

<https://doi.org/10.1038/s44304-024-00032-9>

Using surrogate modeling to predict storm surge on evolving landscapes under climate change

**Mohammad Ahmadi Gharehboragh¹ & David R. Johnson^{1,2}✉**

Planners managing coastal flood risk under a constrained computational budget face a tradeoff. Simulating many time periods or scenarios limits how many storm simulations can be run on each landscape. In this analysis, we present a deep learning model to predict storm surge as a function of storm parameters but also landscape features and boundary conditions (e.g., sea level). It is trained on peak surge elevations from Advanced Circulation (ADCIRC) hydrodynamic simulations of coastal Louisiana in a 2020 baseline and decadal periods from 2030 to 2070 under two morphological and climate scenarios. Leave-one-landscape-out cross-validation yielded a 0.086-m RMSE and 0.050-m MAE over 90 storms per landscape and 94,013 geospatial locations. A two-sided Kolmogorov-Smirnov test comparing annual exceedance probability (AEP) estimates from the model predictions to ADCIRC simulations rejected the null hypothesis that the predicted and ADCIRC AEP values were drawn from the same distribution only 1.1% of the time.

Storm surge, the temporary increase in sea level caused by severe storms, is one of the most destructive components of tropical cyclones (TC), motivating the development of accurate tools and models for its prediction^{1,2}. In recent decades, high-fidelity numerical models have been produced that are able to estimate storm surge generated by tropical cyclone wind fields with high accuracy. However, these hydrodynamic simulations typically are computationally intense, requiring high-performance computing resources, and still there can be considerable biases in their outcome^{3,4}.

Because physically-based models (i.e., solving fluid dynamics equations) like the ADvanced CIRCulation (ADCIRC) model⁵ can be expensive in terms of computational cost, surrogate models have been developed to predict storm surge without incurring the same high computational cost⁶. The use of surrogate models or meta-models has increased rapidly in the field of coastal flood hazard research⁷. Surrogate models are useful tools in different fields because of their ability to emulate the behavior of complex models in their quest to approximate complex systems. Moreover, their computational efficiency makes them a convenient approach for tasks like optimization or modeling large ensembles of events and scenarios^{8,9}.

Risk assessments, using techniques like joint probability methods to estimate a hazard curve, require simulation of a large number of synthetic storm events (i.e., thousands or tens of thousands). Computational limitations motivated advances such as the joint probability method with optimal sampling (JPM-OS)¹⁰⁻¹³ and the use of heuristic algorithms to further reduce the number of simulations required for probabilistic flood risk^{4,14}.

However, these approaches have limits, and consequently, planning studies still face meaningful constraints imposed by computational budgets. When evaluating the benefits of risk reduction infrastructure with long useful life spans, it is indeed important to consider risk and risk reduction over long planning horizons. Protection systems (e.g., levees, dikes, seawalls, pumping stations) must be designed to withstand and mitigate the effects of extreme events over many decades. This increases the necessity to consider uncertainties in factors that, over time, reshape the coastal landscape (e.g., land subsidence, land-use change, impacts of saltwater intrusion on vegetation) and boundary conditions (e.g., sea level rise (SLR)) that determine risk to coastal communities; scenario analyses examine multiple future states of the world with different realizations of uncertain parameters. Integrated coastal management plans like Louisiana's *Comprehensive Master Plan for a Sustainable Coast* (Coastal Master Plan) evaluate the performance of a range of flood protection and coastal restoration projects implemented in different sequences, necessitating the modeling of multiple future time periods¹⁵. Mokrech et al. (2011)¹⁶ stresses the importance of developing an integrated framework to assess long-term coastal impacts and thus make rational management decisions¹⁶. Wamsley et al. (2009)¹⁷ investigated the storm surge and wave reduction benefits of different environmental restoration features (e.g., marsh restoration and barrier island changes), as well as the impact of future wetland

¹School of Industrial Engineering, Purdue University, West Lafayette, IN, USA. ²Department of Political Science, Purdue University, West Lafayette, IN, USA.

✉ e-mail: davidjohnson@purdue.edu

degradation on local conditions, concluding that “consideration of natural features is required” to properly assess flood risk¹⁷.

Studies that include multiple future time periods, states of the world, and project portfolios must evaluate risk on a large number of landscapes, and simple math dictates that under a fixed computational budget, the more landscapes planners want to model, the fewer events can be simulated per landscape. Using a lower-resolution mesh or a model simulating fewer physical processes may be an undesirable solution if it would introduce unacceptable biases and/or uncertainty in storm surge estimates or compromise the ability of the model to resolve the impact of protection features like levees.

In this paper, we introduce a surrogate model using artificial neural networks (ANN) that can be used to resolve this computational constraint. We train the model on synthetic storms simulated on multiple landscapes using the ADCIRC model, including not only the storm parameters as features, but also landscape features (e.g., topographic/bathymetric elevations, canopy) and boundary conditions (e.g., mean sea level). We evaluate the accuracy to predict peak storm surge elevations, as well as the accuracy of annual exceedance probability (AEP) distributions estimated using the predicted surge values, finding that the model is sufficiently accurate for use as a scenario generator in planning studies.

Previous studies have taken various approaches to applying surrogate models for prediction of storm surge and waves. Commonly, this means predicting peak storm surge elevations and peak significant wave heights at many points on a spatial grid as a function of the storm’s characteristics at landfall. Studies vary in their choice of geography and TC characteristics, with the latter typically including parameters such as landfall location, angle of landfall, central pressure, forward velocity, radius of maximum wind-speed, Holland-B parameter and/or tide level. Techniques for the surrogate models include kriging^{4,18}, kriging combined with principle component analysis^{19,20}, support vector regression²¹, and ANN²². Al Kajbaf and Bensi (2020)²¹ provides a comparative assessment of the performance of these techniques²¹.

Many other studies have focused on the use of surrogate models for forecasting future water surface elevations over the course of a storm^{23–29}. Recent works have incorporated SLR into predictions on a static landscape⁶, but landscape morphology plays a significant role in modeling flood inundation and flood risk accurately^{30,31}. In areas exhibiting substantial land subsidence, erosion, barrier island migration, and other phenomena impacting morphology, incorporating SLR is necessary but insufficient for projecting future storm surges and inundation risks. Canopy and vegetation impact wind attenuation and surface friction, as shown in studies of mangrove forests and coastal wetlands on the Gulf coast of South Florida³². Mangrove forests reduced storm inundation areas and restricted surge inundation within a Category 3 hurricane zone, according to the study, finding that the width of the mangrove zone had a nonlinear effect on reducing surge amplitudes.

