)}(t) \,|\, \tilde{R}_{mw}(t);\ r=-1,0,1\}$, with corresponding probability masses 0.3 (for r=-1),

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

0.4 (for r=0) and 0.3 (for r=1) linked to the 15th percentile, 50th percentile and 85th percentile storm size errors, respectively. To unify their presentation, a transformation to the standard Gaussian space is also adopted for the probabilistic ΔR_{mw} description. This leads ultimately to the relationship:

$$\Delta R_{mw}(t) = \begin{cases} \Delta R_{mw}^{(-1)}(t) \,|\, \tilde{R}_{mw}(t) \text{ if } x_1 < \Phi^{-1}(0.3) \\ \Delta R_{mw}^{(0)}(t) \,|\, \tilde{R}_{mw}(t) \text{ if } \Phi^{-1}(0.3) \le x_1 \le \Phi^{-1}(0.7) \\ \Delta R_{mw}^{(1)}(t) \,|\, \tilde{R}_{mw}(t) \text{ if } x_1 > \Phi^{-1}(0.7) \end{cases} \quad (5)$$

where x_1 is a standard Gaussian variable and $\Phi(.)$ denotes the standard Gaussian cumulative distribution function.

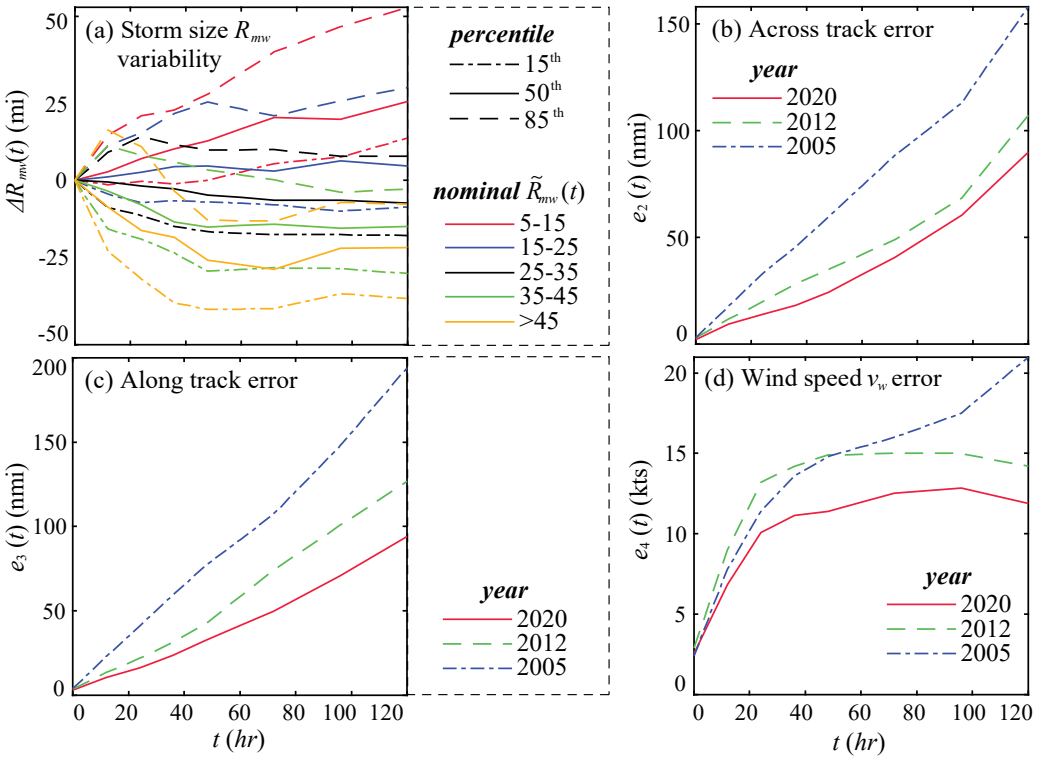


Figure 2. Uncertainty description as a function of time for the storm characteristics: (a) the 15th, 50th and 85th percentiles for the radius of maximum winds for different nominal radius values; (b)–(d) 5-year mean absolute error for the along, cross track variation and the wind speed, respectively, for three different error descriptions [2005, 2012 and 2020 updates of NHC forecast errors].

The vector with components x_i will be denoted herein as **x**. As discussed above, this four-dimensional vector presents the random variables for the uncertainty description. Knowledge of **x** leads to the definition of $\Delta\mathbf{q}(t)$ through Eqs. (3) and (5), which can be then combined with $\tilde{\mathbf{q}}(t)$ to provide the storm features $\mathbf{q}(t)$ through the process described at the end of the previous subsection. The use of $\mathbf{q}(t)$ as input to a numerical model provides subsequently the estimates of the storm surge for the parametrized storm. Notation z_j will be used to denote the peak storm surge at location j of the geographic domain of interest and notation

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

$z_j(\mathbf{x}|\tilde{\mathbf{q}})$ to explicitly denote the dependence of the simulation on the deviation **x** from the nominal storm forecast. To simplify the latter notation, the dependence on $\tilde{\mathbf{q}}$ will be omitted and $z_j(\mathbf{x})$ will be used instead, with the understanding that all results are also dependent upon the nominal forecast $\tilde{\mathbf{q}}$. Therefore, for a given advisory, the variable input for the storm surge prediction model corresponds to vector **x**, which constitutes the model input that the GSA needs to consider. Though statistics of interest include the probability that the surge will exceed different thresholds of interest or the threshold with a specific probability of being exceeded (Hamill et al. 2012; Gonzalez and Taylor 2018), the focus of GSA is directly on the variability of the surge predictions. All these statistics are expressed by propagating the uncertainty from the assumed probability model $p(\mathbf{x})$ to the predicted surge $z_j(\mathbf{x})$.

3 Global sensitivity analysis

For a specific advisory, the n_x-dimensional vector $\mathbf{x} \in \mathbb{R}^{n_x}$ (n_x=4), describing the uncertainty in the storm forecast, defines the input to the numerical simulation, producing the surge estimate $z_j(\mathbf{x})$ for the different geographic locations of interest j=1,…,n_z, where n_z denotes the total number of such locations. We assume that this numerical model is computationally expensive, creating a restriction on the model evaluations that can be considered to perform any probabilistic estimation, and especially a GSA, which as it will be discussed next, introduces a larger degree of complexity. Additionally, we assume that the value of n_z is large, corresponding to over tens of thousand nodes for the surge estimation in typical applications. Let $\mathbf{z} \in \mathbb{R}^{n_z}$ denote the n_z-dimensional output vector, composed of the surge for all examined locations. Recall, as discussed in Section 2.2, that $p(\mathbf{x})$ denotes the probability density function for **x**.

3.1 GSA formulation

The variability of the surge estimates due to forecast errors is quantified through the variance, denoted herein as V^j, expressed for the surge in location j as:

$$Var[z_j] = \int (z_j(\mathbf{x}) - E[z_j])^2 \, p(\mathbf{x}) d\mathbf{x} \tag{6}$$

where Var[.] and E[.] denote the variance and the expectation operators, respectively. The expected value for z_j, representing the expectation operator definition, is:

$$E[z_j] = \int z_j(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \tag{7}$$

Note that Eq. (7) also corresponds to the mean surge under the influence of the forecast errors.

Variance-based GSA considers the decomposition of the total variance $Var[z_j]$ to the contributions coming from each of the inputs x_i as well as from the interactions between all groups of inputs (Sobol 2001).

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

Using superscripts to denote the specific output for the variance representation (in this manuscript the surge at a specific location) and subscripts for the statistics of specific inputs or groups of inputs, the variance decomposition can be expressed as:

$$V^j = \sum_{i=1}^{n} V_i^j + \sum_{i=1}^{n} \sum_{l=i+1}^{n} V_{il}^j + \cdots \tag{8}$$

where V_i^j denotes the contribution from x_i and V_{il}^j denotes the contribution from the interactions between x_i and x_l. The first order contribution is expressed by (Sobol 2001):

$$V_i^j = Var_i[E_{\sim i}[z_j \mid x_i]] \tag{9}$$

where $\mathbf{x}_{\sim i}$ denotes the input vector excluding the x_i input, subscripts i and $\sim i$ are utilized to describe statistics (variance or expectation) with respect to x_i or $\mathbf{x}_{\sim i}$ inputs, respectively, and $E_{\sim i}[z_j \mid x_i]$ denotes the conditional on x_i expectation with respect to input $\mathbf{x}_{\sim i}$. The contribution described by Eq. (9) represents the portion of the total variance that stems from the uncertainty in input x_i. The second-order contribution is given by (Sobol 2001):

$$V_{il}^j = Var_{il}[E_{\sim il}[z_j \mid x_i, x_l]] - V_i^j - V_l^j \tag{10}$$

with similar notation as the one established in Eq. (9), simply extended to multiple inputs. The second-order contribution represents the portion of the total variance that stems strictly from the interactions between inputs x_i and x_l, excluding the portion of the variance that is contributed by each of these inputs individually. Higher-(more than two) order contributions are similarly defined.

Sobol' indices are established by normalizing the variance contributions by the total variance, to define this way the relative contribution. For the jth output and the ith input, the first-order indices are defined as:

$$S_i^j = \frac{Var_i[E_{\sim i}[z_j \mid x_i]]}{Var[z_j]} = \frac{V_i^j}{V^j} \quad \forall i = 1,...,n_x,\ j = 1,...,n_z \tag{11}$$

and quantify the contribution of the ith input to the variability of the jth output without considering its interaction with any of the other inputs. The sensitivity with respect to interactions of multiple inputs can be expressed through higher-order indices. For example, the second-order sensitivity indices between i and l inputs are given by:

$$S_{il}^j = \frac{Var_{il}[E_{\sim il}[z_j \mid x_i, x_l]] - V_i^j - V_l^j}{Var[z_j]} = \frac{V_{il}^j}{V^j} \quad \forall i = 1,...,n_x,\ \ l = i+1,...,n_x,\ j = 1,...,n_z \tag{12}$$

Additionally, one can define the total-effect indices as (Sobol 2001; Saltelli et al. 2010):

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

$$S_{iT}^{j} = 1 - \frac{Var_{\sim i}[E_i[z_j \mid \mathbf{x}_{\sim i}]]}{Var[z_j]} = 1 - \frac{V_{\sim i}^{j}}{V^{j}} \quad \forall i = 1,...,n_x,\ j = 1,...,n_z \tag{13}$$

where subscript notation *T* has been introduced to represent the total-effect indices. These indices quantify the contribution of the *i*th input to the variability of the *j*th output considering the interactions with all possible combinations of the remaining inputs. It is straightforward to show that the total-effect indices represent the sum of all the first-order and higher-order (second, third, …) indices that involve input x_i.

These sensitivity indices provide invaluable information for understanding the influence of the input, in this case of the different types of forecast errors, on the observed surge output (Saltelli et al. 2008). For example, the first-order indices accommodate a relative prioritization of the individual importance of each type of forecast error, revealing what are the important nominal track characteristics, that their change (in future advisories) might drastically impact the expected surge. Higher-order indices answer the same type of question for variation of multiple track characteristics, identifying the correlation between them in the way they impact the surge manifestation. On the other hand, the total-effect indices reveal what type of forecast errors (inputs) are not impactful at all. The uncertainty description for any input that has a total-effect index close to zero can be ignored (the nominal track characteristics can be used), with no impact on the surge predictions. Such an uncertainty reduction, if and when doable, can substantially reduce the computational burden for estimating the expected surge, especially in the current NWS implementation that relies on factorial sampling (Taylor and Glahn 2008; Gonzalez and Taylor 2018). Further discussion on the utility of the sensitivity analysis for storm surge predictions will be provided in the case study examples, after examining an efficient way to estimate them.

In order to simplify the discussion for the different Sobol' index estimation, a unified notation for the variances that are encountered in this estimation is introduced:

$$V_{\mathbf{c}}^{j} = Var_{\mathbf{c}}[E_{\sim \mathbf{c}}[z_j \mid \mathbf{x}_{\mathbf{c}}]] \tag{14}$$

where subscript **c** corresponds to the indices (or index) of inputs that need to be considered, with vector $\mathbf{x}_{\mathbf{c}}$ representing these inputs and $\mathbf{x}_{\sim \mathbf{c}}$ their complement. Dimension of $\mathbf{x}_{\mathbf{c}}$ will be denoted as n_c herein. Note that for the first-order interactions $V_{\mathbf{c}}^{j}$ is directly equal to V_i^j, while for the higher-order interaction $V_{\mathbf{c}}^{j}$ corresponds to the variance with respect to the input $\mathbf{x}_{\mathbf{c}}$ and not to the variances for individual components of $\mathbf{x}_{\mathbf{c}}$ (for example $Var_{il}[E_{\sim il}[z_j \mid x_i, x_l]]$ in the definition of V_{il}^j). For the variance term appearing in Eq. (11) **c**=*i*, for the variance term appearing in Eq. (13) **c**=~*i*, while for the variance term appearing in Eq. (12) **c**=[*i l*]. Each different sensitivity index entails a different definition of **c.** Therefore, the calculation of the Sobol' indices requires the estimation of V^j once, and the estimation of $V_{\mathbf{c}}^{j}$ for the different definitions of **c**, as needed based on the various desired indices we wish to calculate. The estimation of $V_{\mathbf{c}}^{j}$, involving

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

two separate statistical integrals, can be established using a double-loop Monte Carlo Integration (MCI), with the interior loop pertaining to the conditional expectation appearing in Eq. (14), and the exterior to the variance of this expectation. As mentioned above this needs to be separately performed for each examined index, since each pertains to a different definition of **c**. Since the double-loop MCI involves a substantial computational effort, alternative formulations have been considered over the past two decades (Iooss and Lemaître 2015). These formulations include: highly efficient sampling schemes to perform the MCI (Homma and Saltelli 1996; Sobol 2001); approximations using samples from ancillary density functions (Jia and Taflanidis 2016; Li and Mahadevan 2016); approaches that replace the original model with a fast-to-compute surrogate model (Chen et al. 2005; Sudret 2008; Rohmer et al. 2016), even accommodating an entirely analytical estimation of the indices when the surrogate model corresponds to a polynomial chaos expansion (Sudret 2008). Though attractive, these approaches cannot be always integrated easily within existing computational workflows for probabilistic surge estimation, since they might require specialized selection of the storm simulations or might not be able to deal the high-dimensionality of the output. Instead, a completely data-driven approximation is established here for the estimation of $V_{\mathbf{c}}^{j}$, supporting this way an efficient GSA within any probabilistic storm surge prediction setting.

3.2 Efficient GSA implementation

The proposed data-driven estimation of Sobol' indices assumes the availability of a total of k storm surge simulations, for different storm scenarios, represented mathematically by the sample set $\{\mathbf{x}^h; h=1,...,k\}$. The respective output is $\{\mathbf{z}^h; h=1,...,k\}$, with $\mathbf{z}^h=\mathbf{z}(\mathbf{x}^h)$ representing the output vector for model input $\mathbf{x}^h$. These simulations will correspond to the ones employed to perform the probabilistic surge predictions during a landfalling storm (Kyprioti et al. 2021a). Depending on the computational details for facilitating these probabilistic predictions, the samples for **x** may correspond directly to samples from $p(\mathbf{x})$ [direct Monte Carlo implementation setting] or may have been established by some other sampling scheme (Taylor and Glahn 2008; Gonzalez and Taylor 2018). In the latter case, each sample will have a relative probability-weight. Let $\mathbf{X}=[\mathbf{x}^1 \ ... \ \mathbf{x}^k]^T \in \mathbb{R}^{k\times n_x}$ and $\mathbf{Z}=[\mathbf{z}^1 \ ... \ \mathbf{z}^k]^T \in \mathbb{R}^{k\times n_z}$ denote the input and output matrices, respectively, associated with this simulation database. The estimation of Sobol' indices using data [**X**,**Z**] is established by utilizing the generalized Probability-Model GSA (PM-GSA) proposed recently by Hu and Mahadevan (2019). To address the high-dimensionality of the output (multiple locations of interest that the surge needs to be estimated), and avoid the requirement to estimate $V_{\mathbf{c}}^{j}$ for each j explicitly, the dimensionality reduction principles presented in (Li et al. 2020) are adopted. The overall formulation requires extension of the PM-GSA to estimate covariance statistics, and leads to the PCA Probability model-based Sensitivity Analysis (PCA-PSA) algorithm discussed in detail in (J2ung and

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

Taflanidis 2022). Here a brief review is presented, couched within the specific problem under investigation, which is the estimation of the sensitivity indices for storm surge.

Initially, in an effort to reduce the dimensionality of the output, Principal Component Analysis (PCA) is implemented as a dimensionality reduction technique (Jolliffe 2002). Details for the implementation are discussed in Appendix A. PCA in this setting can be interpreted as identifying a low-dimensional vector of latent outputs $\mathbf{y} \in \mathbb{R}^{n_p}$, also called principal components vector, that best explains the variability of the original data $\mathbf{Z}$. Each principal component output $\{y_j; j=1,\ldots,n_p\}$ has an associated importance (also frequently referenced as magnitude) λ_j, representing the portion of data variability explained by the component. This importance introduces a relative prioritization of the components. The dimensionality of the vector $\mathbf{y}$ is $n_p \leq \min(n_z, k-1)$, where min(a,b) denotes the minimum of the two arguments. Instead of using all principal components only the more important ones are retained in order to accommodate larger reduction in the dimension of the output, with the exact number chosen so that these retained components explain a significant portion of the original output variance (say over 99%). In the case studies discussed later, this leads to values of n_p between 30 and 50 showing that PCA can offer a substantial dimensionality reduction for the application examined here. PCA provides ultimately the principal components vector $\mathbf{y}$ and the projection matrix $\mathbf{P} \in \mathbb{R}^{n_z \times n_p}$ facilitating the mapping between $\mathbf{z}$ and $\mathbf{y}$. The observation matrix for the latent outputs for the k available simulations, denoted by $\mathbf{Y} = [\mathbf{y}^1 \ \ldots \ \mathbf{y}^k]^T \in \mathbb{R}^{k \times n_p}$, can be obtained as $\mathbf{Y} = \overline{\mathbf{Z}}\mathbf{P}$, where, as mentioned in Appendix A, $\overline{\mathbf{Z}}$ corresponds to the matrix of original normalized observations (is related to $\mathbf{Z}$). The quantities needed for the estimation of the sensitivity indices for each of the original outputs can be approximated by (Li et al. 2020):

$$\begin{aligned} V^j &\approx \mathbf{P}_{rj} \boldsymbol{\Sigma}_{\mathbf{y}} (\mathbf{P}_{rj})^T \\ V_{\mathbf{c}}^j &\approx \mathbf{P}_{rj} \boldsymbol{\Sigma}_{\mathbf{y}}^{\mathbf{c}} (\mathbf{P}_{rj})^T \end{aligned} \tag{15}$$

where $\mathbf{P}_{rj}$ is the jth row of $\mathbf{P}$, $\boldsymbol{\Sigma}_{\mathbf{y}} = \mathbf{Y}^T\mathbf{Y}/(k-1)$ and $\boldsymbol{\Sigma}_{\mathbf{y}}^{\mathbf{c}}$ is the n_p x n_p covariance matrix for random variables $\{E_{\sim\mathbf{c}}[y_j \mid \mathbf{x}_{\mathbf{c}}]; j=1,...,n_p\}$.

