ADVANCING EARTH AND SPACE SCIENCE

JGR Atmospheres

RESEARCH ARTICLE
10.1029/2022JD037617

Using Neural Networks to Predict Hurricane Storm Surge and to Assess the Sensitivity of Surge to Storm Characteristics

Joseph W. Lockwood[1], Ning Lin[2], Michael Oppenheimer[1,3,4], and Ching-Yao Lai[1,5]

[1]Department of Geoscience, Princeton University, Princeton, NJ, USA, [2]Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA, [3]Princeton School of Public and International Affairs, Princeton University, Princeton, NJ, USA, [4]High Meadows Environmental Institute, Princeton, NJ, USA, [5]Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA

Key Points:

- Artificial Neural network (ANN) models trained with synthetic datasets are able to predict storm surge levels along the US East and Gulf Coasts
- ANN models accurately predict storm surge levels and unlike earlier studies, do not underestimate extreme levels
- Increases in hurricane translation speeds amplify (reduce) storm surge levels in open ocean (semi-enclosed) regions

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. W. Lockwood,
jl115@princeton.edu

Citation:

Lockwood, J. W., Lin, N., Oppenheimer, M., & Lai, C.-Y. (2022). Using neural networks to predict hurricane storm surge and to assess the sensitivity of surge to storm characteristics. *Journal of Geophysical Research: Atmospheres*, *127*, e2022JD037617. https://doi.org/10.1029/2022JD037617

Received 7 AUG 2022
Accepted 10 DEC 2022

Author Contributions:

Conceptualization: Joseph W. Lockwood, Ning Lin, Michael Oppenheimer, Ching-Yao Lai
Data curation: Joseph W. Lockwood, Ning Lin, Ching-Yao Lai
Formal analysis: Joseph W. Lockwood
Funding acquisition: Ning Lin, Michael Oppenheimer, Ching-Yao Lai
Investigation: Joseph W. Lockwood
Methodology: Joseph W. Lockwood, Ning Lin, Michael Oppenheimer, Ching-Yao Lai

Abstract Hurricane storm surge represents a significant threat to coastal communities around the world. Here, we use artificial neural network (ANN) models to predict storm surge levels using hurricane characteristics along the US Gulf and East Coasts. The ANN models are trained with storm surge levels from a hydrodynamic model and physical characteristics of synthetic hurricanes which are downscaled from National Centers for Environmental Prediction (NCEP) reanalysis using a statistical-deterministic hurricane model. The ANN models are able to accurately predict storm surge levels with root-mean-square errors (RMSE) below 0.2 m and correlation coefficients > 0.85. The ANN models trained with the NCEP data also show satisfactory accuracy (RMSE below 0.7 m; correlation > 0.7) in predicting storm surge levels for hurricanes downscaled from future climate projections. Once trained, we use the ANN models to assess the sensitivity of storm surge levels to variations in hurricane characteristics and local geophysical features. Progressively stronger maximum wind speeds and larger outer radius sizes independently increase storm surge levels at all locations along the US East and Gulf Coasts. The response of storm surge levels to changes in hurricane translation speed, however, is found to be sensitive to coastal configuration, with increases in hurricane translation speed amplifying (reducing) storm surge levels in open ocean (semi-enclosed) regions.

Plain Language Summary Hurricane storm surge represents a significant threat to coastal communities around the world. Here, we use a type of machine learning model known as artificial neural network (ANN) models to predict storm surge levels using hurricane characteristics along the US Gulf and East Coasts. The machine learning models are trained with storm surge levels from a hydrodynamic model and physical characteristics of synthetic hurricanes which are downscaled from historical data. The machine learning models are able to accurately predict storm surge levels. The ANN models trained with the NCEP data also show satisfactory accuracy in predicting storm surge levels for hurricanes downscaled from future climate projections. Once trained, we use the ANN models to explore the sensitivity of storm surge levels to variations in hurricane characteristics and local geophysical features. Progressively stronger maximum wind speeds and larger outer radius sizes independently increase storm surge levels at all locations along the US East and Gulf Coasts. The response of storm surge levels to changes in hurricane translation speed, however, is found to be sensitive to coastal configuration, with increases in hurricane translation speed amplifying (reducing) storm surge levels in open ocean (semi-enclosed) regions.