Although these previous studies investigated storm surge surrogate modeling from other perspectives, the impact of the combination of SLR, landscape and TC parameters on storm surge has not been thoroughly investigated. In this study, we aim to fill that gap by developing a surrogate model using deep neural networks for the prediction of peak storm surge elevations from synthetic TCs as a function of their characteristics at landfall in coastal Louisiana, four landscape parameters impacting storm surge, and mean sea level.

The synthetic TCs used in this study are characterized by their overall tracks and five parameters at landfall: forward velocity, radius of maximum windspeed, central pressure, landfall coordinates, and heading. The corpus of 645 synthetic storms was developed by the US Army Corps of Engineers for use in flood risk assessments based on the JPM-OS methodology³³; each synthetic storm’s landfall parameters serve as input data for the predictive model and are provided in Supplementary Information Table S1. We utilize only the landfall parameters because of the synthetic (i.e., idealized) nature of the storm tracks and parameters; synthetic TCs follow their heading at landfall and exhibit similar patterns of decay in intensity as they move

inland, so the variability in synthetic storm behavior and potential to generate storm surge is reasonably captured by their characteristics at landfall. Note that this means our approach would likely yield lower accuracy if the trained models were asked to infer peak storm surge elevations from historic or more realistic TCs.

Hydrodynamic simulations from a coupled ADCIRC + SWAN model were available from Louisiana’s 2023 Coastal Master Plan for all 645 synthetic storms, simulated on the plan’s “Existing Conditions” landscape (i.e., 2020). A subset of 90 synthetic storms was simulated on each of 10 future landscapes representing decadal snapshots (i.e., 2030, 2040, 2050, 2060, 2070) under two different scenarios, a Lower and Higher Scenario, that vary in their assumptions about the rate of SLR, land subsidence, and other environmental factors^{15,34}. For each synthetic storm and landscape, peak storm surge elevations were extracted from the ADCIRC + SWAN simulations at 94,013 locations representing grid points from the Coastal Louisiana Risk Assessment model (CLARA) not located within fully-enclosed protection systems; such points were excluded from analysis because of the additional complexity of flood dynamics in enclosed polders (e.g., overtopping, levee fragility/breaches, pumping, rainfall). The points form a mixed-resolution grid consisting of a regular 1-km resolution grid, with more points added so that every 2010 U.S. Census block contains at least one grid point³⁵. This has the effect of adding higher resolution in densely populated areas. Grid cell polygons are associated with each grid point as Thiessen polygons formed within each census block (i.e., polygons consisting of all pixels in the census block closer to a given grid point than any other in the same block).

Each landscape is characterized by a digital elevation model defining the topography and bathymetry of the study region, as well as rasters defining other inputs to the ADCIRC model (with a resolution of 30 meters): the Manning’s n value (i.e., bottom roughness coefficient), free surface roughness z_0 , and a surface canopy coefficient that captures the reduction in wind stress on water surfaces produced by local vegetation. All landscape characteristics were represented as GeoTIFFs with values extracted at each of the 94,013 grid point locations for use in the surrogate model. Full details regarding the Integrated Compartment Model used to develop the landscape representations are found in White et al. (2019)³⁶ and Reed and White (2023)³⁷, and details regarding the ADCIRC + SWAN model and Louisiana mesh are found in Cobell and Roberts (2021)³⁴ and Roberts and Cobell (2017)³⁸. Mean sea level (NAVD88 m) assumed for each scenario as a boundary condition is provided in Table 1.

Methods

This study used feed-forward ANN models with multilayers and multiple outputs to predict storm surge at each location under current and future landscape conditions. A range of models varying from 128 neurons to 256 neurons was evaluated before selecting the models described here, as specifying too few neurons could impede the learning process while using too many neurons could result in overtraining/overfitting³⁹. Moreover, for all hidden layers, the RELU activation function was chosen with a learning rate of 0.001, and for the last layer, a linear activation function was selected to predict surge values.

Firstly, we examined the value of including landscape parameters in a predictive model of storm surge for a single landscape only. ANN models were trained for current conditions on 645 synthetic storms: a multi-layered feed-forward architecture with four hidden layers and an output layer of 1 dimension was used for predicting peak storm surge elevations at different locations. A “storm-only model” at each location only included the synthetic storm parameters at landfall as inputs. The “full model” included the synthetic storm parameters but also all grid points’ landscape parameters from the current landscape (latitudinal and longitudinal coordinates, topo/bathy NAVD88 elevation, surface canopy, z_0 , and Manning’s n). Sea level was excluded from the full model in this test because the local mean sea level was assumed constant throughout the study region, and thus only has variation when multiple landscapes are taken in as input data. For the storm-only and full models, predicted accuracy is reported for 15% of storms randomly held

Table 1 | Summary of mean sea level assumptions and statistical outcomes for all cases evaluated

Scenario	Year	Mean Sea Level (NAVD88 m)	RMSE (m)	MAE (m)	Correlation	Rejected %
Storm-only Model	2020	0.36	0.314	0.172	0.912	-
Full Model	2020	0.36	0.277	0.077	0.965	-
Higher Scenario	2030	0.46	0.063	0.036	0.998	0.76%
	2040	0.58	0.069	0.037	0.998	0.69%
	2050	0.73	0.076	0.041	0.997	0.56%
	2060	0.92	0.082	0.047	0.997	0.40%
	2070	1.13	0.206	0.134	0.983	4.84%
Lower Scenario	2030	0.44	0.057	0.035	0.998	0.74%
	2040	0.52	0.089	0.044	0.996	1.43%
	2050	0.62	0.064	0.037	0.998	0.46%
	2060	0.73	0.073	0.04	0.997	0.53%
	2070	0.86	0.081	0.045	0.997	0.81%

out in a cross-validation procedure with the remaining 85% of storms used for training and validation.