The second step in the GSA estimation is the calculation of $\boldsymbol{\Sigma}_{\mathbf{y}}^{\mathbf{c}}$ using data [$\mathbf{X}$,$\mathbf{Y}$], which is established by extending the probability model-based approach of Hu and Mahadevan (2019). Adopting a Gaussian mixture as a probability model the process can be summarized as follows. The joint probability density function (PDF) of $\mathbf{x}_{\mathbf{c}}$ and y_j (n_c+1 dimensional PDF) is initially approximated through a multivariate Gaussian Mixture Model (GMM) utilizing subset $[\mathbf{X}_{\mathbf{c}}, \mathbf{Y}_j]$ of the original data, where $\mathbf{X}_{\mathbf{c}}$ corresponds to the data matrix for input components $\mathbf{x}_{\mathbf{c}}$ and $\mathbf{Y}_j$ corresponds to the data vector for principal component y_j

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

(*j*th column of **Y**). Utilizing this GMM the statistics $Var_{\mathbf{c}}[E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}]]$ and $Cov_{\mathbf{c}}[E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}], E_{\sim\mathbf{c}}[y_l \mid \mathbf{x_c}]]$, representing the diagonal and off-diagonal terms, respectively, of $\mathbf{\Sigma}_{\mathbf{y}}^{\mathbf{c}}$ can be obtained through the approach discussed in Appendix B. Both quantities require Monte Carlo Integration (MCI) with final expressions given by Eqs. (B.5) and (B.6), respectively.

Combining these concepts, the PCA-PSA implementation for estimating Sobol' sensitivity indices for the storm surge predictions is summarized as follows:

Step 1: Perform PCA for the surge data matrix **Z** and retain the n_p principal components so that ratio r of the explained variance, given by Eq. (A.2) of Appendix A, is greater than a desired threshold r_o (for example 99% or 99.9%). This process provides the projection matrix **P**, the importance λ_j of each principal component y_j, j=1…,n_p, and the data matrix $\mathbf{Y} = \bar{\mathbf{Z}}\mathbf{P}$ for the principal components.

For *each* of the Sobol' indices perform the following steps.

Step 2: Based on the desired index define subset $\mathbf{x_c}$.

Step 3: For each principal component y_j {j=1,…,n_p}, fit a n_c+1 dimensional GMM to the data [$\mathbf{X_c}$, $\mathbf{Y}_j$] where $\mathbf{X_c}$ corresponds to a matrix that includes the columns of **X** designated by $\mathbf{x_c}$ and $\mathbf{Y}_j$ to the jth column of **Y**. Repeat this n_p times, once for each y_j.

Step 4: Estimate the variance statistics for each j using Eq. (B.5) and covariance statistics using Eq. (B.6) with all conditional expectations estimated according to Eqs. (B.3) and (B.4) using the GMM closed form conditional expressions for the respective output y_j.

Step 5: Calculate the required variance statistics for the original output using Eq. (15) and use that to estimate the Sobol' indices of the original output.

The computationally intensive part of the PCA-PSA implementation is the n_c+1 dimensional GMM fit to the data (Step 2), which for each index needs to be performed n_p times, one time for each retained principal component. Most other calculations rely on simple matrix manipulations, and so they can be performed with a negligible computational burden. The Monte Carlo integration (Step 4) also has a very small burden even for large number of MC samples, since the quantities involved have negligible computational complexity on their own. This shows that the dimensionality reduction offers substantial computational benefits, since only n_p GMM fits are required, instead of n_z that would have been the original demand if no PCA was implemented. Therefore, the proposed implementation can be considered computationally efficient even when the storm output is very large, since it relies on PCA to define a lower dimensional latent output space, to accommodate the GSA estimation.

The accuracy of PCA-PSA is impacted by two sources of errors: the PCA approximation, an error that should be negligible if a sufficient number of principal components is retained (Li et al. 2020), and the probability-model based estimation of the conditional statistics, an error that is expected to have a moderate

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

impact (Hu and Mahadevan 2019). The latter type of error is dependent upon the dimensionality of the probability model that is fitted to the data, which, recall, corresponds to a n_c+1 dimensional PDF, and the amount of information available for developing this probability model (number of simulations k). Therefore, it will always be larger for higher-order and total-effect indices (larger value of n_c), compared to first-order indices (n_c=1). Accuracy of the integrated PCA-PSA formulation was examined in detail in (Jung and Taflanidis 2022), demonstrating that the incorporation of the dimensionality reduction within the probability model-based GSA does not create any additional sources of error. As long as a sufficient number of principal components is utilized, the overall PCA-PSA accuracy is dependent only upon the challenges associated with the GMM approximation for the conditional statistics. In the case studies discussed later, some further validation results will be briefly discussed.

3.3 Aggregated importance indicators across the outputs

The implementation discussed in Section 3.2 supports a computationally efficient estimation of (each of) the Sobol' indices for each output, allowing us to identify the sensitivity and the relative importance of the forecast errors for each location of interest (output). The spatial variability of the sensitivity trends, obtained by estimating the indices for the different outputs, can reveal some interesting insights, as will be also demonstrated in the case study examples presented later. Beyond the spatial varying sensitivity results, though, some averaging of these results across the entire geographic domain of interest might be also of interest. The latter provides aggregated importance indicators for each input (forecast error), revealing their influence on the expected surge across the entire coastal region that the storm is expected to impact.

Averaging of the results for individual outputs z_j, to obtain the desired aggregated importance indicators, should incorporate appropriate weights. This can be justified by the fact that locations with small experienced surge variation cannot be given equal weight to locations for which the surge variability is significant. Since the proposed GSA is variance-based, these weights should be evidently selected based on the surge variance for each location. Note that identical definitions have been previously proposed (Rohmer et al. 2016) for the quantification of aggregated sensitivity. For the first-order indices, the aggregated sensitivity index, denoted S_i for the x_i input, using as weighting the total variance for location j, is then obtained as:

$$S_i = \frac{\sum_{j=1}^{n_z} V^j S_i^j}{\sum_{j=1}^{n_z} V^j} \tag{16}$$

with similar expressions holding for all other higher-order indices.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

An alternative formulation can be established by using directly the sensitivity indices for the retained principal components. In this case, each component y_j is weighted by its relative importance λ_j. The PCA-based aggregated first-order index corresponds to:

$$\underline{S}_i = \frac{\sum_{j=1}^{n_p} \lambda_j \underline{S}_i^j}{\sum_{j=1}^{n_p} \lambda_j} \tag{17}$$

where $\underline{S}_i^j$ is the first-order sensitivity index for the jth principal component, given by:

$$\underline{S}_i^j = \frac{Var_i[E_{\sim i}[y_j \mid x_i]]}{\mathbf{Y}_j^T \mathbf{Y}_j / (k-1)} \tag{18}$$

and is easily calculated using existing information [Eq. (B.5)]. If only the aggregated importance indicators are of interest, and not the geographical variation of the sensitivity as well, then the PCA-based importance of Eq. (17) provides a simplified estimation, as it does not require the transformation of the sensitivity information to the original output space through Eq. (15), or the estimation of the covariance statistics of Eq. (B.6). The important question is of course if there are differences between the aggregated importance indices that are identified through the use of Eq. (16) (original output) or Eq. (17) (principal components), a topic that will be further explored in the case studies examined later in this manuscript.

4 Illustrative case studies overview

4.1 Case study characteristics

Five different historical storms are examined as illustrative case studies, corresponding to: hurricane Katrina (2005), hurricane Gustav (2008), hurricane Irene (2011), super storm Sandy (2012) and hurricane Arthur (2014). These storms are chosen to offer a comprehensive GSA demonstration examining both landfalling (Katrina, Gustav and Sandy) and bypassing (Irene and Arthur) storms, across different geographic regions, that include both the Gulf of Mexico (Katrina and Gustav) and the North Atlantic (Irene, Sandy and Arthur). At least two storms are selected in each subcategory (bypassing/landfalling or Gulf/North Atlantic), to also facilitate comparisons between them. For each storm, two different NHC advisories will be utilized: (i) for Katrina, advisories 23 and 18 are used, corresponding, respectively, roughly to 24 hr and 48 hr before the storm makes landfall, (ii) for Gustav, advisories 29 and 26 are used, corresponding, respectively, roughly to 24 hr and 48 hr before the storm makes landfall, (iii) for Sandy advisories 26 and 20 are used, corresponding, respectively, roughly to 36 hr and 72 hr before landfall, (iv) for Irene advisories 30 and 26 are used, corresponding, respectively, roughly to 24 hr and 48 hr before the storm bypasses New York and starts gradually loosing strength, while (v) for Arthur advisories 14 and 10

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

are used, corresponding, respectively, roughly to 24 hr and 48 hr before the storm, similarly to Irene, bypasses New York and starts gradually losing strength. To simplify the presentation, the earlier advisory (smaller advisory indexing number), representing the time further from landfall, will be denoted by A_1 and the later advisory (larger advisory number), representing time closer to landfall, will be denoted by A_2. The tracks for all these storms and advisories are shown in Figure 3.

The consideration of different advisories allows us to examine how the GSA results evolve as the nominal track characteristics change. It also supports an investigation of the impact of the magnitude of the forecast errors on the GSA; advisory A_1 represents ultimately a case with a larger relative influence of forecast uncertainties, corresponding -according to Figure 2- to storm characteristics (track, intensity and size) that have a larger variability by the time the storm reaches landfall -when the storm surge is expected to reach its peak value across a larger part of the domain. This is also demonstrated by the cross track variability shown in Figure 3. The comparison across the advisories is influenced though, not only by the size of the forecast uncertainties, but also by the differences in the nominal advisories. In order to investigate more clearly the influence of the forecast uncertainties, two different scenarios are examined for the uncertainty characterization: (a) the storm occurring in its actual year, denoted as U_1; and (b) the storm occurring in 2020, denoted as U_2. In each case, the NWS forecast errors of that specific year are utilized. Since, as evident from Figure 2, the forecast error magnitude for the track and the intensity (but not for the storm size) are decreasing across the years, smaller uncertainties will be considered for these storm features for uncertainty characterization scenario U_2. For consistency, the scaling parameter a in the forecast error definition of Eq. (4) is taken as 0.7979 according to the current recommendations (Gonzalez and Taylor 2018), even though for some of the years that will be examined (for example for 2005) a smaller value of a=0.6745 was utilized (at that point) by the NWS. Figure 3 shows for each storm advisory the nominal NHC forecast, as well as the tracks that correspond to one standard deviation from the nominal track based on the identified cross track variability for scenario U_1.

According to the current NWS probabilistic framework (Gonzalez and Taylor 2018), for each advisory the information available is assumed to correspond to: the storm track, $DP(t)$ and $v_w(t)$ for the past hurricane history (t<0), and the storm track and $v_w(t)$ for future forecasts (t>0). Based on the provided DP and v_w, the storm size R_{mw} is estimated for the hurricane history. The current estimate for t=0, $R_{mw}(0)$, is kept constant for future predictions (t>0), representing the nominal storm size forecast for t>0. For the relationship among v_w, DP and R_{mw} results from the study (Knaff and Zehr 2007) are utilized. The same relationships are used to provide the $DP(t)$ predictions for t>0, based on the values of $v_w(t)$ and $R_{mw}(t)$ for the respective time t.

The output **z** considered in all comparisons corresponds to the peak storm surge without considering any tidal effects. The numerical model utilized for the surge predictions corresponds to a surrogate model of ADCIRC (ADvanced CIRrculation Model for Shelves, Coastal Seas, and Estuaries) (Luettich et al.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

1992). For the implementation discussed here, this approximation is considered to correspond to the exact numerical model, and within the PCA-PSA framework is only utilized to perform the k simulations in order to obtain the sample set [**X**, **Z**]. Its use instead of ADCIRC is necessitated by the extensive case studies considered and the need to estimate reference solutions for the Sobol' indices for some of these cases, to validate the PCA-PSA accuracy. Details for the development of the surrogate models for the North Atlantic, (used for hurricanes Irene, Arthur and Sandy), and for the Louisiana region, (used for hurricanes Katrina and Gustav) can be found in (Kyprioti et al. 2021b) and (Jia et al. 2016), respectively. The surrogate models were developed using databases of ADCIRC simulations for the two aforementioned regions of interest and are available by the U.S Army Corps of Engineers through their Coastal Hazards System (Nadal-Caraballo et al. 2020). The accuracy of the established surrogate models is very high, with correlation coefficient with respect to the original database of over 98.5%, providing a very high degree of confidence for their use within the case studies considered here. The part of the ADCIRC grid within the domain of impact for these storms includes n_z=1,860,021 nodes for the North Atlantic case studies and n_z =1,552,341 for the New Orleans case studies. These numbers showcase the large dimensionality of the examined output, stressing the GSA-related challenges. The portion of these nodes that correspond to a depth less than 5 m or 10 m, is 78.2% (for 5m) and 90.4% (for 10m) for the North Atlantic case studies, and 83.4% for (5m) and 92.8% (for 10m) for the Louisiana case studies. This shows that a large portion of the examined outputs corresponds to near-shore nodes, belonging to regions with potentially complex geomorphologies, creating interesting GSA trends as will be discussed later.

The omission of tides is an assumption that was made in past studies that have examined storm surge GSA applications (Resio et al. 2017; Sochala et al. 2020; Ayyad et al. 2021), as it accommodates a focus on the storm physics, and for our study is additionally dictated by the need to use surrogate models in order to accommodate the extensive validation aspects that will be investigated here. This omission, nevertheless, needs to be acknowledged as a limitation in our work. Although tides are a deterministic process, that is not directly influenced by the forecast errors, there is a secondary effect related to the timing of the peak storm surge and how the latter is coupled with the tidal phases to produce the resultant peak water elevation (tides + storm surge). Therefore, the influence of the forecast error of the along track variation on the storm forward speed and consequently on the temporal variation of the storm surge (affecting ultimately the coupling with the tidal phases), is expected to be impacted by an additional consideration of tides. As such, results regarding the forecast errors for the along track variability presented in the case studies should be treated as a lower bound for the sensitivity analysis if tidal effects were to be considered. Of course the degree that the inclusion of tides will impact the sensitivity to the along track variability forecast errors is dependent upon the tide magnitude, something that exhibits both geographical and seasonal variability.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

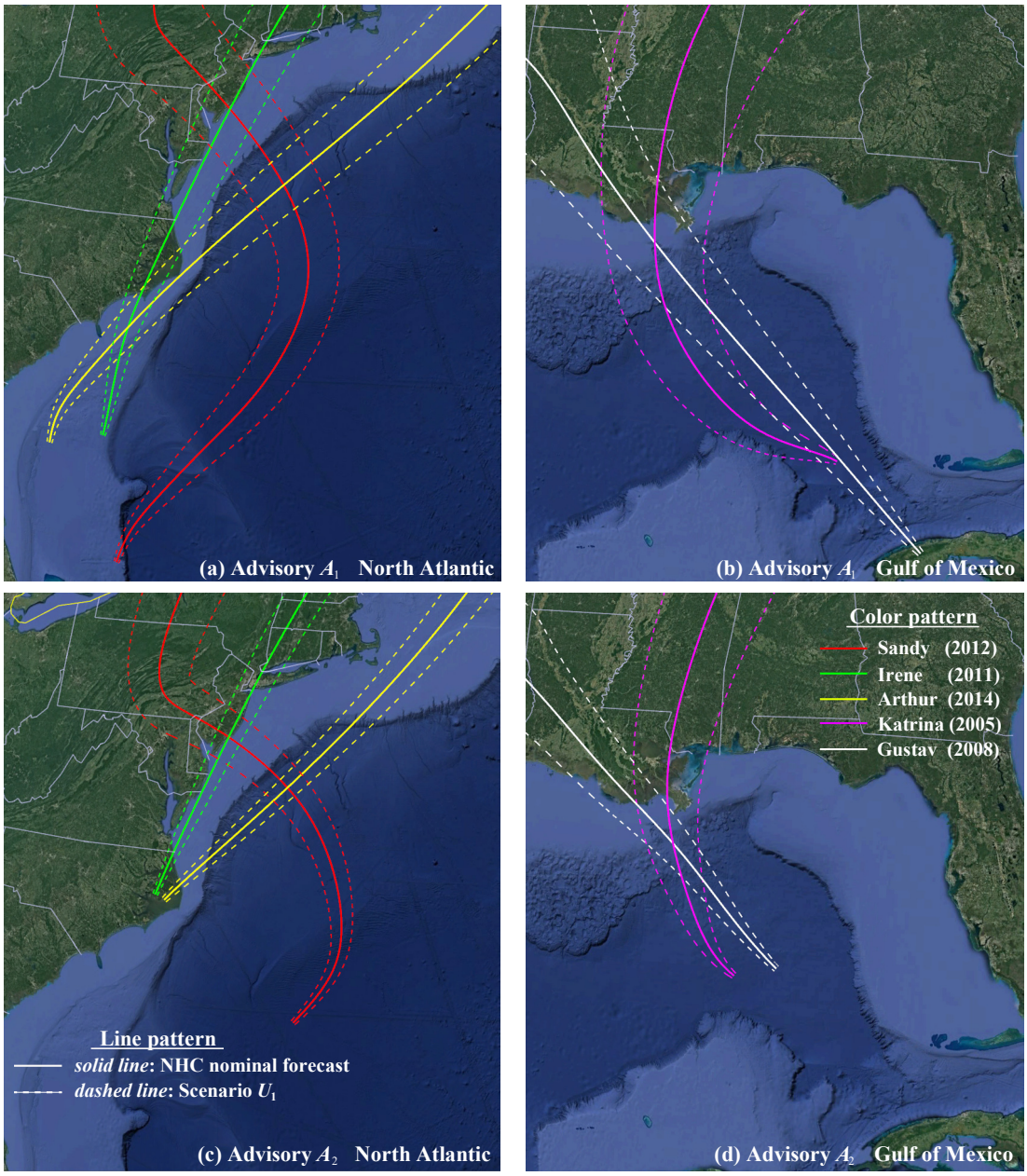


Figure 3. Storm tracks for the case study storms in the North Atlantic [left column], corresponding to hurricanes Irene (2011), Sandy (2012) and Arthur (2014), and the Gulf [right column], corresponding to hurricanes Katrina (2005) and Gustav (2008), for two different advisories in each case [A_1: top row, A_2: bottom row]. NHC nominal forecast (solid line) as well as the track corresponding to one standard deviation for the cross track variability based on the forecast errors of the year of the storm (uncertainty characterization scenario U_1 in dashed line).