1. Introduction

Hurricane-induced coastal flooding in the U.S. has historically caused extensive damage and loss of life, with much of the destruction resulting from hurricane storm surges. A storm surge is the elevation or depression of the sea surface with respect to the predicted tide caused by changes in atmospheric pressure or by strong winds pushing water toward the shore (Gregory et al., 2019). In shallow coastal regions, storm surge levels are a function of both the physical characteristics of the hurricane (e.g., wind speed, size) and the local geophysical features (e.g., coastline geometry and bathymetry).

Studies have simulated storm surge events primarily using either hydrodynamic models or data-driven approaches. The advantage of hydrodynamic models is that they can be used to understand the relative contributions of different processes in extreme events (Marsooli & Lin, 2018; Muis et al., 2019), for storm surge predictions at locations where there are no tide gauges and for simulating unobserved events that are expected to occur rarely.

Journal of Geophysical Research: Atmospheres 10.1029/2022JD037617

Project Administration: Ning Lin, Michael Oppenheimer, Ching-Yao Lai
Resources: Joseph W. Lockwood, Ning Lin
Software: Joseph W. Lockwood, Ching-Yao Lai
Supervision: Ning Lin, Michael Oppenheimer, Ching-Yao Lai
Validation: Joseph W. Lockwood
Visualization: Joseph W. Lockwood, Michael Oppenheimer, Ching-Yao Lai
Writing – original draft: Joseph W. Lockwood, Ning Lin, Michael Oppenheimer, Ching-Yao Lai
Writing – review & editing: Joseph W. Lockwood, Ning Lin, Michael Oppenheimer, Ching-Yao Lai

Hydrodynamic models, however, require large computational resources and are currently limited in their ability to simulate large ensembles of events globally and in near-term forecasting (Fox-Kemper et al., 2021). To overcome these limitations associated with hydrodynamic modeling, studies have used data-driven approaches, that quantify the relationship between the predictand (e.g., storm surge) and relevant predictors (e.g., wind speed, sea-level pressure), to model coastal water levels (Brunner et al., 2020; Tadesse et al., 2020; Tiggeloven et al., 2021). Data-driven approaches are able to predict storm surge levels without the need for knowledge of the underlying physical processes, reducing computational cost.

Increases in computational power and in the sophistication of machine learning techniques have lead to recent advances in data-driven modeling of storm surge (Bruneau et al., 2020; Chen et al., 2022; Guillou & Chapalain, 2021; Ramos Valle et al., 2021; Reichstein et al., 2019; Tadesse et al., 2020; Tiggeloven et al., 2021). Recently, Bruneau et al. (2020) used neural network (NN) models to predict hourly storm surge levels at global tide gauge stations using a variety of atmospheric and oceanic reanalysis products as inputs to the NN models. Tiggeloven et al. (2021) expanded upon the NN models by using more complex deep learning models (such as convolutional NN) to construct ensembles of water levels at global tide gauge stations. Lee et al. (2021) used a convolutional NN model combined with principal component analysis and a k-means clustering for the Chesapeake Bay area to predict peak storm surges from synthetic tropical cyclone data.

Results from these studies suggest that machine learning modeling approaches can model storm surge levels in a computationally efficient way and with high accuracy compared to physics based models. However, large errors in machine learning storm surge predictions are often linked to the passage of hurricanes (particularly high intensity storms) in low latitude regions (Bruneau et al., 2020; Tadesse et al., 2020; Tiggeloven et al., 2021), likely resulting from the fact that hurricanes are poorly resolved by most course resolution reanalysis products (Bloemendaal et al., 2019) and due to the limited set of hurricane events used to train the data-driven models. Additionally, climate change is expected to alter hurricane characteristics into the future, including their translation speeds and maximum wind speeds (Knutson et al., 2020). This may prove to be a challenge for predictions made by data-driven models, as models that are trained on one population (e.g., historical reanalysis) generally have poor accuracy on different temporal and spatial populations (e.g., future climates) (Chattopadhyay et al., 2020).