Next, to investigate the impacts of climate change and the slowly evolving landscape, we trained the full model using the 645 synthetic storms from the 2020 landscape condition, as well as the 10 future landscapes, each with the same 90 synthetic storms. Predictive accuracy of the full model was evaluated utilizing leave-one-landscape-out cross-validation (LOOCV) on the future landscapes; in other words, for each fold of the CV procedure, the model was trained on the 2020 landscape and 9 of the 10 future landscapes ($n = 1455$ storms), with predictions made on the tenth future landscape ($n = 90$ storms). We did this to reflect a real-world use case in which the full model could serve as a scenario generator, training on a set of landscapes run through ADCIRC and then predicting outcomes in novel landscapes. The current conditions landscape's 645 storms were included to represent a realistic case in which a larger suite of synthetic TCs could be run on a single landscape as an input to a storm selection process that would identify the subset of 90 storms to run on other landscapes. Using 100 epochs to train the model, the entire process was executed on an AMD Epyc 7662 CPU at 2.0 GHz, taking less than 7 h for training to be completed in preparation for making predictions on a new landscape. For all folds in the cross-validation process, it took 70 h. Once trained, less than 4 min is needed to generate predictions for a novel landscape.

Finally, we also wanted to know how errors in the predicted peak storm surge from each synthetic storm propagate to differences in the estimated annual probability distribution of experiencing storm surge of varying elevations. This is ultimately what planners may care about when making decisions about flood protection projects. For this task, we employed the Coastal Louisiana Risk Assessment (CLARA) model, an implementation of JPM-OS which is the model used to estimate flood hazard for Louisiana's Coastal Master Plan^{35,40}. Full details on the CLARA model's methodology are in Johnson et al. (2023)³⁵; in this analysis, we compared peak surge elevation exceedance curves (i.e., surge elevations as a function of AEP) generated from the simulated surge elevations from ADCIRC to exceedance curves generated from the predicted surge elevations from the LOOCV procedure. The resulting empirical distributions were compared using a two-sample Kolmogorov-Smirnov (K-S) test, which calculates the maximum difference between two empirical samples' cumulative distribution functions to test a null hypothesis that they have been drawn from the same underlying probability distribution function⁴¹. CLARA produces estimates of surge exceedances at 23 return periods ranging from a 50% AEP to 0.005% AEP (i.e., the 2-year event to the 2000-year event), so the two-sample KS test dictates that the null hypothesis be rejected at significance level α if

$$\sup_x |F_{\text{ADCIRC}}(x) - F_{\text{ANN}}(x)| > \sqrt{-\ln\left(\frac{\alpha}{2}\right)} \cdot \frac{1}{23} \quad (1)$$

where \sup is the supremum over x , $F_{\text{ADCIRC}}(x)$ and $F_{\text{ANN}}(x)$ are the sample CDFs associated with the ADCIRC simulations and ANN predictions, respectively.

Results

The ANN model that includes landscape parameters performs markedly better than the model with only storm parameters when predicting surge from relatively intense storms, as shown for two illustrative storms in Fig. 1. Each pane plots ADCIRC-simulated values against the ANN-predicted values at ~3500 points on a west-to-east transect at 29.8° N, a latitude selected for its nearly continuous series of grid points uninterrupted by major water bodies or enclosed polders. It is worth to mention that the sudden decrease in points with simulated surge below 0.36 NAVD88 m is due to this being the mean sea level assumed for the current conditions landscape. Grid points over water are initialized at this level, leaving very few points along the chosen transect with lower topographic elevation, typically due to being pumped and drained. Blue points represent the Full Model which includes landscape parameters, and red points represent the Storm-Only Model which excludes them. The left-hand pane shows Storm 495, a weaker TC with central pressure of 975 mb at landfall, while the right-hand pane shows the much stronger storm 11 with a landfalling central pressure of 905 mb (landfall parameters for all 645 synthetic storms are provided as Supplementary Table 1).

Across all 645 synthetic storms and grid points in the current conditions landscape, the Storm-only Model reached an overall RMSE of 0.31 m, while the Full Model achieved an RMSE of 0.28 m (Table 1). While this does represent an improvement of over 10 percent, primarily the result of greater accuracy for larger surge elevations, we expected the difference between these models to be minimal when trained only on the current conditions landscape. This is because of the lack of variation in landscape parameters over the synthetic storms at each point and contrastingly greater variation in TC parameters.

Examining the spatial pattern of the Full Model when trained on current and future scenarios, we see that points with higher RMSE over all storms and landscapes are generally further inland (Fig. 2). This is expected, given that such points generally have fewer storms in the corpus that produce wetting, and we did not employ any dry-node correction techniques like those used in ref. 42 or ref. 18; instead, non-wetting observations were simply removed from the training set. The model also performed less accurately in areas with more complex topography and hydrodynamics, such as in unpopulated wetlands in the Atchafalaya River Basin (between 91° and 92° W longitude on the northern portion of the model domain), where the ADCIRC model also has greater uncertainty and bias when validated against historic TCs³⁴.

Considering the RMSE of the Full Model averaged over all landscapes and synthetic storms, the RMSE at 90% of grid points is less than 0.18 m, at

Fig. 1 | Simulated versus predicted storm surge for representative synthetic storms. Simulated versus predicted storm surge for storm 495 (left pane) and storm 11 (right pane) in the cases where the ANN model input includes only storm parameters (red) and both storm and landscape parameters (blue) grid points along a transect at 29.8° N. Note: Axis ranges vary between left and right panes.

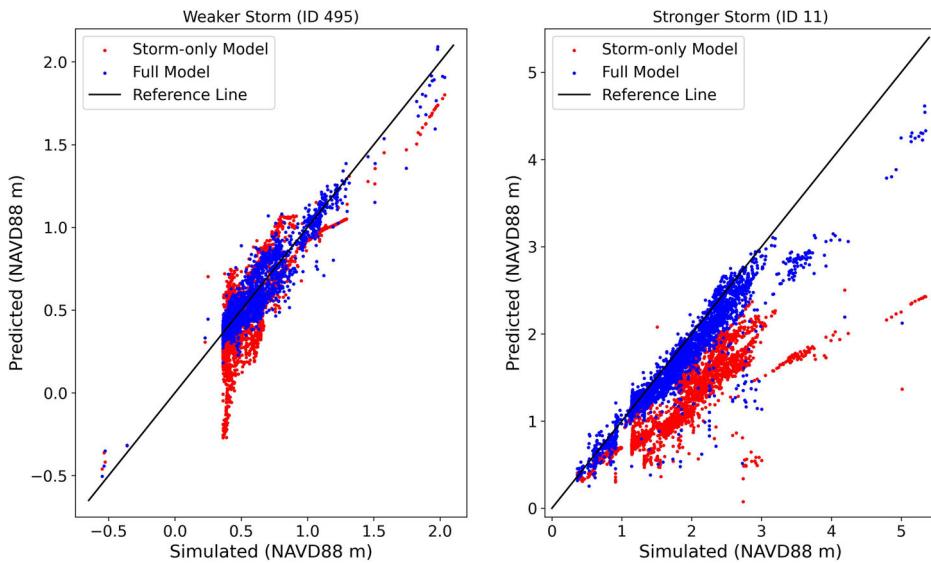
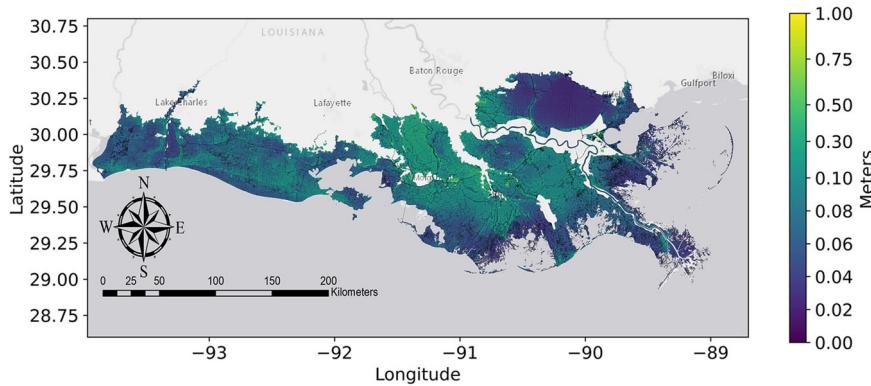