4.2 GSA implementation computational details

For the PCA-PSA implementation the number of storm simulations is chosen as k=500 with samples obtained directly from $p(\mathbf{x})$. This number corresponds to a typical number of storm simulations currently employed in probabilistic surge forecasting (Kyprioti et al. 2021a). The threshold r_o for selection of PCA components is chosen as 99.9% in order to minimize the PCA truncation error, leading to n_p in the range of

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

20-50 across the different storms and advisories. Therefore, the PCA offers a substantial reduction with respect to the dimension of the output, and significant computational benefits for the GSA implementation. The GMM fit is performed using the Expectation-Maximization algorithm (Bishop 2006) with an adaptive selection of the number of mixture components based on the minimization of the Bayesian Information Criterion (BIC) (McNicholas and Murphy 2008). The number of samples for the MCI is chosen large, equal to N=5000.

To establish validation of the PCA-PSA for some of the case studies, the estimation of the first-order and total-effect indices using an efficient numerical implementation of the double-loop MCI is considered (Homma and Saltelli 1996; Sobol 2001). As a validation metric, the average accuracy across the output (Sobol' index at different locations) is quantified using the normalized root mean squared error, denoted as *nrmse* herein. If S^j denotes the reference results for the Sobol' index of interest for output z_j obtained through the double-loop MCI and $\tilde{S}^j$ the approximation of the same index using PCA-PSA, the *nrmse* is given by:

$$nrmse = \frac{\sqrt{\frac{1}{n_z}\sum_{j=1}^{n_z}(S^j - \tilde{S}^j)^2}}{\max_{j=1,\ldots,n_z}(S^j) - \min_{j=1,\ldots,n_z}(S^j)} \tag{19}$$

5 Illustrative case studies: results and discussion

The output sensitivity to each of the four forecast errors is calculated for each of the case study examples using the PCA-PSA algorithm. To simplify the discussions, all results will be presented with respect to the variation of the underlying storm feature, ΔR_{mw}, Δs_{cross}, Δs_{along} and Δv_w, instead of the respective variable x_i which, as discussed in Section 2, represents the equivalent variable input in the storm simulations.

5.1 PCA-PSA validation for probabilistic surge estimation GSA

Validation of PCA-PSA is established first for all storms looking at advisory A_1 and uncertainty characterization scenario U_1. Validation results averaged across the entire output domain are presented for the first-order and total-effect indices in Table 1 for each of the four inputs. Results demonstrate that for probabilistic surge estimation applications, very high-accuracy is established by the PCA-PSA implementation for the first-order indices, but this accuracy is reduced for the total-effect indices. This agrees with the trends identified in (Jung and Taflanidis 2022) and should be attributed, as also mentioned earlier, to the higher dimensionality of the PDF fit that is required for these indices (Hu and Mahadevan 2019). In the application examined here, with a 4-dimensional input, first-order indices entail a GMM fit for a 2-dimensional PDF, while total-effect indices for a 4-dimensional PDF. The accuracy for total-effect indices can be improved if a larger value is utilized for k (storm surge simulations), but as discussed earlier

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

this would be impractical since the intention here is to integrate a data-driven GSA within the existing computational workflows for probabilistic surge estimation, and the typical value used by these workflows (Kyprioti et al. 2021a) is in the range of 500 total numerical simulations.

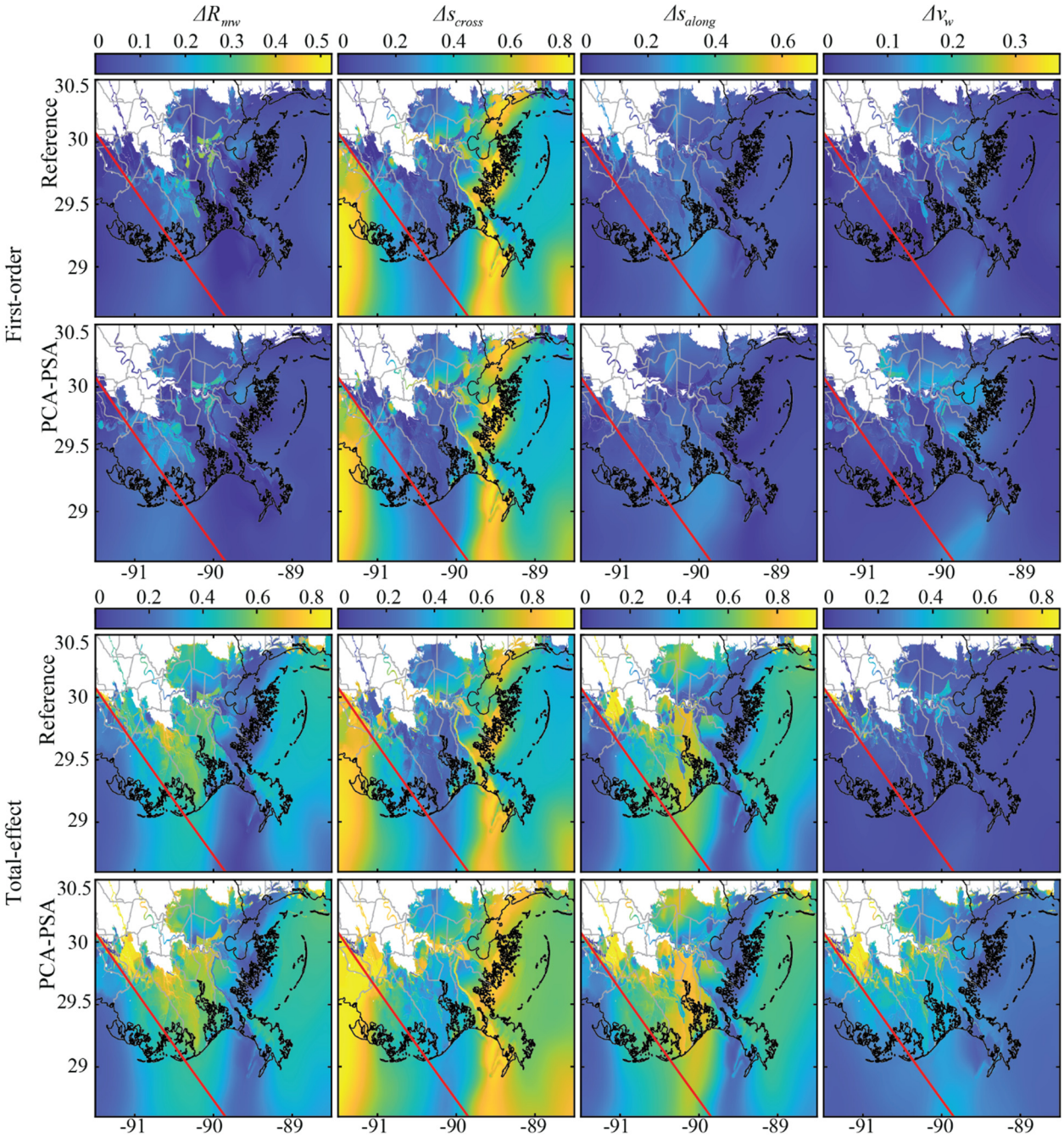


Figure 4. Comparison of GSA estimates over the spatial domain of interest for hurricane Gustav (2008) for advisory A_1 and uncertainty characterization U_1. The first two rows present the first-order indices, comparing the reference estimates [first row] to the PCA-PSA estimates [second row], while the last two rows present the total-effect indices establishing similar comparisons. The red line in all subplots corresponds to the forecasted storm track.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

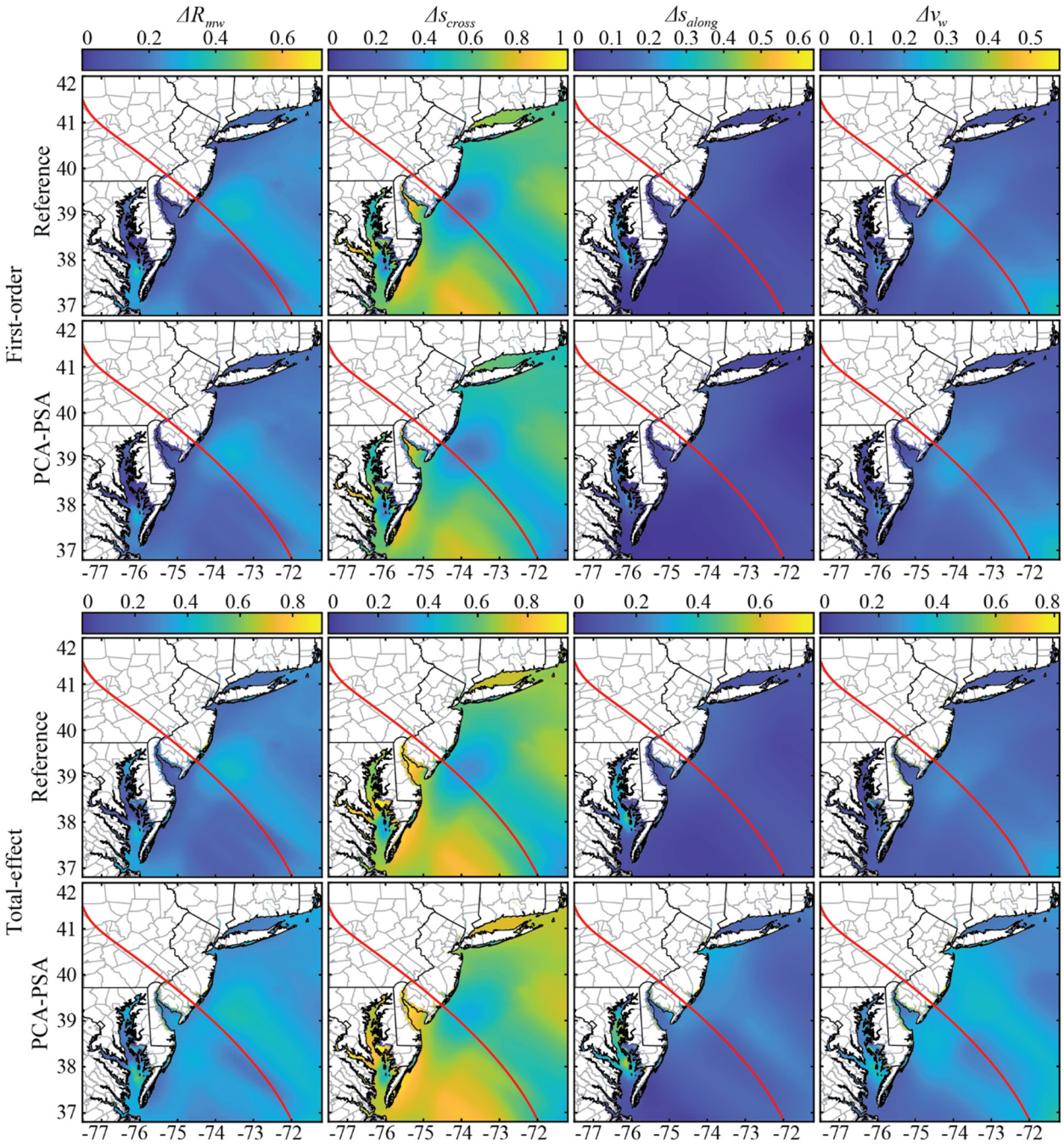


Figure 5. Comparison of GSA estimates over the spatial domain of interest for super storm Sandy (2012) for advisory A_1 and uncertainty characterization U_1. The first two rows present the first-order indices, comparing the reference estimates [first row] to the PCA-PSA estimates [second row] while the last two rows present the total-effect indices establishing similar comparisons. The red line in all subplots corresponds to the forecasted storm track.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

Table 1. Validation of PCA-PSA results, expressed through the normalized root mean squared error, for all storms for advisory A_1 and uncertainty characterization U_1.

Storm	First-order indices				Total-effect indices			
	ΔR_{mw}	Δs_{cross}	Δs_{along}	Δv_w	ΔR_{mw}	Δs_{cross}	Δs_{along}	Δv_w
Katrina	0.029	0.093	0.128	0.036	0.175	0.057	0.183	0.203
Gustav	0.040	0.033	0.069	0.012	0.134	0.214	0.124	0.247
Irene	0.043	0.033	0.030	0.028	0.168	0.176	0.208	0.189
Sandy	0.042	0.048	0.040	0.032	0.220	0.203	0.256	0.254
Arthur	0.085	0.048	0.104	0.012	0.244	0.255	0.249	0.256

The bigger question related to the reduced accuracy for the total-effect index estimates, is how this reduced accuracy might impact GSA insights. Figures 4 and 5 present the spatial distribution of the first-order and total-effect indices over the domain of interest, for hurricane Gustav (Figure 4) and super storm Sandy (Figure 5). In all subplots, a red line corresponds to the forecasted storm track. Results for the other storm case studies are similar and not presented due to space constraints. The reference estimates are compared to the estimates obtained through PCA-PSA in both figures. Results clearly show that the spatial distribution for both the first-order and total-effect indices is practically identical for both approaches, meaning that GSA insights will not be impacted by any (in this case small) loss of accuracy when the corresponding sensitivity indices are estimated through PCA-PSA. Validation that will be presented later, in section 5.3, with respect to the aggregated indices will similarly show very good agreement between the reference and the PCA-PSA estimates. Overall these comparisons showcase that the proposed PCA-PSA implementation can support an accurate and highly efficient estimation of global sensitivity indices for probabilistic storm surge estimation applications.

5.2 GSA results for different storms and advisories

The validated PCA-PSA computational framework is now applied to all the storm case studies for both advisories and uncertainty characterizations. This section focuses on the spatial distribution of the sensitivity indices, with the following section (Section 5.3) discussing the aggregated importance indicators. Also, the discussion focuses on first-order and total-effect indices, representing as discussed earlier the two most important and commonly used GSA measures, though some results will be also briefly presented for higher-order interaction indices.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

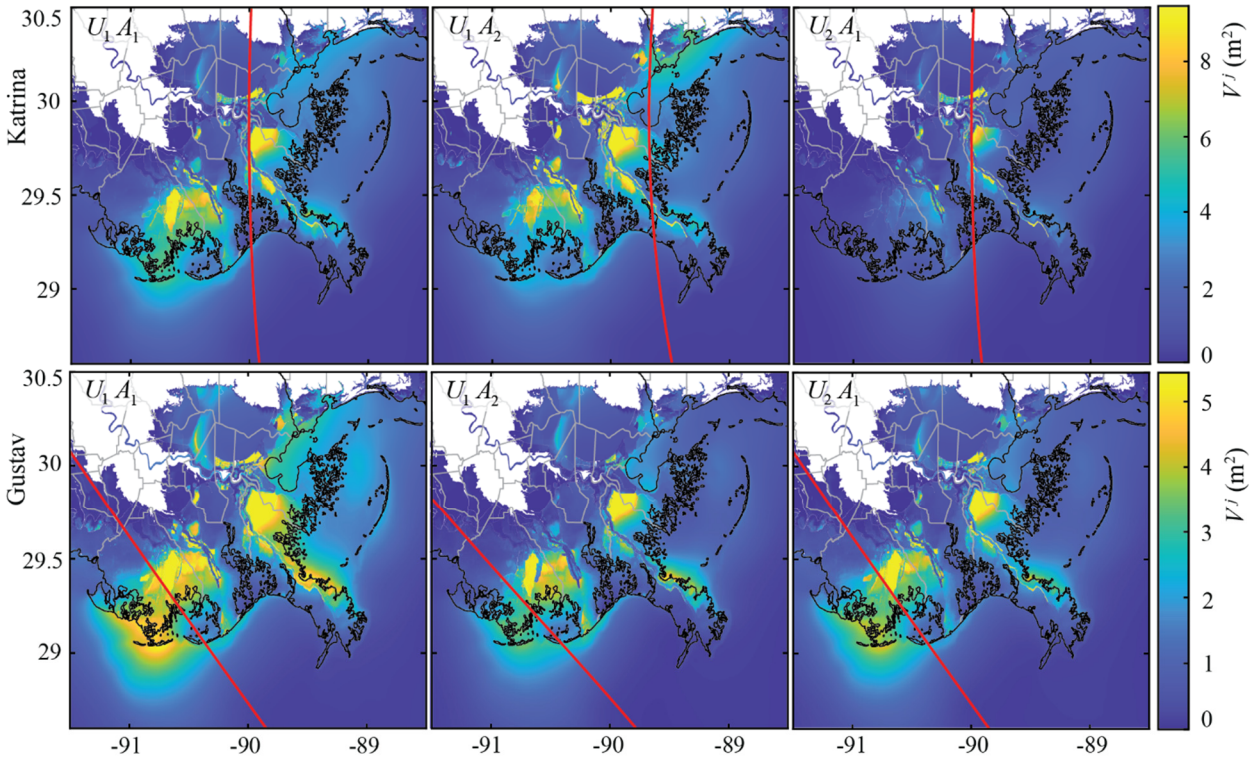


Figure 6. Distribution of the storm surge variance over the spatial domain of interest, for storms in the Gulf, corresponding to hurricanes Katrina (2005) [top row] and Gustav (2008) [bottom row]. Each column presents a different combination of advisory and uncertainty characterization cases.