In addition to storm surge prediction, many studies have used hydrodynamic models to assess the response of storm surge levels to variations in hurricane characteristics and geophysical features (Ayyad et al., 2021; Irish et al., 2008; Lin et al., 2010; Liu et al., 2019; Ramos Valle et al., 2021; Rego & Li, 2009). Studies that assess the response of storm surge to hurricane characteristics often overlook the correlation between storm characteristics or the impact of geophysical setting (e.g., Ayyad et al., 2021; Nielsen, 2009; Pandey and Rao, 2019; Peng et al., 2004; Rego and Li, 2009). Such analysis is usually performed by varying one of the hurricane characteristics, for example, wind speed, by a certain percentage while assuming all other characteristics, for example, size and translation speed, constant, for a specific location. On the other hand, studies that assess the sensitivity of storm surge to geophysical setting often model a very limited set of hurricane scenarios (e.g., Irish et al., 2008; Peng et al., 2004; Zhang and Li, 2019).

Notably, a number of studies have suggested that storm surge levels increase as a hurricane moves progressively faster (Ayyad et al., 2021; Irish et al., 2008; Liu et al., 2019; Rego & Li, 2009), while other studies have found that storm surge levels decrease with increases in the hurricane speed (Liu et al., 2019; Peng et al., 2004; Weisberg & Zheng, 2006). Recently, using 42 years of tidal records and landfall tropical cyclone best tracks in Japan, Islam and Takagi (2021) found strong positive correlations between hurricane surge and speed along open coastlines, but negative correlations in semi-enclosed bays. Using Hurricane Mathew (2016) and ADvanced CIRCulation model (ADCIRC), Thomas et al. (2019) found that increasing the storm speed would have caused higher surges on the open coast. They found that in estuaries and bays reductions in storm speed would enhance surge levels as more water being pushed into these regions. It is therefore plausible that the different responses of storm surges to hurricane translation speed may be related to coastal configuration, however given that these studies are limited in scope and events, considerable uncertainty remains.

After model training, machine learning approaches can also be used to quantify the relationship between input and output variables. Ramos Valle et al. (2021) were, to our knowledge, the first study to use deep learning techniques to assess the sensitivity of storm surge to hurricane characteristics. They implemented an artificial neural network (ANN) model to predict storm surge levels for 200 category 1–2 synthetic hurricanes for five sites located in the Mid-Atlantic Bight region. Their ANN model was trained with synthetic hurricane characteristics modeled using the Hybrid Weather Research and Forecasting model and with storm surge levels modeled using the ADCIRC. Using the ANN model, they then identified hurricane intensity and position as the most important variables for accurate storm surge predictions.

In this paper, we use neural network models to (a) predict hurricane storm surge levels using hurricane characteristics and (b) assess the sensitivity of storm surge levels to changes in hurricane characteristics and geophysical features. We expand upon the framework of Ramos Valle et al. (2021) by training the ANN models with a larger (5018) set of category 1–5 synthetic hurricanes along the entire US East and Gulf Coast. These hurricanes are downscaled from years 1980–2005 of NCEP reanalysis using a statistical-deterministic hurricane model. To assess the ability of the ANN models in predicting hurricane storm surges under climate change, we test the ANN models trained with the synthetic NCEP data on hurricanes downscaled from Coupled Model Intercomparison Project Phase 6 (CMIP6) models under a high emissions scenario. Finally, we use the ANN models to quantify the sensitivity of storm surge levels to geophysical location and to changes in the characteristics of the synthetic storms including their size, maximum wind speed and translation speed.

2. Methodology

2.1. Climatological-Hydrodynamic Modeling

We use 5018 synthetic hurricanes generated in Gori et al. (2022) from National Centers for Environmental Prediction (NCEP) reanalysis for the historical time period between 1980 and 2005. The hurricanes are downscaled using the statistical/deterministic hurricane model of Emanuel et al. (2008). For each hurricane, the outer radius at which the cyclonic wind speed goes to zero (henceforth outer radius; Ro) is randomly drawn from an empirical lognormal distribution (Chavas et al., 2015). Wind fields are developed based on the maximum wind speed (V_{max}) and Ro of each synthetic hurricane using the physics-based parametric model of Chavas et al. (2015). We use the model of Holland (1980) to develop pressure fields, which is related to the maximum wind speed, R_{max}, and pressure deficit.