Fig. 2 | RMSE values across all landscapes and synthetic storms for all grid points in the study domain. Darker colors indicate lower RMSE values, i.e., locations of greater agreement between simulated and predicted storm surge.



99% of grid points less than 0.38 m, and at 99.9% of grid points less than 0.79 m (Fig. 3). Over the ten future scenarios used in the LOOCV procedure, the Full Model produced a grand RMSE of 0.086 m and grand mean absolute error (MAE) of 0.050 m (Table 1). These results compare favorably to the calibration and validation results from the ADCIRC + SWAN model used to generate the hydrodynamic simulations, which reported a standard error in simulated high-water marks of ~0.46 m over seven historical storms (hurricanes Katrina, Rita, Gustav, Ike, Isaac, Nate, and Harvey)³⁴. Further analysis incorporated into the 2023 Coastal Master Plan estimated an average standard error of 0.15 m in peak surge elevations over the grid cells included in this analysis (authors' own calculations).

Figure 4 further disaggregates the model predictions to show the frequency distribution of predicted versus simulated surge elevations in each of the future landscapes over the synthetic storms and grid points. The overall distributions appear nearly indistinguishable except in the Higher scenario's 2070 landscape, the most extreme scenario with respect to its assumptions about mean sea level and cumulative land subsidence. That this scenario would be an outlier compared to the others is intuitive, given its more extreme assumptions about environmental conditions; in this sense, the 2070 Higher scenario landscape is subject to the common difficulty of extrapolating beyond training data in the leave-one-landscape-out experimental design. That said, the directionality of the difference is somewhat counterintuitive. In this scenario, the predicted storm surge is on average, greater than the simulated values, though the primary non-linear difference in the scenario is an accelerating rate of SLR. Despite

this acceleration, it appears that the Full Model overestimates the gradient in storm surge associated with changes in mean sea levels.

To fully understand the RMSE error over a range of storm surge levels with respect to a fixed datum and relative to prevailing ground, Fig. 5 and Fig. 6 are provided. They show the RMSE over all grid points by surge elevation relative to NAVD88 and above prevailing ground, respectively. Prevailing ground elevations are calculated as the median topographic elevation of land pixels in each grid cell polygon. In grid cells containing open water, such as over Lake Pontchartrain, the prevailing elevations used are the mean sea level associated with the specific landscape, representing the prevailing water surface elevation storm surge would build upon. Simulated ADCIRC observations are binned into 0.1-m intervals with respect to their surge elevations or depths, and RMSE of predictions for the locations and storms in each bin are shown on the vertical axis. Based on Fig. 5, it can be seen that all scenarios show approximately similar performance, except for 2070 of the Higher scenario, which has slightly larger RMSE values. Additionally, it can be observed that larger surge elevation values do not necessarily correspond to larger RMSE values. Moreover, the overall performance of all scenarios in the lower scenarios was similar, with a maximum RMSE of 0.4 m and a mean RMSE of less than 0.1 m. Furthermore, in the higher scenario, except for Scenario 2070, the rest have a similar mean RMSE of less than 0.1 m. Generally, the model is least accurate for smaller surge elevations, meaning that extremes which would cause more damage are still captured relatively well.

This is further confirmed by Fig. 6, which is similar to Fig. 5 but with the bins on the horizontal axis representing surge elevation relative to prevailing

Fig. 3 | Exceedance percentage of RMSE values by grid point. Each point represents the RMSE for a particular spatial grid point, averaged across all landscapes and synthetic storms.