To better frame the discussions, Figures 6 and 7 present the distribution of the surge variance over the domains of interest, separately for the storms in the Gulf (Figure 6) and the storms in the North Atlantic (Figure 7). In all subplots, herein, the forecasted storm track is depicted with a red line. The variance shown in these figures quantifies the relative importance of the different regions within the domain of interest; regions with larger variance represent the parts of the domain that the probabilistic surge estimates will exhibit larger degree of variability and will be ultimately impacted substantially more by the relative sensitivity to the different storm features. It also represents the weights that will be used for the estimation of the aggregated importance indicators according to the discussions in Section 3.3. Results are presented for both advisories for uncertainty characterization U_1, and for advisory A_1 for uncertainty characterization U_2. The variance distribution is impacted, as expected, by the interaction between the storm evolution and local coastal features. Larger variability is observed when the size of the storm forecast errors, and therefore the variability in the probabilistic description of the storm features, increases. This is evident when comparing the results corresponding to uncertainty characterizations U_1 to U_2, while it can be also observed in the comparisons between advisories A_1 and A_2, since the former, representing a time further from landfall, will be typically characterized by larger forecast errors at the time peak surge is manifested across the bigger parts of the domain the storm impacts. Note that the change of the nominal storm features also greatly

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

influences results between advisories, so any differences observed between them cannot be solely attributed to the impact of the forecast error size. For example, this dependence contributes to larger variability for superstorm Sandy in substantial parts of the domain of interest when advisory A_2 is compared to advisory A_1.

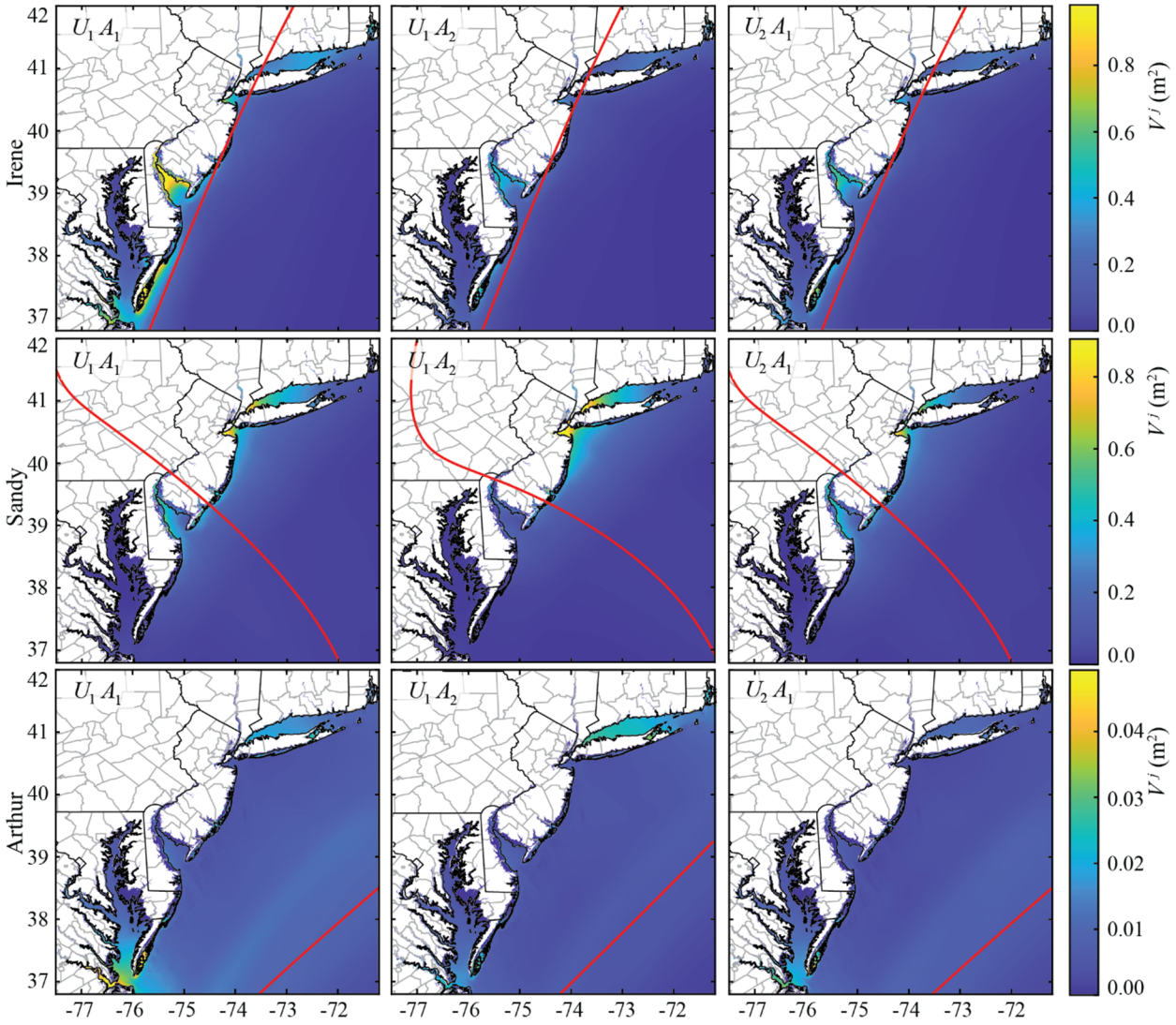


Figure 7. Distribution of the storm surge variance over the spatial domain of interest, for storms in the North Atlantic, corresponding to hurricanes Irene (2011) [top row], Sandy (2012) [middle row] and Arthur (2014) [bottom row]. Each column presents a different combination of advisory and uncertainty characterization cases.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

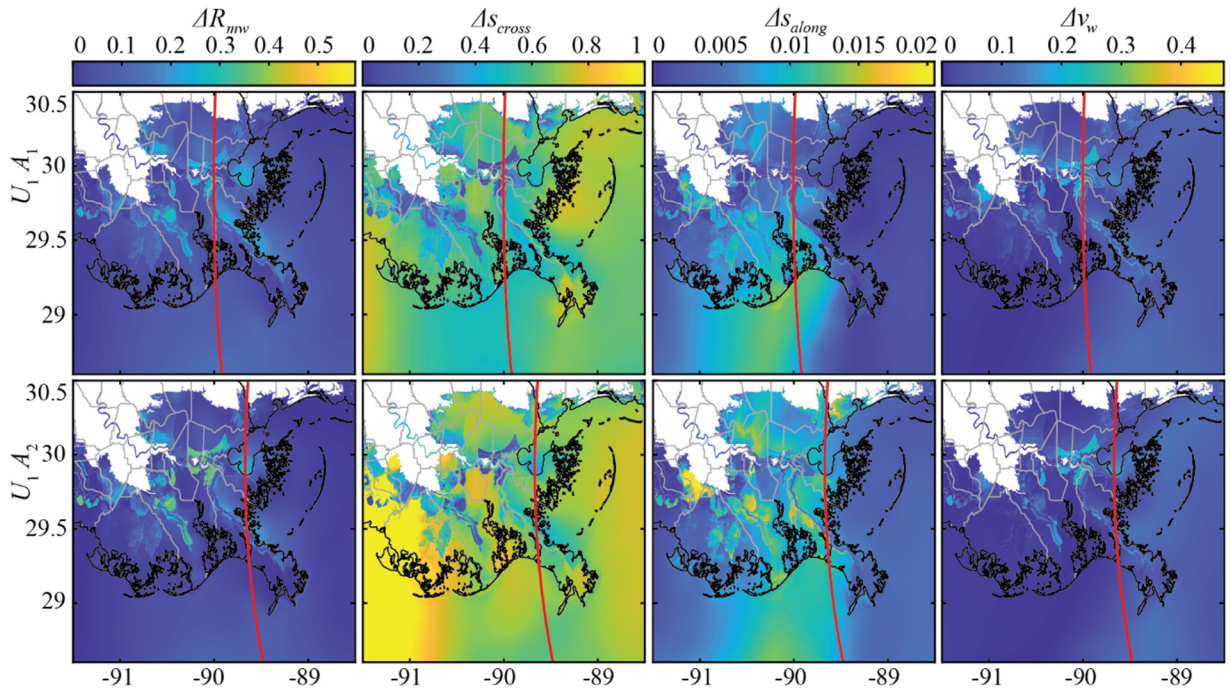


Figure 8. Distribution of the first-order sensitivity indices over the spatial domain of interest for hurricane Katrina (2005) for advisories A_1 [top row] and A_2 [bottom row] with uncertainty characterization U_1.

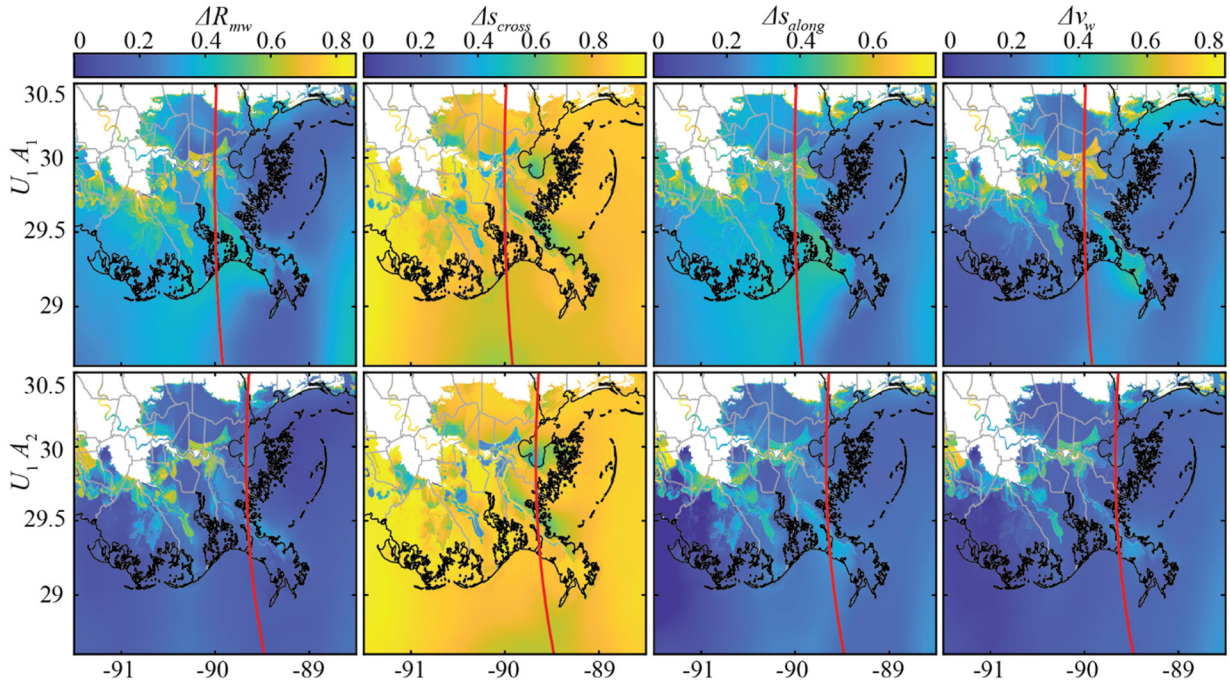


Figure 9. Distribution of the total-effect sensitivity indices over the spatial domain of interest for hurricane Katrina (2005) for advisories A_1 [top row] and A_2 [bottom row] with uncertainty characterization U_1.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

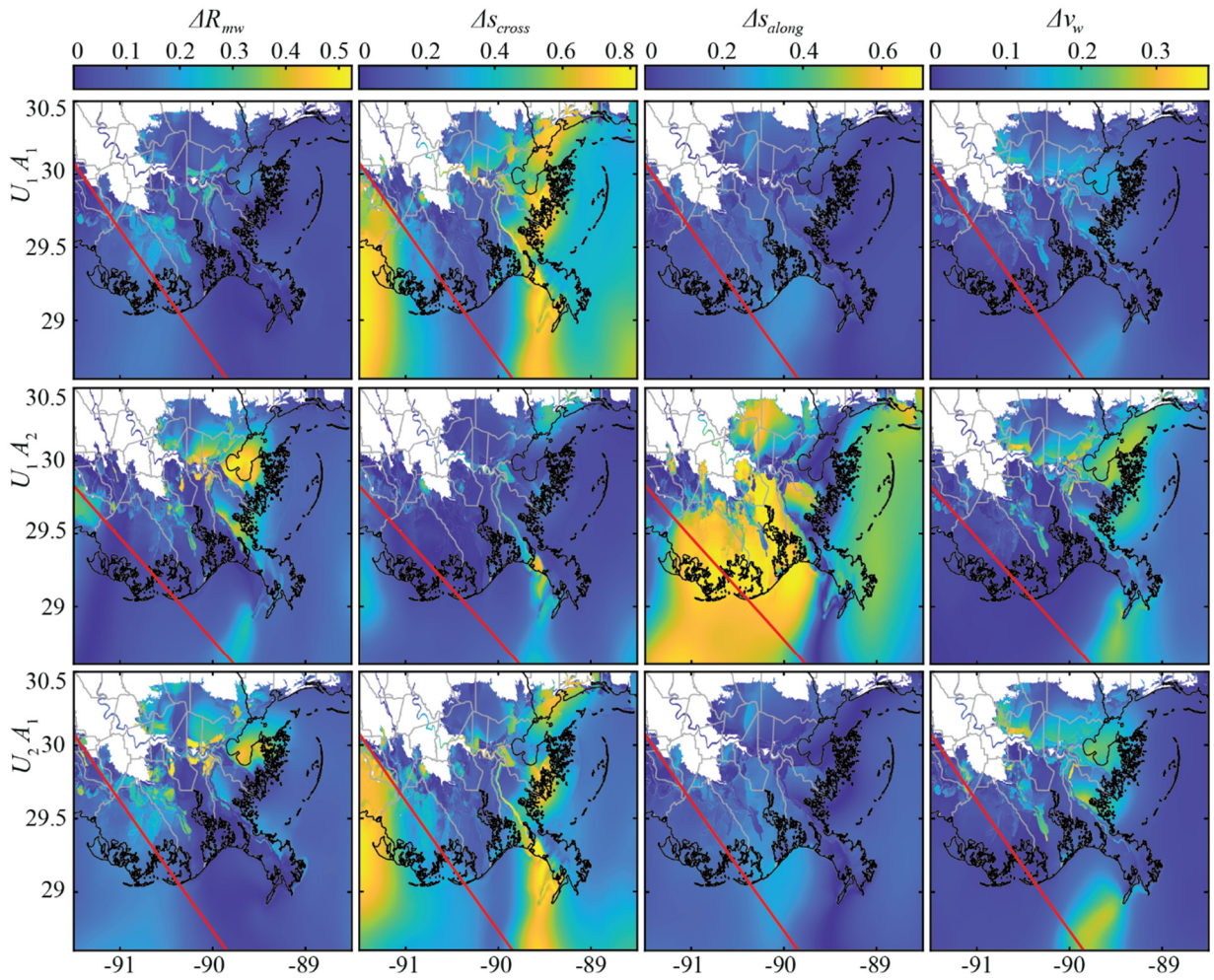


Figure 10. Distribution of the first-order sensitivity indices over the spatial domain of interest for hurricane Gustav (2008) for a combination of advisory and uncertainty characterization cases, corresponding to U_1A_1 [top row], U_1A_2 [middle row] and U_2A_1 [bottom row].

Figures 8-17 present the distribution of first-order and total-effect indices within the geographical domain of interest for all storm case studies, with each pair of figures presenting, respectively, the first-order and total-effect indices for a specific storm; Figures 8-9 for hurricane Katrina, Figures 10-11 for hurricane Gustav, Figures 12-13 for hurricane Irene, Figures 14-15 for super storm Sandy and Figures 16-17 for hurricane Arthur. For all storms, results are presented for both advisories for uncertainty characterization U_1, while for selected storms (Gustav, Irene, Sandy) results are also presented for uncertainty characterization U_2 for advisory A_1. Trends for the omitted storms in the latter case are identical to the ones presented (and not explicitly reported here due to space limitations). In each figure, columns correspond to different storm features, and rows to different advisory or uncertainty characterization cases. Note that in all figures the same scale is used for each storm feature (across the different rows) to allow easier comparisons, but different scale might be used for the different storm features (across the different

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

columns) for clarity of illustration, and this difference should be considered when comparing sensitivity indices across the storm features.