To assess the ability of the ANN models in predicting hurricane storm surge under climate change, we test the ANN models trained with the historical NCEP data in predicting storm surges for 5000 synthetic hurricanes downscaled from years 2070–2100 of SSP5-8.5. Specifically, we model 1,000 surge events from hurricanes downscaled from simulations of each of five CMIP6 climate models: Canadian Earth System Model, Centre National de Recherches Météorologiques (CNRM), EC-Earth Consortium Model (ECEARTH) and Max Planck Institute Earth System Model (MPI) and The Institute Pierre Simon Laplace Climate Model (IPSL).

We simulate time series of storm surge levels for each synthetic hurricane using the 2D depth-integrated version of the ADCIRC model (Luettich et al., 1992). We utilize an unstructured computational mesh developed by Marsooli and Lin (2018) that spans the entire North Atlantic basin. Tides and waves are neglected to keep the focus on storm surge. Other model parameters are set following Marsooli and Lin (2018). The mesh resolution is 1 km in the nearshore on the U.S. East and Gulf Coasts, and gradually increases to 5 and 15 km at water depths of 20 and 50 m, respectively. In deeper waters, the mesh resolution varies between 15 and 50 km in the Gulf of Mexico and between 15 and 100 km in the Atlantic Ocean. Other model parameters are set following Marsooli and Lin (2018). For each hurricane event, we extract timeseries of storm surge levels for 300 sites that are spaced approximately 25 km apart (each site will be used as the location of one ANN model) along the US East and Gulf coasts.

2.2. Training Data and Neural Network Architecture

We train a ANN model at each of the 300 sites along the US East and Gulf Coasts spaced approximately 25 km apart (Figure 1). We train each ANN model with hourly data of synthetic hurricanes within 350 km of each site. We find that the accuracy of the ANN models exhibit little sensitivity to the radius from each site (Table S1 in Supporting Information S1). Following Ramos Valle et al. (2021), the input layer of each NN model consisted of six input parameters related to the physical characteristics of the synthetic hurricanes, including the maximum wind speed (V_{max}), Ro, bearing, translation speed, and track position as given by latitude and longitude. The output data (or target variable) is hourly storm surge levels (Figure 1a).