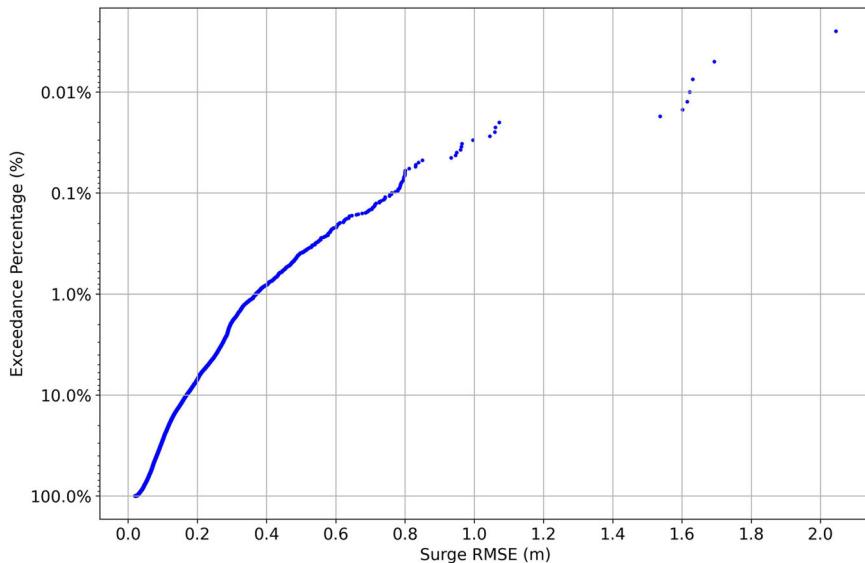
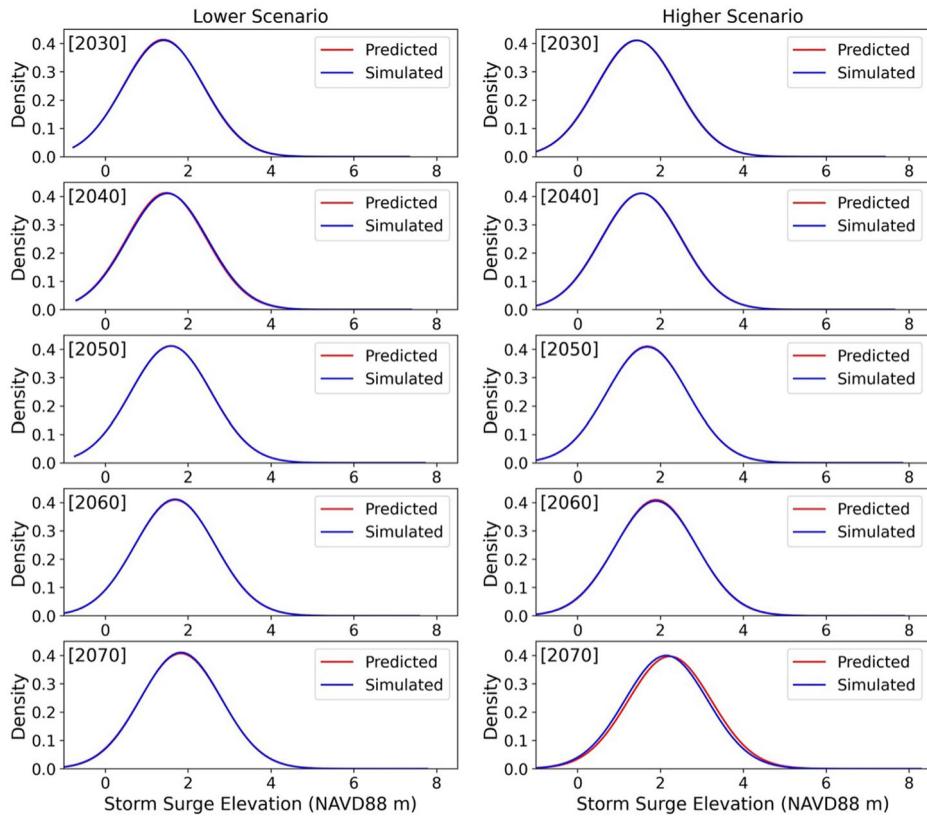


Fig. 4 | Frequency distribution of peak storm surge elevation for all future landscapes. Red indicates the distribution of predicted values, while blue indicates the distribution of simulated values.



ground elevation (i.e., surge depth instead of surge elevation relative to a fixed datum). We see in these results that most scenarios exhibit an RMSE below 0.1 m consistently for surge elevations above prevailing ground. In general, the model predictions are actually less accurate for storm surge values below prevailing ground, likely due to a relatively smaller set of observations in the training data. Based on Fig. 5 and Fig. 6, it can be concluded that for most scenarios, the performance of the model is acceptable across a range of surge elevations, including extreme values which would cause more damage.

Examining the hazard aggregated over multiple TCs, the errors associated with surge predictions do not appear to meaningfully compound once aggregated to AEP curves, in the sense that the RMSEs over all grid points at a range of return periods are in a similar range to the RMSEs over all grid points and synthetic storms (between 0.05 and 0.1 m for all landscapes but the most extreme, as shown in Fig. 7). The RMSE generally is larger at lower AEPs, consistent with an intuition that prediction is more challenging for extreme events associated with storm surge values near the upper bounds of observations in the simulated training sets.

Considering the full AEP distribution of storm surge at each point and in each landscape, the two-sample K-S tests further indicate that the surge predictions are accurate enough to usefully inform probabilistic risk studies. In eight of the ten future landscapes, the null hypothesis, that the empirical distributions generated with the ADCIRC simulations and the ANN

predictions are drawn from the same underlying probability distribution, is rejected at level $\alpha = 0.05$ for less than one percent of the grid points (Table 1). This table also reports a MAE over the grid points below 0.05 m and correlation between simulated and predicted values over 0.99 for all landscapes but the 2070 Higher Scenario.

From Table 1, it is evident that by incorporating landscape parameters into the ANN model, storm surge can be predicted accurately for a variety of different scenarios.

Figure 8 highlights the spatial pattern of points that rejected the null hypothesis of the two-sided K-S test for an illustrative landscape, the Higher scenario in 2060. Red points indicate the locations where the test rejects the null hypothesis at $\alpha = 0.05$, while the blue points indicate locations where the evidence fails to rule out the possibility of the hazard estimates coming from the same underlying AEP distribution.

Discussion

We have presented a machine learning-based surrogate model of peak storm surge elevations that yields predictions of comparable or greater accuracy, when compared to ADCIRC simulations, than the ADCIRC model relative to historic observations used for model calibration and validation. The addition of future landscapes with variation in landscape parameters and mean sea level conditions provides more features and heterogeneity in training data to improve the model's accuracy. In a

Fig. 5 | RMSE over all grid points and storms, as a function of surge elevation. Each line represents a different year (varying by color) of the Lower (left pane) and Higher (right pane) scenarios.

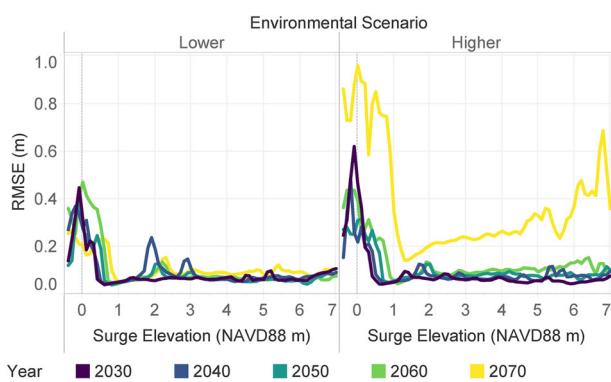


Fig. 6 | RMSE over all grid points and storms, as a function of surge depth relative to prevailing ground. Each line represents a different year (varying by color) of the Lower (left pane) and Higher (right pane) scenarios.