Results reveal significant differences in the spatial distribution of the sensitivity indices, indicating that the different forecast errors impact differently the probabilistic surge estimates within the different sub-regions of the geographical domain of impact for each storm. The trends change across storms, and even across the different advisories or uncertainty characterization cases, though some similar characteristics can be observed. The overall behavior regarding the spatial distribution of the indices appears to be more complex in the Gulf storms (Figures 8-11), owing to the more complicated coastal morphology and, for the storms examined here, to the flood protection systems close to the New Orleans. For the North Atlantic storms (Figures 12-17) trends are more consistent, with first-order indices for features Δs_{along} and Δv_w, impacting the speed and the intensity of storm, being larger in regions close to the nominal track (red line in the plots), and first-order indices for features Δs_{cross} and ΔR_{mw}, impacting the track and the size of the storm, respectively, being typically larger for subdomains away from the nominal track. The latter behavior intuitively makes sense: characteristics that dictate the deviation of the storm from the nominal advisory, as well as its size, defining, therefore, the geographical domain the storm can impact, are more influential for areas that are away from the nominal path. For the Gulf storms, similar underlying trends can be observed, though they are not as clear, owing to the aforementioned more complex distribution of the sensitivity indices in this geographical region. Total-effect indices across both regions of the North Atlantic and the Gulf demonstrate more complicated trends, since in this case the interaction of each storm feature with the remaining ones is included. For some subdomains for which the first-order indices are relatively large, the total-effect indices seem to follow the same underlying trends, since a strong trend has been established by the first-order indices that drives the trends for the total-effect indices (substantial contribution to the total-effect indices comes from the large first-order indices). For other subdomains, though, the accumulated contribution of higher-order interactions can be significant (and constitute a substantial portion of the total-effect indices), and has the potential to change some of the trends established by the first-order indices when examining the spatial distribution of the total-effect indices.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

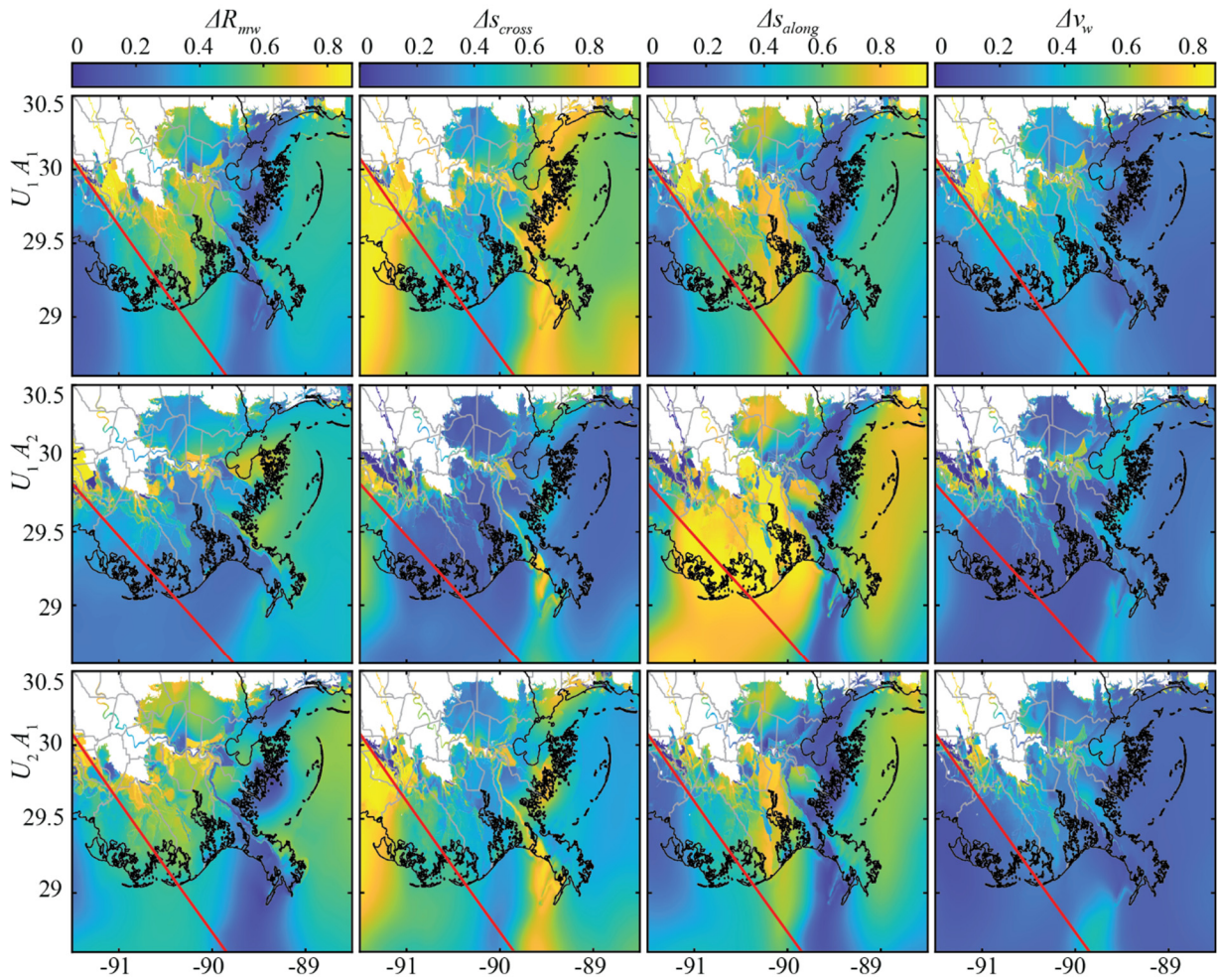


Figure 11. Distribution of the total-effect sensitivity indices over the spatial domain of interest for hurricane Gustav (2008) for a combination of advisory and uncertainty characterization cases, corresponding to U_1A_1 [top row], U_1A_2 [middle row] and U_2A_1 [bottom row].

Comparisons between the first-order and the total-effect indices (compare the subplots in each pair of figures corresponding to the same storm) show that total-effect indices are consistently large for all storm features for the landfilling storms (Katrina, Gustav, Sandy), even for those storm features (Δs_{along} in most cases) that some first-order indices are small. This indicates that higher-order interactions are impactful for the latter type of storm features. The same observation can be made for all the storm features for bypassing storms (Irene, Arthur) apart from Δs_{along}, for which even the total-effect indices remain moderately low. This indicates that for bypassing storms some uncertainty reduction can be potentially accomplished by ignoring the variability along the storm track (removing Δs_{along} from the uncertainty description in the problem formulation). Of course, as discussed earlier, the identified sensitivity with respect to Δs_{along} should be considered as a lower bound due to the potential complex coupling with tides for the peak surge manifestation, which is ignored in this study. Therefore, additional investigations are warranted before such a recommendation is made with respect to any uncertainty reduction. The results from the extensive case

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

studies indicate that all types of forecast errors are important for landfalling storms, and it seems that all these forecast errors might also be important for bypassing storms, depending on the actual importance of Δs_{along} if tides are explicitly considered; no uncertainty reduction seems to be possible. Comparisons of the first-order indices across the different storm features indicate that Δs_{cross} is consistently the forecast error with the relative larger importance across all storms and advisories examined. This trend will be further showcased in the next section, when the aggregated importance indicators across the entire geographic domain (instead the geographic distribution) are examined, but it agrees with the sensitivity trends identified by the NWS across the years, and the fact that the factorial sampling in P-Surge uses a substantially larger number of representative values for Δs_{cross} compared to the other forecast errors (Taylor and Glahn 2008), recognizing its bigger influence in the storm surge variation. The proposed GSA can support this important insight with a complete data-driven implementation for each advisory, instead of the use of parametric studies across multiple storm cases (which is the source of the NWS insights that guided the aforementioned choices in the P-Surge implementation), demonstrating this way the significant benefits GSA can accommodate for probabilistic surge estimation.

Overall, these discussions additionally showcase the different insights obtained from each type of sensitivity index. First-order indices provide insights for the individual importance of each examined storm feature, offering an understanding of how deviations of that feature from the nominal advisory impact the predicted surge. On the other hand, total-effect indices reveal the overall importance of that feature considering any interactions with the remaining ones, providing also guidance for potential uncertainty reduction, if the total-effect index is globally small for a certain storm feature.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

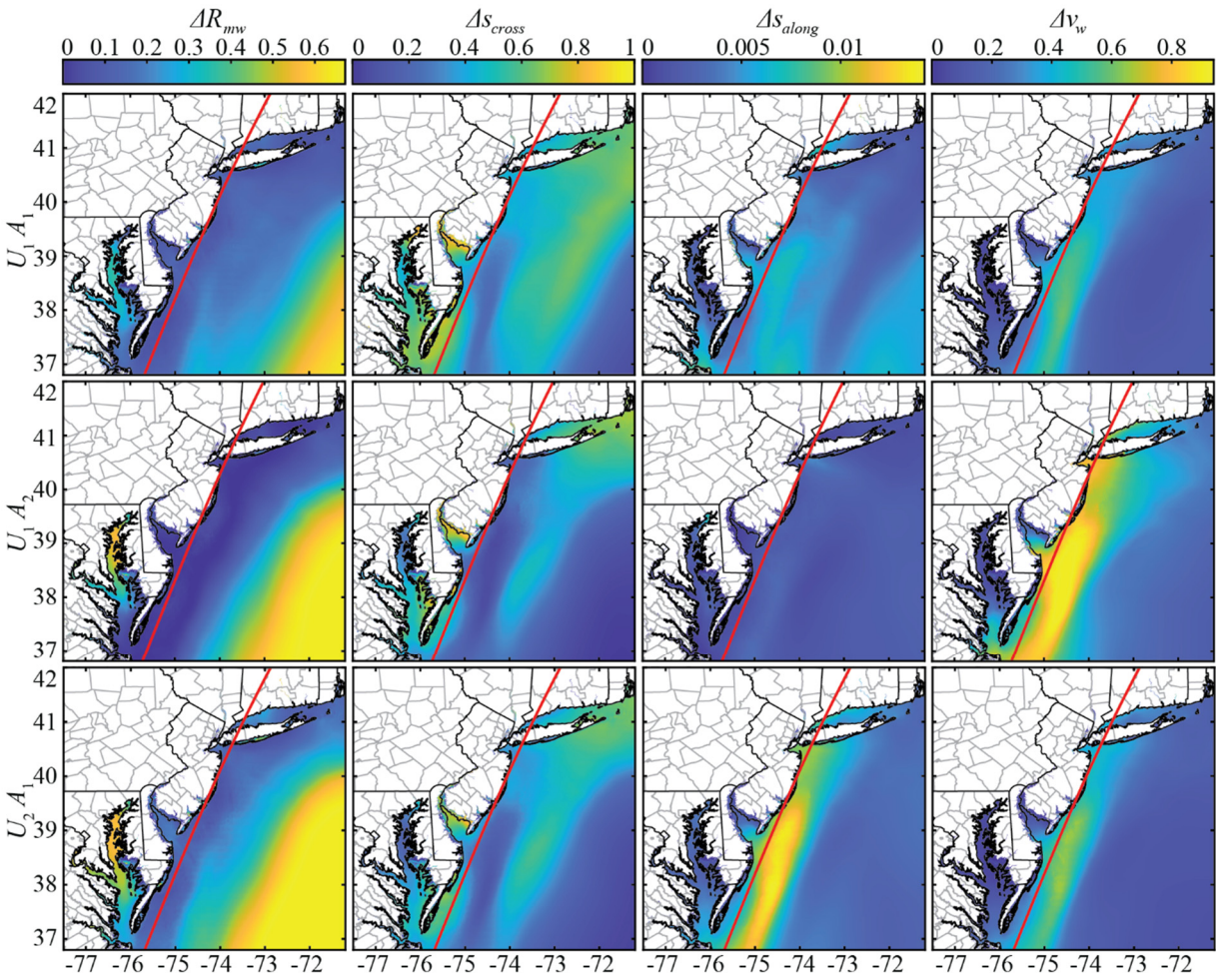


Figure 12. Distribution of the first-order sensitivity indices over the spatial domain of interest for hurricane Irene (2011) for a combination of advisory and uncertainty characterization cases, corresponding to U_1A_1 [top row], U_1A_2 [middle row] and U_2A_1 [bottom row].

Moving now to comparisons across the different advisories, i.e. across the first two rows in each of the figures, we can observe that although some general similarities exist, differences in the magnitude and spatial distribution characteristics of the indices are also very common. The patterns change across the storms and between the different types of indices (first-order or total-effect), showcasing that consistent trends cannot be predicted. The similarities indicate that partial sensitivity information can be shared across the advisories, with foundational insights about the relative importance of forecast errors remaining consistent. This also means that information from one advisory can be used to guide decisions that will be made for the uncertainty propagation step in future advisories. On the other hand, the differences showcase the need of repeating the GSA across the various advisories as the storm evolves. This stresses the importance of an efficient and versatile computational framework for the GSA implementation, a requirement satisfied by the established data-driven PCA-PSA approach that can be seamlessly integrated

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

within existing computational workflows for probabilistic surge estimation without adding any significant computational cost.

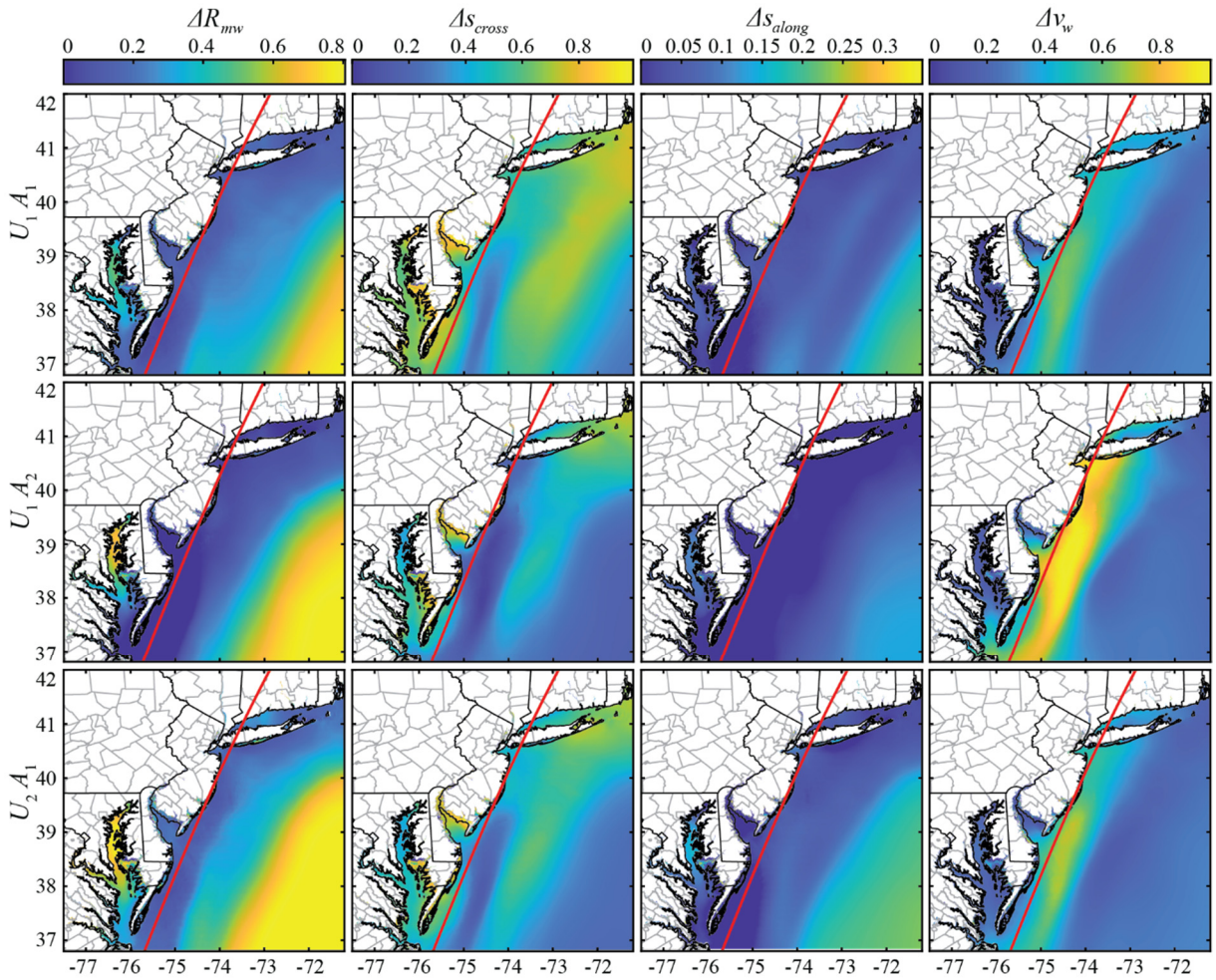


Figure 13. Distribution of the total-effect sensitivity indices over the spatial domain of interest for hurricane Irene (2011) for a combination of advisory and uncertainty characterization cases, corresponding to U_1A_1 [top row], U_1A_2 [middle row] and U_2A_1 [bottom row].

Comparison between the two uncertainty characterization cases for advisory A_1 for hurricanes Gustav (Figures 10 and 11), Irene (Figures 12 and 13) and Sandy (Figures 14 and 15) show a potential influence by the magnitude of the forecast errors (source of differences between the two uncertainty characterization scenarios). This influence is case-dependent, with differences being larger for hurricane Irene. The possibility of the sensitivity trends changing when the uncertainty description for the forecast errors changes is of course expected, but also stresses the importance of repeating the GSA when the magnitude of the forecast errors is updated, instead of relying on insights and trends established while using older error estimates. The efficiency of the proposed PCA-PSA framework readily accommodates this need.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

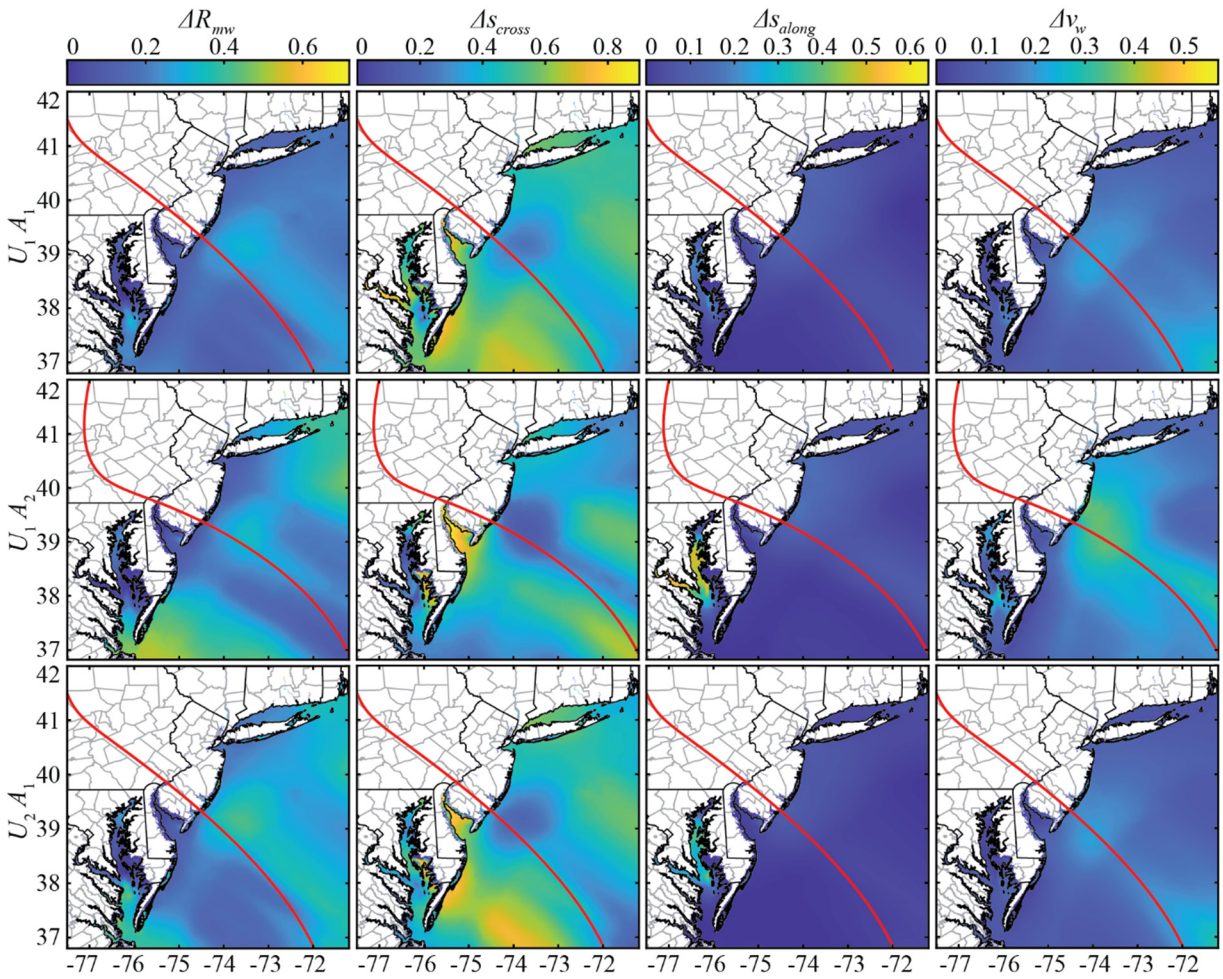


Figure 14. Distribution of the first-order sensitivity indices over the spatial domain of interest for super storm Sandy (2012) for a combination of advisory and uncertainty characterization cases, corresponding to U_1A_1 [top row], U_1A_2 [middle row] and U_2A_1 [bottom row].

Finally Figures 18 and 19 present some results for selected higher-order indices, Figure 18 for storms in the Gulf region and Figure 19 for storms in the North Atlantic region. Results are presented only for advisory A_1 and uncertainty characterization U_1, and for only four combinations of storm features (i.e. four different higher-order GSA indices) in each figure. The combinations presented are the ones exhibiting larger values for the higher-order interactions for at least one of the storms examined. The intention here is not to establish an exhaustive investigation to obtain in depth insights into the interaction across storm features for different type of storms (as performed earlier for the first-order and total-effect indices), rather simply to demonstrate the potential of GSA to accommodate insights for storm-specific interaction for the forecast errors. Results demonstrate that such interactions between the forecast errors can be important (non-negligible values for the sensitivity indices reported), and exhibit significant spatial variability. Trends are not necessarily consistent across different storms even within the same geographic region. Therefore, any insights about the sensitivity interactions of the storm features should focus on each separate storm and

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

advisory. Since the number of such interactions, corresponding to the combination of different storm features, is large, the efficiency of the proposed PCA-GSA formulation is key in performing the analyses to accommodate the desired insights. It should be noted, though, that GSA formulations typically focus on first-order and total-effect indices, since identifying trends for the higher-order interactions is always more challenging. Still, the proposed GSA formulation provides the ability to efficiently estimate higher-order interaction effects.

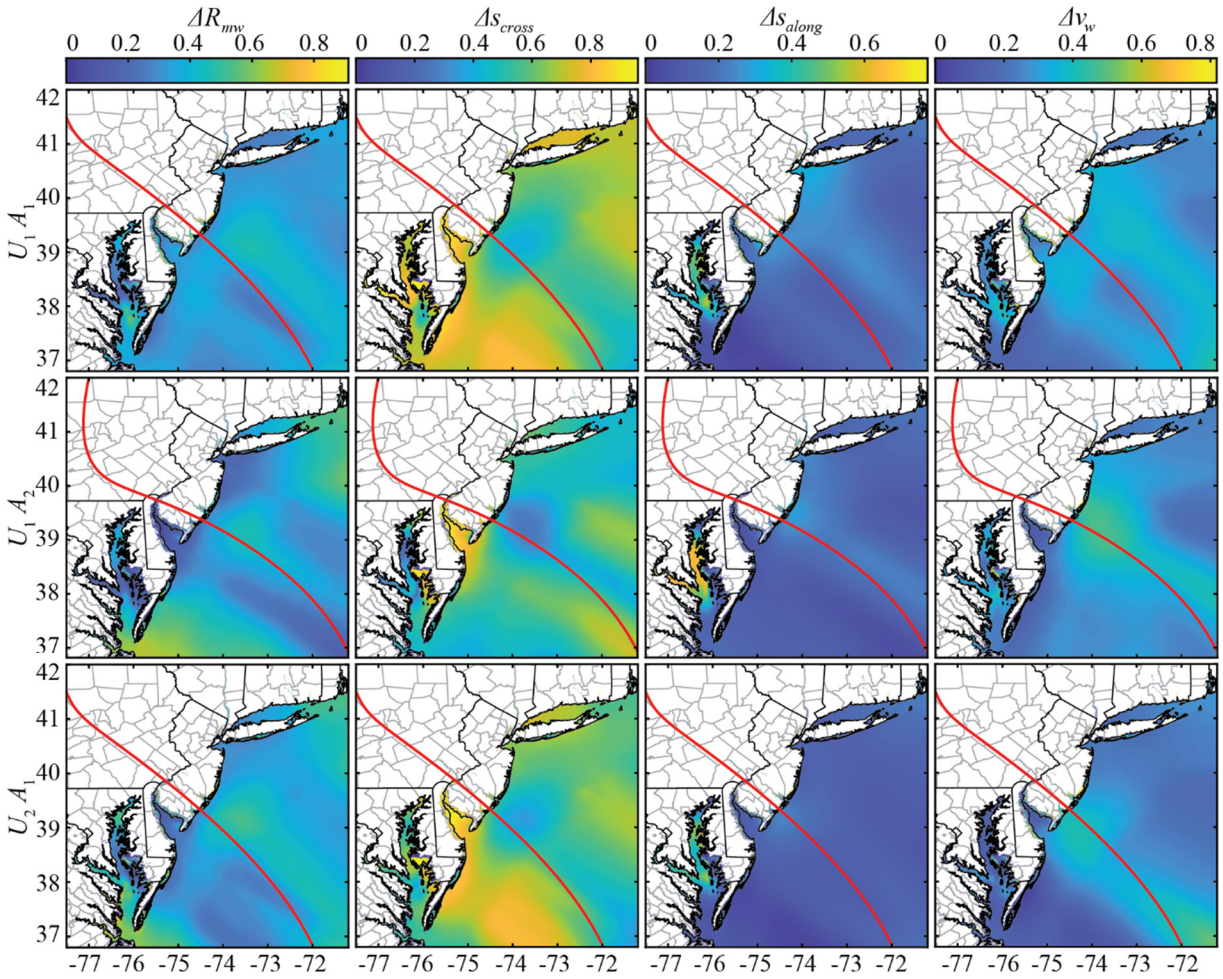


Figure 15. Distribution of the total-effect sensitivity indices over the spatial domain of interest for super storm Sandy (2012) for a combination of advisory and uncertainty characterization cases, corresponding to U_1A_1 [top row], U_1A_2 [middle row] and U_2A_1 [bottom row].

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

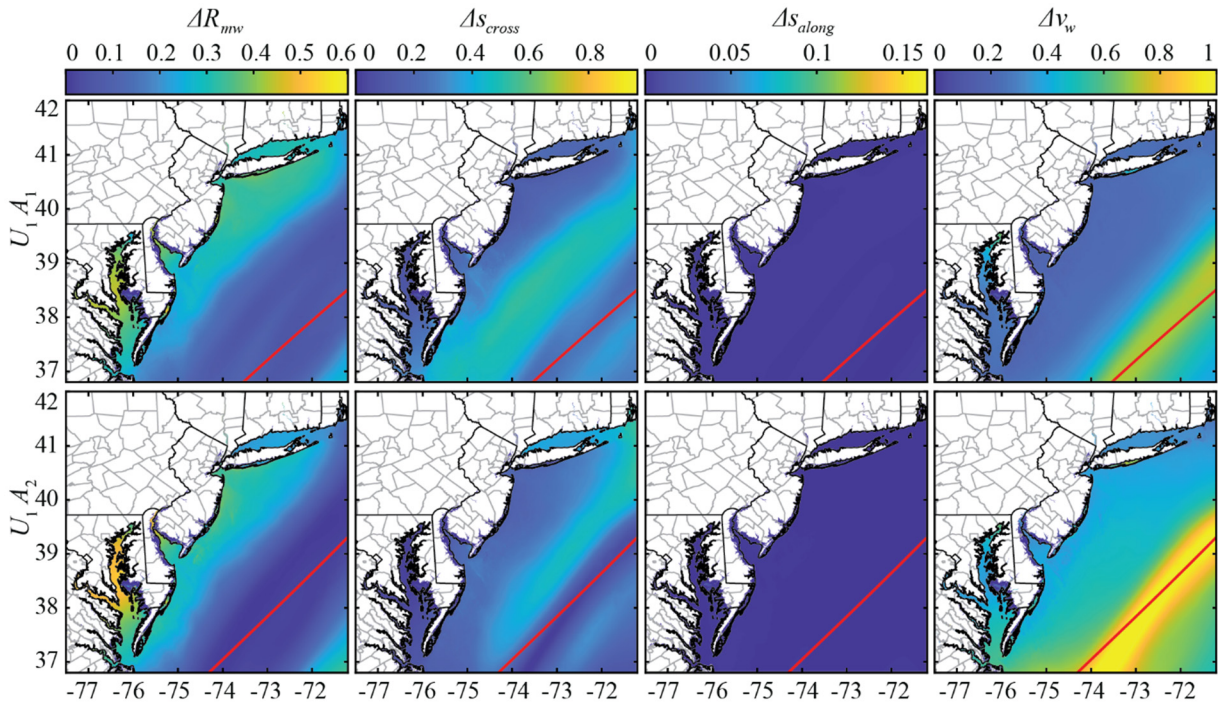


Figure 16. Distribution of the first-order sensitivity indices over the spatial domain of interest for hurricane Arthur (2014) for advisories A_1 [top row] and A_2 [bottom row] with uncertainty characterization U_1.

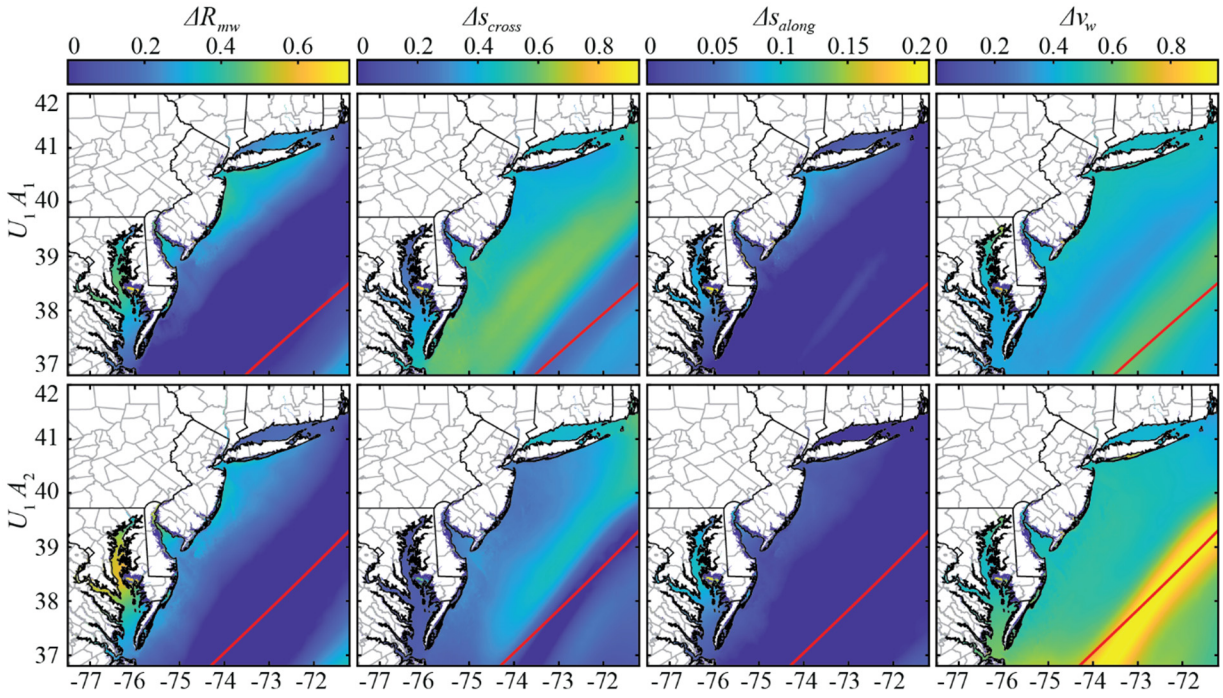


Figure 17. Distribution of the total-effect sensitivity indices over the spatial domain of interest for hurricane Arthur (2014) for advisories A_1 [top row] and A_2 [bottom row] with uncertainty characterization U_1.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

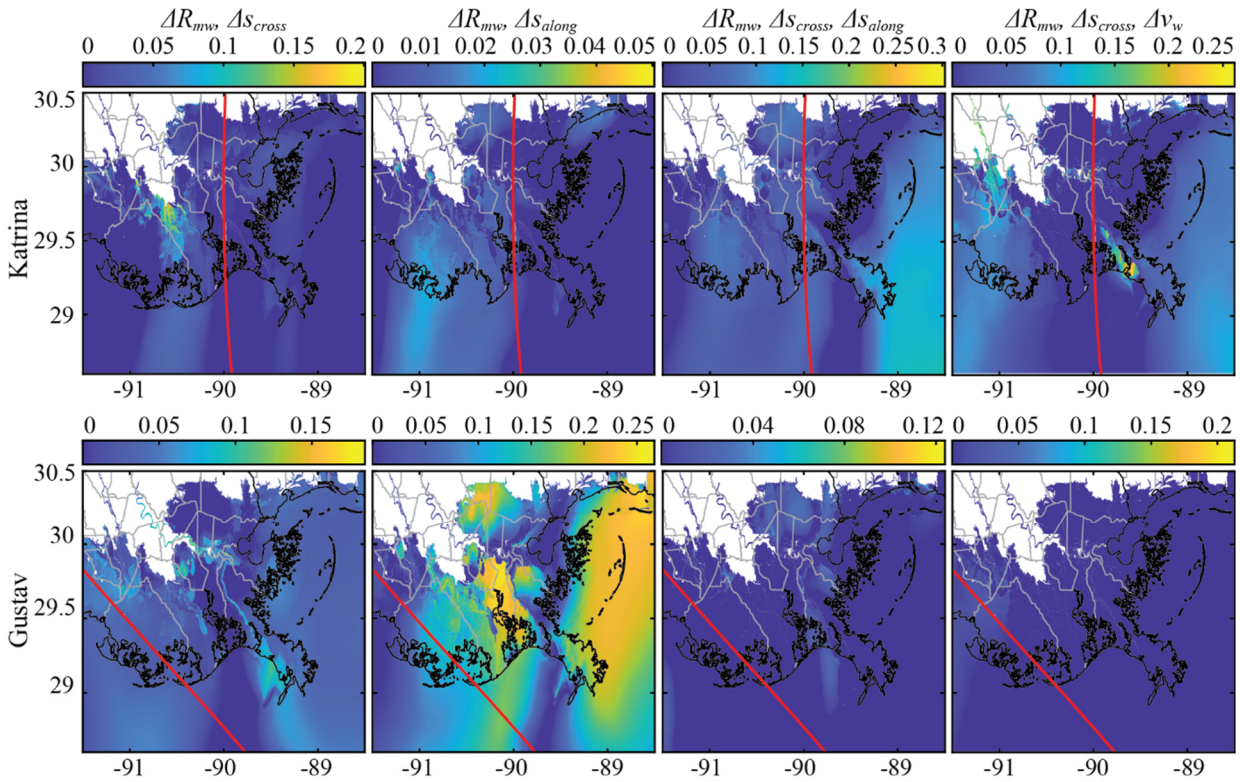


Figure 18: Distribution of some higher-order sensitivity indices over the spatial domain of interest for storms in the Gulf, corresponding to hurricanes Katrina (2005) [top row] and Gustav (2008) [bottom row] for advisory A_1 and uncertainty characterization U_1.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