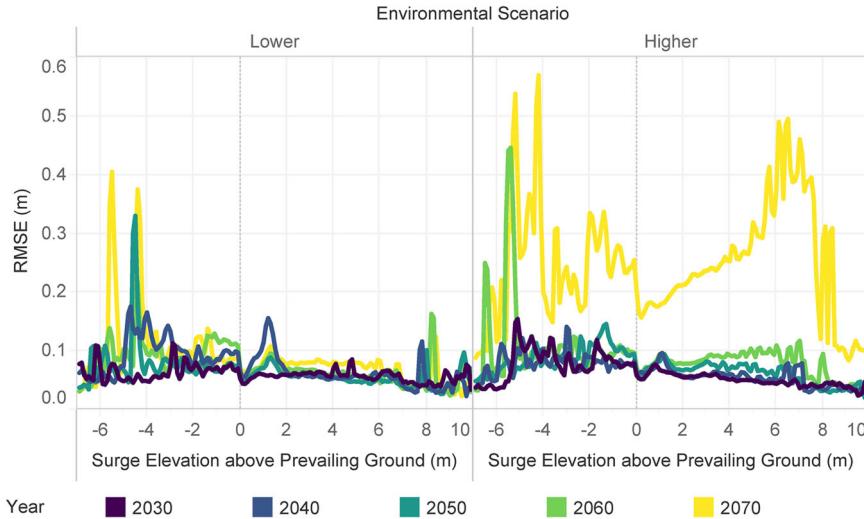


Fig. 7 | RMSE over all grid points, by annual exceedance probability and landscape. Each line represents a different year (varying by color) of the Lower (left pane) and Higher (right pane) scenarios.

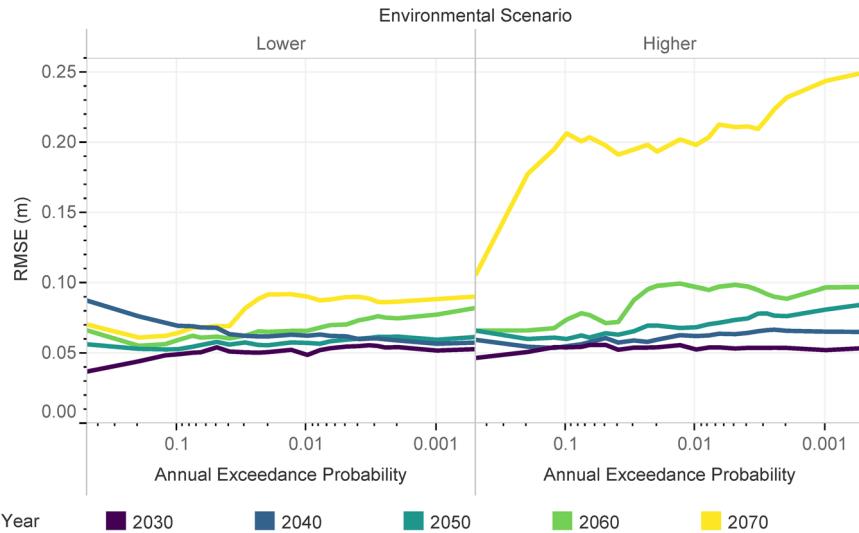
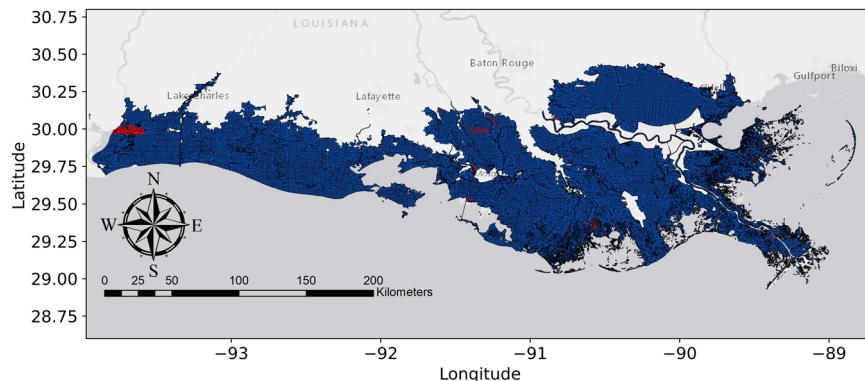


Fig. 8 | Results of two-sample K-S test for the year 2060 of the Higher scenario. Red points indicate locations where the null hypothesis is rejected.



LOOCV exercise, the model produced a MAE of ~ 0.04 m in nine out of ten of the future landscapes, with the exception being the year 2070 of the Higher scenario, the most extreme landscape with respect to having the greatest SLR and land subsidence. A two-sided K-S test failed to reject the null hypothesis, that points on hazard curves generated by the ADCIRC simulations and ANN predictions are drawn from the same underlying distribution, at less than 1% of grid points in eight out of ten of the future landscapes. It is worth mentioning that, in addition to the capabilities of the developed surrogate models, part of the residual sensitivity is attributable to variability in the training dataset being used; as such, the accuracy of surrogate models trained on simulation data from ADCIRC will be limited by the accuracy of the calibrated and validated simulations.

This highlights an important caveat: this analysis utilized ADCIRC simulations that were readily available from Louisiana's 2023 Coastal Master Plan, meaning that the scenarios and time periods were not chosen with the idea of using ADCIRC to train a surrogate model already in mind. This work therefore represents a proof of concept where the ANN model produced predictions suitable for planning studies from a training set of convenience. If planners are interested in estimating risk over a 50-year planning horizon ending in 2070, it may be that accuracy could be improved for the same computational cost by replacing one of the "intermediate" landscapes with a landscape corresponding to the year 2080 instead, to mitigate the challenges of ML models in extrapolating beyond data in their training set.

This also has implications for the storm selection process, given that the 90 storms simulated in each future landscape were chosen by comparing hazard curves to the curve associated with the full 645-storm suite in the current conditions landscape⁴³. Prior research has suggested a difficulty in using ML methods to predict extreme storm surge elevations^{44,45}. While accurate reproduction of extreme individual events is important in controlling the overall RMSE and MAE of the model, extreme storms (i.e., with lower central pressures at landfall) are relatively more rare in occurrence, thus having smaller probability masses when contributing to AEP distributions and making smaller contributions to expected annual damage calculations.