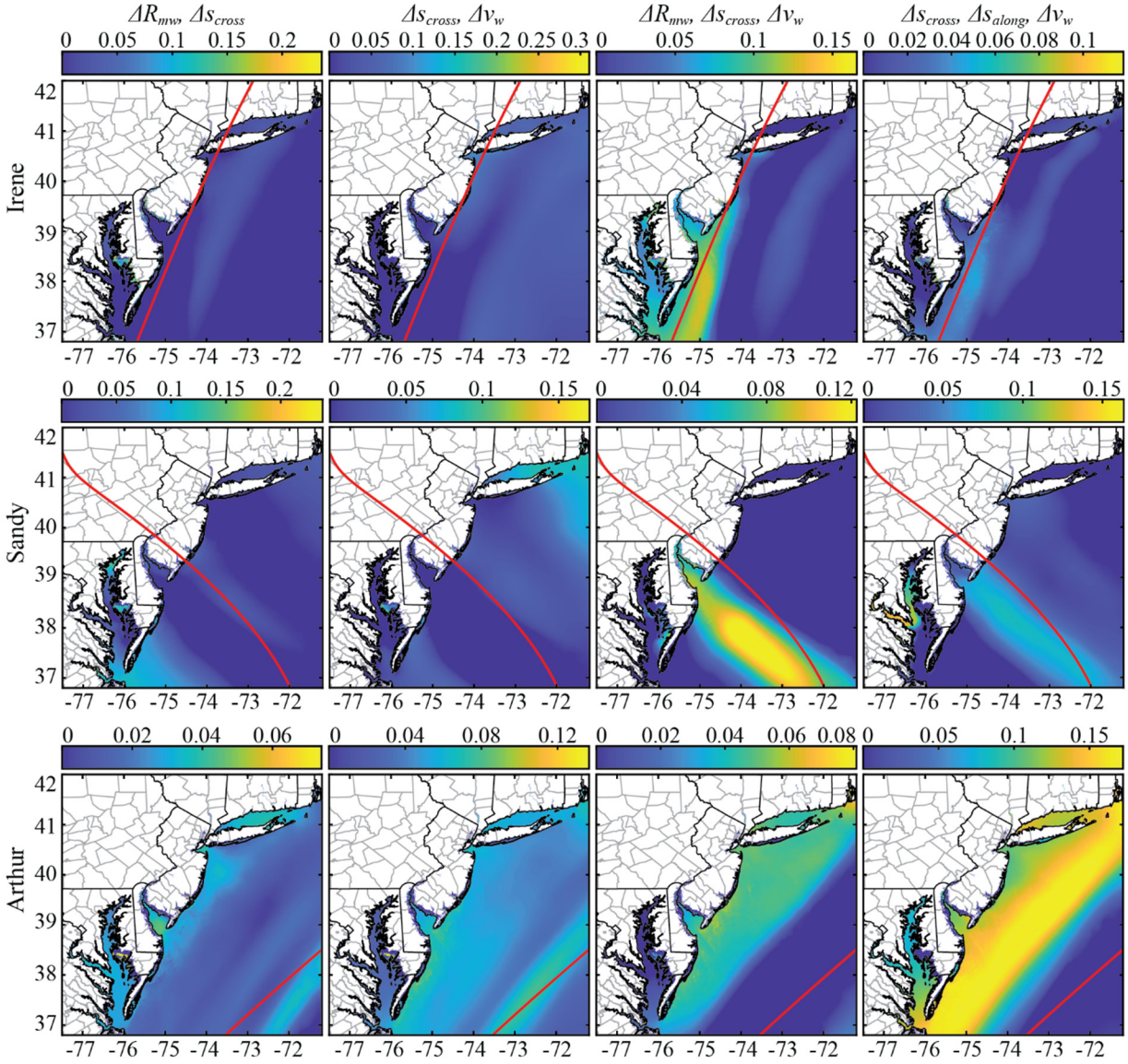


Figure 19: Distribution of some higher-order sensitivity indices over the spatial domain of interest for storms in the North Atlantic, corresponding to hurricanes Irene (2011) [top row], Sandy (2012) [middle row] and Arthur (2014) [bottom row] for advisory A_1 and uncertainty characterization U_1.

5.3 Aggregated importance indicators

Discussions in the previous section focused on the spatial distribution of the sensitivity indices. This section examines the aggregated sensitivity indices introduced in Section 3.3, which quantify the aggregated importance of each of the storm features across the entire geographic domain that the storm impacts, using the variance (presented in Figures 6 and 7) as a weighting factor of the individual (for each location) sensitivity indices.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

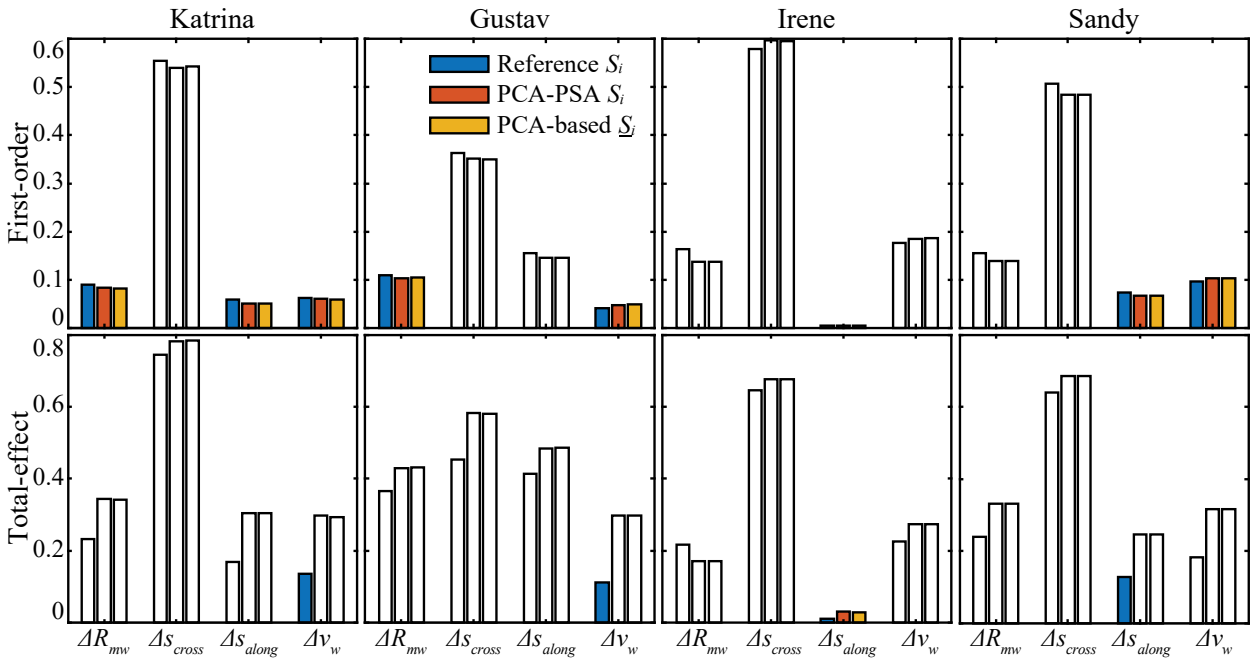


Figure 20: Comparison of the aggregated first-order [top row] and total-effect [bottom row] sensitivity indices estimated by the different approaches for selected storm case studies for advisory A_1 and uncertainty characterization U_1. Approaches examined are: reference and PCA-PSA estimates for indices corresponding to the original output S_i, as well as PCA-based estimates $\underline{S}_i$ obtained in the latent space.

First, the accuracy of the different methods for calculating the aggregated importance indicators is investigated in Figure 20. Specifically, the aggregated first-order and total-effect indices are presented using three different estimation approaches: using the original output based on either the double-loop Monte Carlo reference estimates or the PCA-PSA estimates, or using the latent output leveraging partial PCA-PSA results. Results are presented only for advisory A_1 and uncertainty characterization scenario U_1, though any established trends below are identical for all other cases. Looking first at the results for the sensitivity indices estimated for the original output, it is evident that the PCA-PSA implementation provides a very good match to the reference estimates for the first-order aggregated indices. For the total-effect indices, a small loss of accuracy is observed, with PCA-PSA overestimating the aggregated indices, though the relative importance trends are consistent across both approaches. These comparisons offer an additional validation of the PCA-PSA GSA estimates, in this case with respect to the identification of the aggregated importance of each of the storm features. Moreover, the comparison of the PCA-PSA estimates based on the original and the latent output (comparing red and yellow bars), demonstrates good agreement for both the first-order and total-effect indices across all examined cases. This is a very important result, showing that if the aggregated importance of the storm features is the only required information, without the explicit need to calculate the spatial distribution of the sensitivity indices, then the estimation can be performed using the principal component conditional variance statistics. This accommodates significant computational

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

savings as the transformation back to the original output or the estimation of conditional covariance statistics for the latent outputs is not warranted.

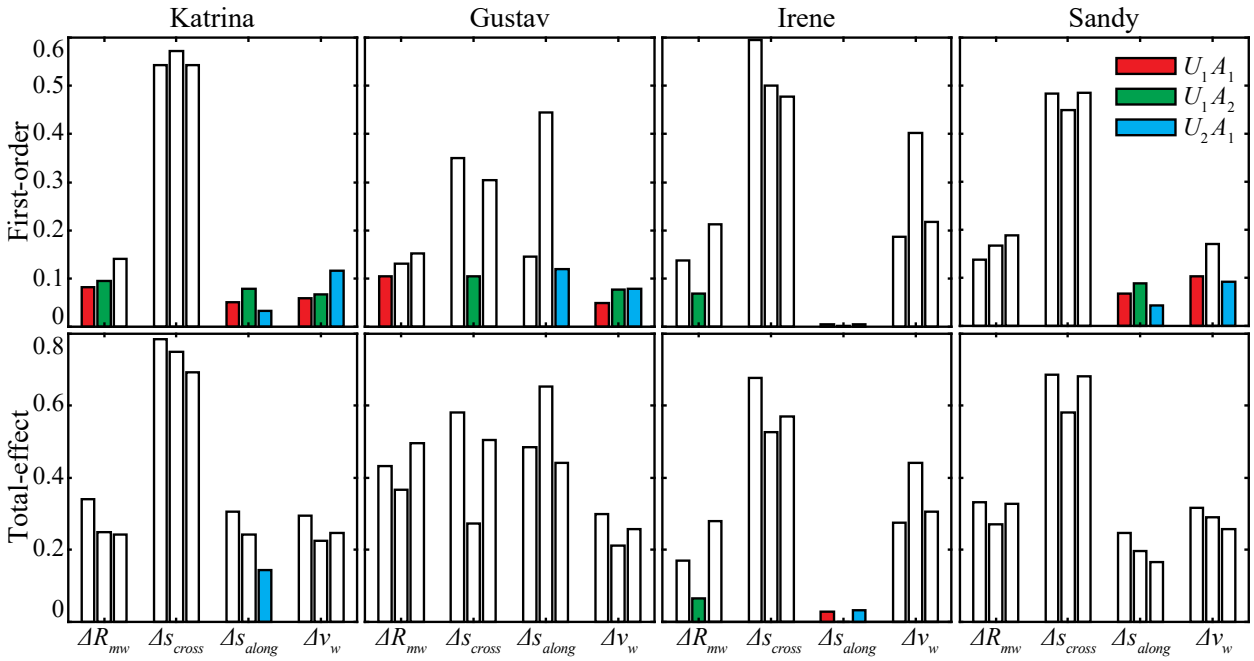


Figure 21: Aggregated first-order [top row] and total-effect [bottom row] sensitivity indices for selected case study storms for different combinations of advisory and uncertainty characterization: U_1A_1, U_1A_2 and U_2A_1.

Finally, Figure 21 compares the aggregated importance indicators (first-order and total-effect) for the different advisories and uncertainty characterizations. The comparisons between the aggregated sensitivity indices verify the trends identified when similar comparisons across advisories and uncertainty characterization cases were established based on the spatial distribution of the indices. Although basic trends remain relatively consistent across the different advisories, demonstrating how GSA insights from one advisory can be potentially leveraged to understand the importance of the forecast errors for future advisories, some differences also exist. The magnitude of these differences is case-dependent and can be potentially significant, as evident, for example by the significant increase of importance for Δs_{along} for Gustav between advisories A_1 and A_2. The comparison of importance for individual storm features shows that Δs_{cross} always holds a dominant importance, while on the other hand, Δs_{along} appears insignificant for bypassing storms. Both these trends agree with the ones identified in Section 5.2 when examining the spatial distribution of the sensitivity indices. This agreement showcases how the proposed aggregated sensitivity indices can help quantify the global importance of each of the forecast errors across the geographic domain of interest, with no need to establish and examine any spatial distribution maps.

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

6 Conclusions

The application of a variance-based global sensitivity analysis (GSA), quantified through the estimation of Sobol' indices, was investigated in this paper to explore the importance of storm forecast errors in probabilistic peak surge predictions. The formulation of the probabilistic surge estimates adopted the framework currently employed by the National Weather Service (NWS), considering the variability in four different storm features: size, intensity, along- and cross-track position of the storm eye. The GSA was formulated through the estimation of first- and higher-order sensitivity indices and of total-effect indices. Each type of index offers different insights with respect to the relative importance of each of the storm features individually or through the inclusion of the correlation with other storm features, and can provide valuable insights regarding the anticipated surge variability. A computationally efficient and versatile GSA framework was presented, termed PCA-PSA, established by combining Principal Component Analysis (PCA) and a probability-based estimation of the variance of conditional expectations. Principal component Analysis (PCA) is leveraged as a dimensionality reduction technique, to address the requirement of estimating the sensitivity indices for multiple locations within the geographic domain of interest, while the probability-based variance estimation accommodates the constraint of establishing the GSA using only a small number of hydrodynamic numerical simulations, since the associated computational burden of such simulations is significant. The established PCA-PSA framework is completely data-driven, and can utilize any given set of hydrodynamic simulations, and therefore can be seamlessly integrated within any computational workflow for probabilistic surge estimation, without requiring additional specialized simulations to accommodate the GSA implementation. Beyond the estimation of individual sensitivity indices for each location within the geographic domain of interest, aggregated importance indicators were also estimated, quantifying the overall importance of each storm feature. This was established through an averaging of the individual indices using variance-based weighting. A computational efficient formulation of the aggregated indicators was also proposed, using characteristics available for the principal components, and was shown to provide consistent results with the aggregated indices based on the original high-dimensional output.

An illustrative case study was considered, examining the GSA for different advisories and magnitude of forecast errors for five different storms in the North Atlantic and the Gulf regions. The influence of tides was ignored across the case studies, something that should be recognized as limitation of the work, impacting in particular the sensitivity with respect to the along-tract forecast errors. The validation of the proposed PCA-PSA implementation was first examined showing good overall accuracy for the problem considered here; excellent agreement was observed for the first-order indices, and while the overall accuracy was reduced for the total-effect indices, the correct sensitivity trends were identified. Comparisons of individual GSA results offered the following insights:

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

- The spatial distribution of the sensitivity indices shows significant variability, indicating that the different forecast errors affect differently the surge estimates within sub-regions of storm impact.
- The spatial distribution of the indices appears to be more complex in the Gulf region, or, generalizing, across coasts with more complex geomorphologies.
- Storm characteristics that dictate the deviation of the storm from the nominal advisory and/or its size, are more influential for areas that are away from the nominal storm path.
- Total-effect indices demonstrate more complex trends compared to first-order indices, indicating important higher-order interactions among the forecast errors.
- All types of forecast errors are important for landfalling storms, and it seems that all these forecast errors might also be important for bypassing storms, depending on the (increased) importance of along-track errors if tides are explicitly considered.
- The cross-track error has been consistently identified as the forecast error with the relative largest importance.
- For a specific storm, similarities across advisories indicate that partial sensitivity information can be shared across them. Still, the reported differences showcase the need of repeating the GSA across the various advisories as the storm evolves.
- GSA trends depend on the magnitude of the forecast errors. This means that the sensitivity analysis needs to be repeated when such errors are updated, instead of relying on insights and trends established while using older error estimates.

Overall, the study stresses the insights GSA can offer and the importance of a computationally efficient GSA framework, like the data-driven PCA-PSA proposed here, that can be repeated fast and accurately for each specific storm advisory, or forecast error quantification.

7 Statements and Declarations

Funding

This research was funded by the National Oceanographic and Atmospheric Administration (NOAA), under the grant number NA19OAR0220089. The views and opinions expressed in this paper are those of the authors and do not represent NOAA's official position.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

Conceptualization, A.A.Taflanidis, W.Jung and A.P.Kyprioti; methodology, W.Jung, and A.A. Taflanidis; software, W. Jung, and E.Adeli; validation, W.Jung, and A.A.Taflanidis; data curation, W.Jung

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

and A.P.Kyprioti; writing—original draft preparation, A.A.Taflanidis, W.Jung and A.P.Kyprioti; writing—review and editing, A.A.Taflanidis, W.Jung, A.P.Kyprioti, and E.Adeli; funding acquisition, A.A.Taflanidis. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

Authors would like to thank Arthur Taylor for offering guidance with respect to the sampling implemented in version 2.7 of P-Surge, for providing access to the NHC forecast error statistics and advisories used in this study, and for his feedback related to the importance of the along track variability of storms when tides are incorporated in the analysis. They would also like to thank the Army Corps of Engineers, Coastal Hydraulics Laboratory of the Engineering Research and Development Center for providing access to the databases that were used to develop the surrogate models for the storm surge predictions.

Data Availability

Data utilized for the creation of the surrogate models were provided by the Army Corps of Engineers through the Coastal Hazards System directory https://chs.erdc.dren.mil/default.aspx. Data for the storm advisories can be found in https://rammb-data.cira.colostate.edu. All other data used are reported in the manuscript.

Appendix A: Review of PCA implementation

Define central normalized output $\overline{z}_j^h = z_j^h - \mu_{z_j}$ for the hth storm simulation, where the mean surge over the database is estimated as:

$$\mu_{z_j} = \frac{1}{k}\sum_{h=1}^{k} z_j^h \tag{A.1}$$

and consider the eigenvalue problem for the covariance matrix $\overline{\mathbf{Z}}^T\overline{\mathbf{Z}}$ associated with the normalized observation matrix $\overline{\mathbf{Z}}$, whose hth column corresponds to the normalized output $\overline{\mathbf{z}}^h$ (with elements $\overline{z}_j^h$). The solution of the eigenvalue problem provides the vector of latent outputs (principal components) with the jth latent output denoted as y_j. The corresponding eigenvalue λ_j represents the portion of the total variance of the original data $\mathbf{Z}$ that can be explained by y_j, while the respective eigenvector $\mathbf{P}_j \in \mathbb{R}^{n_z}$ facilitates the mapping from $\mathbf{z}$ to y_j. To accommodate a larger dimensionality reduction only the principal components corresponding to the n_p largest eigenvalues are retained, with n_p chosen such that the ratio

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

$$r = \frac{\sum_{j=1}^{n_p} \lambda_j}{\sum_{j=1}^{\min(n_z, k-1)} \lambda_j} \tag{A.2}$$

is greater than some threshold r_o (for example 99%). The vector with the retained principal components is denoted by **y**, while the relationship between **y** and **z** is $\mathbf{z} = \mathbf{P}\mathbf{y} + \boldsymbol{\mu}_z + \boldsymbol{\tau}$ where $\mathbf{P} \in \mathbb{R}^{n_z \times n_p}$ is the projection matrix with the *j*th column corresponding to the eigenvectors $\mathbf{P}_j \in \mathbb{R}^{n_z}$, $\boldsymbol{\mu}_z$ is the vector with elements μ_{z_j}, and $\boldsymbol{\tau}$ corresponds to the approximation error.

PCA maximizes the variance in the original data that has been preserved when considering only n_p components, while it also minimizes the total squared reconstruction error between **z** and $\mathbf{P}\mathbf{y} + \boldsymbol{\mu}_y$ over the database. The latter corresponds to the truncation error over the database, stemming from the fact that not all principal components are retained.

Appendix B: Details for estimation of $\boldsymbol{\Sigma}_{\mathbf{y}}^{\mathbf{c}}$ using Gaussian Mixture Model (GMM)

As discussed in Section 3.2, the efficient estimation of $\boldsymbol{\Sigma}_{\mathbf{y}}^{\mathbf{c}}$ starts with establishing a GMM fit for the n_c+1 dimensional PDF of $\mathbf{x_c}$ and y_j utilizing subset $[\mathbf{X_c}, \mathbf{Y}_j]$ of the original data. The GMM parameters can be estimated using the Expectation Maximization (EM) algorithm (Moon 1996), whereas for selecting the number of mixture components, techniques relying on the Bayesian Information Criterion (BIC) can be utilized to avoid overfitting the available data (McNicholas and Murphy 2008). The probability-weights of the original samples for $\mathbf{x_c}$ [if they do not directly correspond to samples from $p(\mathbf{x})$] should be explicitly used when performing the GMM fit.

This GMM approximation of the joint PDF of $\mathbf{x_c}$ and y_j is written as:

$$p(\mathbf{x_c}, y_j) = \sum_{q=1}^{Q} w_q^{(j)} N(\mathbf{x_c}, y_j \mid \boldsymbol{\mu}_q^{(j)}, \boldsymbol{\Sigma}_q^{(j)}) \tag{B.1}$$

where Q is the number of mixture components, $w_q^{(j)}$, $\boldsymbol{\mu}_q^{(j)}$, and $\boldsymbol{\Sigma}_q^{(j)}$ are the weight, mean vector and covariance matrix of the *q*th GMM component, respectively, and $N(\mathbf{v}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a Gaussian PDF with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ evaluated at **v**. Superscript (*j*) is used for all relevant GMM parameters to distinguish that the GMM fit pertains to $\mathbf{x_c}$ and y_j combination (stresses dependence on each specific output *j*). The mean and covariance of each component can be partitioned as:

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

$$\boldsymbol{\mu}_q^{(j)} = \begin{bmatrix} \boldsymbol{\mu}_{q,\mathbf{x_c}}^{(j)} \\ \mu_{q,y_j}^{(j)} \end{bmatrix}$$
$$\boldsymbol{\Sigma}_q^{(j)} = \begin{bmatrix} \boldsymbol{\Sigma}_{q,\mathbf{x_c}}^{(j)} & \boldsymbol{\Sigma}_{q,\mathbf{x_c}y_j}^{(j)} \\ \boldsymbol{\Sigma}_{q,y_j\mathbf{x_c}}^{(j)} & \boldsymbol{\Sigma}_{q,y_j}^{(j)} \end{bmatrix} \tag{B.2}$$

where subscript $\mathbf{x_c}$ or y_j (or combined) is used to denote the specific random variable of the joint PDF that the statistic pertains to. Note that $\boldsymbol{\Sigma}_{q,y_j}^{(j)}$ is a scalar (the variance of y_j for the GMM fit) but for notational consistency it is denoted as a covariance matrix.

Based on this fitted GMM, the conditional expectation $E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}]$ can be expressed in a closed form solution leveraging the conditional distribution of the fitted Gaussian mixture probability model as (Hu and Mahadevan 2019):

$$E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}] = \sum_{q=1}^{Q} w_q^{(j)}(\mathbf{x_c}) \cdot \mu_{q,y_j|\mathbf{x_c}}^{(j)} \tag{B.3}$$

where the equations of the conditional statistics $w_q^{(j)}(\mathbf{x_c})$ and $\mu_{q,y_j|\mathbf{x_c}}^{(j)}$ are obtained for each component of the GMM through the conditional statistics of the respective Gaussian as:

$$w_q^{(j)}(\mathbf{x_c}) = \frac{w_q^{(j)} N(\mathbf{x_c} \mid \boldsymbol{\mu}_{q,\mathbf{x_c}}^{(j)}, \boldsymbol{\Sigma}_{q,\mathbf{x_c}}^{(j)})}{\sum_{q'=1}^{Q} w_{q'}^{(j)} N(\mathbf{x_c} \mid \boldsymbol{\mu}_{q',\mathbf{x_c}}^{(j)}, \boldsymbol{\Sigma}_{q',\mathbf{x_c}}^{(j)})}$$
$$\mu_{q,y_j|\mathbf{x_c}}^{(j)} = \mu_{q,y_j}^{(j)} + \boldsymbol{\Sigma}_{q,y_j\mathbf{x_c}}^{(j)} \left(\boldsymbol{\Sigma}_{q,\mathbf{x_c}}^{(j)}\right)^{-1} (\mathbf{x_c} - \boldsymbol{\mu}_{q,\mathbf{x_c}}^{(j)}) \tag{B.4}$$

Finally Monte Carlo Integration (MCI) is used to estimate the elements of $\boldsymbol{\Sigma}_{\mathbf{y}}^{\mathbf{c}}$. Using N total samples of $\mathbf{x_c}$, generated from density $p(\mathbf{x_c})$, with $\mathbf{x_c}^{(g)}$ denoting the gth sample, the MCI expression for the diagonal (variance) elements of $\boldsymbol{\Sigma}_{\mathbf{y}}^{\mathbf{c}}$ is:

$$Var_{\mathbf{c}}[E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}]] \approx \frac{1}{N}\sum_{g=1}^{N}\left\{E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}^{(g)}] - \left(\frac{1}{N}\sum_{g=1}^{N} E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}^{(g)}]\right)\right\}^2 \tag{B.5}$$

whereas the expression for the off-diagonal elements is:

$$\begin{aligned} &Cov_{\mathbf{c}}[E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}],\ E_{\sim\mathbf{c}}[y_l \mid \mathbf{x_c}]] \\ &\approx \frac{1}{N}\sum_{g=1}^{N}\left\{E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}^{(g)}] - \left(\frac{1}{N}\sum_{g=1}^{N} E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}^{(g)}]\right)\right\}\left\{E_{\sim\mathbf{c}}[y_l \mid \mathbf{x_c}^{(g)}] - \left(\frac{1}{N}\sum_{g=1}^{N} E_{\sim\mathbf{c}}[y_l \mid \mathbf{x_c}^{(g)}]\right)\right\} \\ &\approx \frac{1}{N}\sum_{g=1}^{N} E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}^{(g)}] E_{\sim\mathbf{c}}[y_l \mid \mathbf{x_c}^{(g)}] - \left(\frac{1}{N}\sum_{g=1}^{N} E_{\sim\mathbf{c}}[y_j \mid \mathbf{x_c}^{(g)}]\right)\left(\frac{1}{N}\sum_{g=1}^{N} E_{\sim\mathbf{c}}[y_l \mid \mathbf{x_c}^{(g)}]\right) \end{aligned} \tag{B.6}$$

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

References

Ayyad M, Hajj MR, Marsooli R (2021) Spatial Variation in Sensitivity of Hurricane Surge Characteristics to Hurricane Parameters. Journal of Engineering Mechanics 147 (10):04021070

Bishop CM (2006) Pattern recognition and machine learning. Springer, New York, NY

Blake ES (2018) The 2017 Atlantic hurricane season: catastrophic losses and costs. Weatherwise 71 (3):28-37

Chen JH, Lin SJ, Magnusson L, Bender M, Chen X, Zhou L, Xiang B, Rees S, Morin M, Harris L (2019) Advancements in hurricane prediction with NOAA's next-generation forecast system. Geophysical Research Letters 46 (8):4495-4501

Chen W, Jin R, Sudjianto A (2005) Analytical variance-based global sensitivity analysis in simulation-based design under uncertainty. Journal of Mechanical Design 127 (5):875-886

Dangendorf S, Marcos M, Wöppelmann G, Conrad CP, Frederikse T, Riva R (2017) Reassessment of 20th century global mean sea level rise. Proceedings of the National Academy of Sciences 114 (23):5946-5951

Fu G, Kapelan Z, Reed P (2012) Reducing the complexity of multiobjective water distribution system optimization through global sensitivity analysis. Journal of Water Resources Planning and Management 138 (3):196-207

Glahn B, Taylor A, Kurkowski N, Shaffer WA (2009) The role of the SLOSH model in National Weather Service storm surge forecasting. National Weather Digest 33 (1):3-14

Gonzalez T, Taylor A (2018) Development of the NWS'Probabilistic Tropical Storm Surge Model. Paper presented at the 33rd Conference on Hurricanes and Tropical Meteorology, Ponte Vedra, FL,

Hamill TM, Brennan MJ, Brown B, DeMaria M, Rappaport EN, Toth Z (2012) NOAA's future ensemble-based hurricane forecast products. Bulletin of the American Meteorological Society 93 (2):209-220

Homma T, Saltelli A (1996) Importance measures in global sensitivity analysis of nonlinear models. Reliability Engineering & System Safety 52 (1):1-17

Hu Z, Mahadevan S (2016) Global sensitivity analysis-enhanced surrogate (GSAS) modeling for reliability analysis. Structural and Multidisciplinary Optimization 53 (3):501-521

Hu Z, Mahadevan S (2019) Probability models for data-driven global sensitivity analysis. Reliability Engineering & System Safety 187:40-57

Iooss B, Lemaître P (2015) A review on global sensitivity analysis methods. In: Uncertainty management in simulation-optimization of complex systems. Springer, pp 101-122

Irish JL, Resio DT, Ratcliff JJ (2008) The influence of storm size on hurricane surge. Journal of Physical Oceanography 38 (9):2003-2013

Javeline D, Kijewski-Correa T (2019) Coastal homeowners in a changing climate. Climatic Change 152 (2):259-274

Jelesnianski CP, Chen J, Shaffer WA (1992) SLOSH: Sea, lake, and overland surges from hurricanes. NOAA Technical Report, NWS 48. US Department of Commerce, National Oceanic and Atmospheric Administration,

Jelesnianski CP, Taylor A (1973) A preliminary view of storm surges before and after storm modifications. NOAA technical memorandum. vol 3. Environmental Research Laboratories, Weather Modification Program Office, Washington, DC

Jia G, Taflanidis AA (2016) Efficient evaluation of Sobol' sensitivity indices utilizing samples from an auxiliary probability density function. Engineering Mechanics 142 (5):04016012

Jia G, Taflanidis AA, Nadal-Caraballo NC, Melby J, Kennedy A, Smith J (2016) Surrogate modeling for peak and time dependent storm surge prediction over an extended coastal region using an existing database of synthetic storms. Natural Hazards 81 (2):909-938

Jolliffe IT (2002) Principal component analysis. Springer series in statistics, 2nd edn. Springer, New York

Jung W, Taflanidis AA (2022) Efficient global sensitivity analysis for high-dimensional outputs combining data-driven probability models and dimensionality reduction techniques. Reliability Engineering & System Safety. https://doi.org/10.1016/j.ress.2022.108805

Kennedy AB, Westerink JJ, Smith J, Taflanidis AA, Hope M, Hartman M, Tanaka S, Westerink H, Cheung KF, Smith T, Hamman M, Minamide M, Ota A (2012) Tropical cyclone inundation potential on the Hawaiian islands of Oahu and Kauai. Ocean Modelling 52-53:54–68

Kijewski-Correa T, Taflanidis A, Vardeman C, Sweet J, Zhang J, Snaiki R, Wu T, Silver Z, Kennedy A (2020) Geospatial environments for hurricane risk assessment: applications to situational awareness and resilience planning in New Jersey. Frontiers in Built Environment 6:549106

Klotzbach PJ, Schreck III CJ, Collins JM, Bell MM, Blake ES, Roache D (2018) The extremely active 2017 North Atlantic hurricane season. Monthly Weather Review 146 (10):3425-3443

Knaff JA, Zehr RM (2007) Reexamination of tropical cyclone wind–pressure relationships. Weather and Forecasting 22 (1):71-88

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

Kyprioti AP, Adeli E, Taflanidis AA, Westerink JJ, Tolman HL (2021a) Probabilistic Storm Surge Estimation for Landfalling Hurricanes: Advancements in Computational Efficiency Using Quasi-Monte Carlo Techniques. Journal of Marine Science and Engineering 9 (12):1322
Kyprioti AP, Taflanidis AA, Nadal-Caraballo NC, Campbell M (2021b) Storm hazard analysis over extended geospatial grids utilizing surrogate models. Coastal Engineering:103855
Li C, Mahadevan S (2016) An efficient modularized sample-based method to estimate the first-order Sobol′ index. Reliability Engineering & System Safety 153:110-121
Li M, Wang R-Q, Jia G (2020) Efficient dimension reduction and surrogate-based sensitivity analysis for expensive models with high-dimensional outputs. Reliability Engineering & System Safety 195:106725
Lim Y-K, Schubert SD, Kovach R, Molod AM, Pawson S (2018) The roles of climate change and climate variability in the 2017 Atlantic hurricane season. Scientific Reports 8 (1):1-10
Luettich RA, Jr. , Westerink JJ, Scheffner NW (1992) ADCIRC: An advanced three-dimensional circulation model for shelves, coasts, and estuaries. Report 1. Theory and methodology of ADCIRC-2DDI and ADCIRC-3DL. Dredging Research Program Technical Report DRP-92-6, U.S Army Engineers Waterways Experiment Station, Vicksburg,MS
Marshall JS, Daniel, Lyda A, Roueche D, Davis B, Djima W, Heo Y, Kijewski-Correa T, Moravej M, Rittelmeyer B, Salman A, Prevatt D, Robertson I, Mosalam K (2019) StEER - Hurricane Dorian: Field Assessment Structural Team (FAST-1) Early Access Reconnaissance Report (EARR). DesignSafe-CI. https://doi.org/10.17603/ds2-4616-1e25.,
McNicholas PD, Murphy TB (2008) Parsimonious Gaussian mixture models. Statistics and Computing 18 (3):285-296
Moon TK (1996) The expectation-maximization algorithm. IEEE Signal Processing Magazine 13 (6):47-60
Nadal-Caraballo NC, Campbell MO, Gonzalez VM, Torres MJ, Melby JA, Taflanidis AA (2020) Coastal Hazards System: A Probabilistic Coastal Hazard Analysis Framework. Journal of Coastal Research 95 (sp1):1211-1216
Pasch RJ, Roberts DP, Blake ES (2020) The 2019 Atlantic Hurricane Season: An Active and Destructive Year. Weatherwise 73 (3):32-39
Peng M, Xie L, Pietrafesa LJ (2004) A numerical study of storm surge and inundation in the Croatan–Albemarle–Pamlico Estuary System. Estuarine, Coastal and Shelf Science 59 (1):121-137
Ramos-Valle AN, Curchitser EN, Bruyère CL (2020) Impact of tropical cyclone landfall angle on storm surge along the Mid-Atlantic bight. Journal of Geophysical Research: Atmospheres 125 (4):e2019JD031796
Resio DT, Powell NJ, Cialone MA, Das HS, Westerink JJ (2017) Quantifying impacts of forecast uncertainties on predicted storm surges. Natural Hazards 88 (3):1423-1449
Robertson IN, Riggs HR, Yim SC, Young YL (2007) Lessons from Hurricane Katrina storm surge on bridges and buildings. Journal of Waterway, Port, Coastal, and Ocean Engineering 133 (6):463-483
Rohmer J, Lecacheux S, Pedreros R, Quetelard H, Bonnardot F, Idier D (2016) Dynamic parameter sensitivity in numerical modelling of cyclone-induced waves: a multi-look approach using advanced meta-modelling techniques. Natural Hazards 84 (3):1765-1792
Saltelli A (2002) Sensitivity analysis for importance assessment. Risk Analysis 22 (3):579-590
Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications 181 (2):259-270
Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) Global sensitivity analysis: the primer. John Wiley & Sons,
Sobol' IM (1990) On sensitivity estimation for nonlinear mathematical models. Matematicheskoe Modelirovanie 2 (1):112-118
Sobol' IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and computers in simulation 55 (1-3):271-280
Sochala P, Chen C, Dawson C, Iskandarani M (2020) A polynomial chaos framework for probabilistic predictions of storm surge events. Computational Geosciences 24 (1):109-128
Sudret B (2008) Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering & System Safety 93 (7):964-979
Taylor AA, Glahn B (2008) Probabilistic guidance for hurricane storm surge. In: 19th Conference on probability and statistics.
Vetter C, Taflanidis A (2014) Comparison of alternative stochastic ground motion models for seismic risk characterization. Soil Dynamics and Earthquake Engineering 58:48-65

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)

Wood KM, Klotzbach PJ, Collins JM, Schreck CJ (2019) The record-setting 2018 eastern North Pacific hurricane season. Geophysical Research Letters 46 (16):10072-10081
Zhang C, Li C (2019) Effects of hurricane forward speed and approach angle on storm surges: an idealized numerical experiment. Acta Oceanologica Sinica 38 (7):48-56

Author Personal Copy (DOI: 10.1007/s11069-022-05598-z)