Consequently, adoption of surrogate models as scenario generators would also benefit from a rigorous consideration of how optimal sampling techniques could be extended to include heterogeneity in landscape parameters and boundary conditions. When planning an analysis that will span a range of future states of the world, it is likely that greater computational efficiency could be achieved by sampling different synthetic storm events on each landscape, rather than simulating the same 90 storms as was done for the Coastal Master Plan. We note that in this analysis, the surrogate model does benefit from predicting storm surge on the same 90 storms in the left-out landscape that were present in the training data from other landscapes during the LOOCV procedure. Accuracy would almost certainly be lower if called upon to predict storm surge in both novel landscapes and for storms with novel landfall parameters.

All the future landscapes used for training our ANN came from the Coastal Master Plan's Future Without Action scenarios (i.e., no additional

projects implemented on the landscape). This means we have restricted our predictions to analyzing a slowly evolving landscape without major coastal management interventions. As such, developing a surrogate modeling framework to capture the impacts of projects such as constructing or upgrading levees and floodwalls is a more challenging problem and subject for future research. However, the surrogate model developed in this study would also have utility in evaluating the flood risk impacts of coastal restoration projects that affect landscape morphology over time scales ranging from the immediate (e.g., beach nourishment) to decadal (e.g., river diversions).

This study enables better modeling of future climate and environmental conditions by policy makers and water resource managers. Moreover, the developed model makes it possible to evaluate risk under a greater number and range of future scenarios and time periods, opening the door to the use of computationally expensive models like ADCIRC for planning studies utilizing techniques for decision-making under deep uncertainty (DMDU) that require or benefit from the use of large ensembles of future states of the world⁴⁶. We note, though, that not all institutions have the resources and planning capacity to adopt DMDU approaches or to generate the same number of landscapes used for training data in this study. Even with a more limited set of future scenarios, the concept introduced here, of incorporating landscape and boundary condition features to train a surrogate model over multiple futures, could be applied to improve risk assessments that currently employ simplifying assumptions such as linearly interpolating storm surge or hazard over time or applying a bathtub model of SLR.

Data availability

Data for this project are available at DesignSafe-CI, <https://doi.org/10.17603/ds2-0ksb-yy40>.

Code availability

Model source code, written in Python 3 and R, is available at https://github.com/mohammadahmadi1995/Storm_Surge.

Received: 24 December 2023; Accepted: 21 October 2024;

Published online: 14 November 2024

References

1. Bilskie, M. V. et al. Dynamic simulation and numerical analysis of hurricane storm surge under sea level rise with geomorphologic changes along the northern Gulf of Mexico. *Earth's Future* **4**, 177–193 (2016).
2. Jia, G. et al. Surrogate modeling for peak or time-dependent storm surge prediction over an extended coastal region using an existing database of synthetic storms. *Nat. Hazards* **81**, 909–938 (2015).
3. Plumlee, M., Asher, T. G., Chang, W. & Bilskie, M. V. High-fidelity hurricane surge forecasting using emulation and sequential experiments. *Ann. Appl. Stat.* **15**, 460–480 (2021).
4. Zhang, J., Taflanidis, A. A., Nadal-Caraballo, N. C., Melby, J. A. & Diop, F. Advances in surrogate modeling for storm surge prediction: storm

selection and addressing characteristics related to climate change. *Nat. Hazards* **94**, 1225–1253 (2018).

5. Luettich, R. A., Westerink, J. J. & Scheffner, N. W. *ADCIRC: an Advanced Three-dimensional Circulation Model for Shelves, Coasts, and Estuaries*. Report 1, Theory and methodology of ADCIRC-2DD1 and ADCIRC-3DL (U.S. Army Engineer Waterways Experiment Station, 1992).
6. Kyprioti, A. P., Taflanidis, A. A., Nadal-Caraballo, N. C. & Campbell, M. O. Incorporation of sea level rise in storm surge surrogate modeling. *Nat. Hazards* **105**, 531–563 (2021).
7. Kyprioti, A. P., Taflanidis, A. A., Nadal-Caraballo, N. C., Yawn, M. C. & Aucoin, L. A. Integration of node classification in storm surge surrogate modeling. *J. Mar. Sci. Eng.* **10**, 551 (2022).
8. Bartz-Beielstein, T. & Zaefferer, M. Model-based methods for continuous and discrete global optimization. *Appl. Soft Comput.* **55**, 154–167 (2017).
9. Bekasiewicz, A., Koziel, S. & Pankiewicz, B. Accelerated simulation-driven design optimisation of compact couplers by means of two-level space mapping. *IET Microw. Antennas Propag.* **9**, 618–626 (2015).
10. Resio, D. T. White paper on estimating hurricane inundation probabilities. in *Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System. Appendix 8: Hazard Analysis* Vol. VIII (U.S. Army Corps of Engineers, 2007).
11. Resio, D. T., Irish, J. & Cialone, M. A surge response function approach to coastal hazard assessment—part 1: basic concepts. *Nat. Hazards* **51**, 163–182 (2009).
12. Toro, G. R., Resio, D. T., Divoky, D., Nedoroda, A. W. & Reed, C. Efficient joint-probability methods for hurricane surge frequency analysis. *Ocean Eng.* **37**, 125–134 (2010).
13. Yang, K., Paramygin, V. & Sheng, Y. P. An objective and efficient method for estimating probabilistic coastal inundation hazards. *Nat. Hazards* **99**, 1105–1130 (2019).
14. Fischbach, J. R., Johnson, D. R. & Kuhn, K. Bias and efficiency tradeoffs in the selection of storm suites used to estimate flood risk. *J. Mar. Sci. Eng.* **4**, 10 (2016).
15. Coastal Protection and Restoration Authority. *Coastal Master Plan: Appendix B: Scenario Development and Future Conditions*. **18**, https://coastal.la.gov/wp-content/uploads/2023/01/E1_2023StormSurge_Waves_Report_Jan2021.pdf (2023).
16. Mokrech, M. et al. The Tyndall coastal simulator. *J. Coast. Conserv.* **15**, 325–335 (2011).
17. Wamsley, T. V., Cialone, M. A., Smith, J. M., Ebersole, B. A. & Grzegorzewski, A. S. Influence of landscape restoration and degradation on storm surge and waves in southern Louisiana. *Nat. Hazards* **51**, 207–224 (2009).
18. Kyprioti, A. P. et al. Improvements in storm surge surrogate modeling for synthetic storm parameterization, node condition classification and implementation to small size databases. *Nat. Hazards* **109**, 1349–1386 (2021).
19. Jia, G. et al. Surrogate modeling for peak or time-dependent storm surge prediction over an extended coastal region using an existing database of synthetic storms. *Nat. Hazards* **81**, 909–938 (2016).
20. Jia, G. & Taflanidis, A. A. Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment. *Comput. Methods Appl. Mech. Eng.* **261–262**, 24–38 (2013).
21. Al Kajbaf, A. & Bensi, M. Application of surrogate models in estimation of storm surge: a comparative assessment. *Appl. Soft Comput.* **91**, 106184 (2020).
22. Chen, W.-B., Liu, W.-C. & Hsu, M.-H. Predicting typhoon-induced storm surge tide with a two-dimensional hydrodynamic model and artificial neural network model. *Nat. Hazards Earth Syst. Sci.* **12**, 3799–3809 (2012).
23. De Oliveira, M. M. F., Ebecken, N. F. F., De Oliveira, J. L. F. & De Azevedo Santos, I. Neural network model to predict a storm surge. *J. Appl. Meteorol. Climatol.* **48**, 143–155 (2009).
24. Kim, S., Pan, S. & Mase, H. Artificial neural network-based storm surge forecast model: practical application to Sakai Minato, Japan. *Appl. Ocean Res.* **91**, 101871 (2019).
25. Lee, J.-W., Irish, J. L., Bensi, M. T. & Marcy, D. C. Rapid prediction of peak storm surge from tropical cyclone track time series using machine learning. *Coast. Eng.* **170**, 104024 (2021).
26. Lee, T.-L. Neural network prediction of a storm surge. *Ocean Eng.* **33**, 483–494 (2006).
27. Lee, T.-L. Predictions of typhoon storm surge in Taiwan using artificial neural networks. *Adv. Eng. Softw.* **40**, 1200–1206 (2009).
28. Rajasekaran, S., Gayathri, S. & Lee, T.-L. Support vector regression methodology for storm surge predictions. *Ocean Eng.* **35**, 1578–1587 (2008).
29. Sztobryn, M. Forecast of storm surge by means of artificial neural network. *J. Sea Res.* **49**, 317–322 (2003).
30. Bates, P. D., Horritt, M. S. & Fewtrell, T. J. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *J. Hydrol.* **387**, 33–45 (2010).
31. Bilskie, M. V., Hagen, S. C., Medeiros, S. C. & Passeri, D. L. Dynamics of sea level rise and coastal flooding on a changing landscape. *Geophys. Res. Lett.* **41**, 927–934 (2014).
32. Zhang, K. et al. The role of mangroves in attenuating storm surges. *Estuar. Coast. Shelf Sci.* **102–103**, 11–23 (2012).
33. Nadal-Caraballo, N. C. et al. Coastal hazards system: a probabilistic coastal hazard analysis framework. *J. Coast. Res.* **95**, 1211–1216 (2020).
34. Cobell, Z. & Roberts, H. J. *Coastal Master Plan: Attachment E1: Storm Surge and Waves Model Improvements*. **56**, https://coastal.la.gov/wp-content/uploads/2023/05/B_ScenarioDevelopmentFutureConditions_Jan2023_v3.pdf (2021).
35. Johnson, D. R. et al. *Coastal Master Plan: Attachment C11: 2023 Risk Model*. **33**, https://coastal.la.gov/wp-content/uploads/2023/04/C11_2023RiskModel_Jan2023_v2.pdf (2023).
36. White, E. D., Reed, D. J. & Meselhe, E. A. Modeled sediment availability, deposition, and decadal land change in coastal Louisiana marshes under future relative sea level rise scenarios. *Wetlands* **39**, 1233–1248 (2019).
37. Reed, D. J. & White, E. D. *Appendix C: Use of Predictive Models in the 2023 Coastal Master Plan*. Baton Rouge, Louisiana, Louisiana Coastal Protection and Restoration Authority (2023).
38. Roberts, H. J. & Cobell, Z. *Coastal Master Plan Modeling: Attachment C3-25.1—Storm Surge 110*. Baton Rouge, Louisiana, Louisiana Coastal Protection and Restoration Authority (2017).
39. Jammoussi, I. & Ben Nasr, M. A hybrid method based on extreme learning machine and self organizing map for pattern classification. *Comput. Intell. Neurosci.* **2020**, 1–9 (2020).
40. Johnson, D. R., Fischbach, J. R. & Ortiz, D. S. Estimating surge-based flood risk with the coastal Louisiana risk assessment model. *J. Coast. Res.* **109–126**. https://doi.org/10.2112/SI_67_8 (2013).
41. Smirnov, N. V. Estimate of deviation between empirical distribution functions in two independent samples. *Bull. Mosc. Univ.* **2**, 3–16 (1939).
42. Shisler, M. P. & Johnson, D. R. Comparison of methods for imputing non-wetting storm surge to improve hazard characterization. *Water* **12**, 1420 (2020).
43. Fischbach, J. R., Johnson, D. R., Wilson, M. T., Geldner, N. B. & Stelzner, C. *2023 Coastal Master Plan: Model Improvement Report, Risk Assessment*. Baton Rouge, Louisiana, Louisiana Coastal Protection and Restoration Authority (2021).
44. Hashemi, M. R., Spaulding, M. L., Shaw, A., Farhadi, H. & Lewis, M. An efficient artificial intelligence model for prediction of tropical storm surge. *Nat. Hazards* **82**, 471–491 (2016).
45. Ramos-Valle, A. N., Curchitser, E. N., Bruyère, C. L. & McOwen, S. Implementation of an artificial neural network for storm surge forecasting. *J. Geophys. Res. Atmos.* **126**, e2020JD033266 (2021).

46. Johnson, D. R. & Geldner, N. B. Contemporary decision methods for agricultural, environmental, and resource management and policy. *Annu. Rev. Resour. Econ.* **11**, 19–41 (2019).

Acknowledgements

This work was supported by the U.S. National Science Foundation under awards 2238060 and 2118329. We thank members of the modeling teams for Louisiana's Comprehensive Master Plan for a Sustainable Coast and funding from the Coastal Protection and Restoration Authority, under the 2023 Coastal Master Plan's Master Services Agreement, for the original simulation data used in this study. Any opinions, findings, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding entities.

Author contributions

D.R.J. designed the conceptual framework and analysis plan and led the writing of the paper. M.A.G. developed, debugged, and refined the model, executed the analysis, and contributed to writing the paper. Both authors co-produced figures and reviewed the completed manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44304-024-00032-9>.

Correspondence and requests for materials should be addressed to David R. Johnson.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024