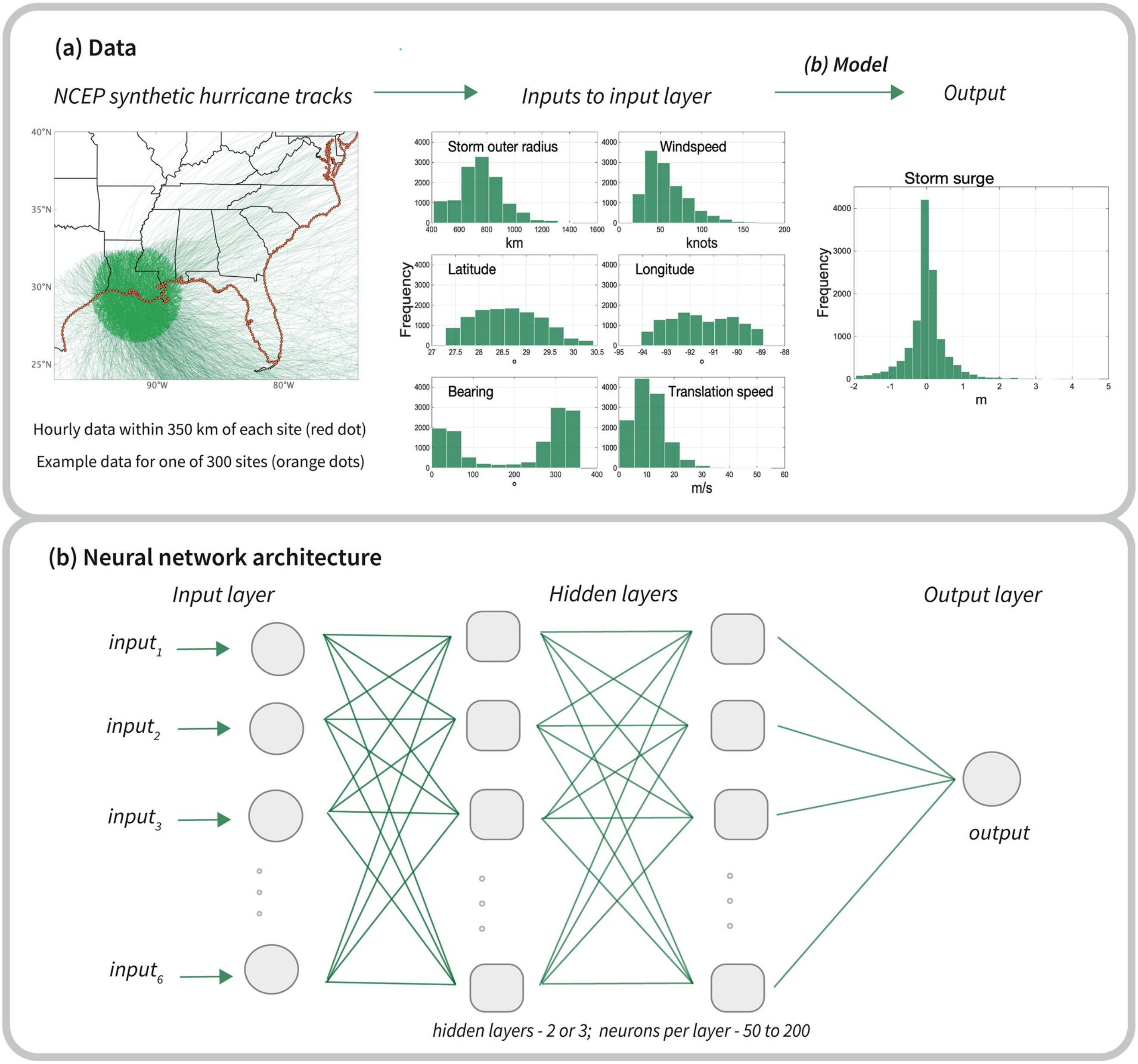

Figure 1. (a) Model data and (b) neural network architecture. See Section 2.2 for the details of the neural network hyperparameters tested.

We use V_{max} as the metric for hurricane intensity instead of central pressure, as generally a hurricane's central pressure directly contributes less than 15 per cent to the magnitude of the storm surge (Horsburgh et al., 2011). Potential issues associated with the multicollinearity of input variables can make interpretation of ANN models challenging (Apley, 2020). To address this issue, we use each hurricane's outer radius to characterize the size of the hurricane, as Ro is sampled from a lognormal distribution and is therefore uncorrelated with the other hurricane characteristics. We use Ro as the metric for hurricane size instead of the radius of maximum wind speed (R_{max}) as R_{max} is positively correlated with V_{max} (Chavas et al., 2016).

ANN models are a type of supervised modeling technique in which both input and outputs are provided, allowing the model to learn and determine the relationships between them. Following Ramos Valle et al. (2021), we use a ANN model framework (in contrast to more sophisticated models such as Long Short-Term Memory and Convolution NN models) to allow for both computational efficiency in training the models with a large number of events and interpretability of the ANN models through the use of accumulated local effects (ALE) plots

(Section 2.3; see Text S1 in Supporting Information S1 for a comparison between this model and other commonly used data-driven modeling techniques for storm surge modeling). ANN models contain many interconnected units (neurons) that can extract linear and non-linear relationships in the data. The ANN architecture comprises multiple layers: an input layer, a variable number of intermediate layers, known as hidden layers, and an output layer (Figure 1b). The relationship between the input and output of a neuron (denoted by i) in a hidden layer in the ANN can be written as (Ramos Valle et al., 2021):

$$\mathrm{AF} = b_i + \sum_{j=1}^{n} x_j^{in} w_j \tag{1}$$

$$x_i^{out} = f(\mathrm{AF}) \tag{2}$$

where x_j^{in} and x_i^{out} are input and output data of the neuron i and AF represents the input to the activation function, $f(AF)$. n is the total number of neuron from the previous hidden layer connected to neuron i in the current hidden layer. The weights and the bias are given by w_j and b_i, respectively.

Each neuron receives an input, processes it through an activation function, and produces an output. The activation function must be non-linear for the ANN model to learn beyond linear relationships. While a sigmoid activation function is used for the last layer, the hidden layers consists of Rectified Linear Units activation functions. The output from the sigmoid activation function is re-scaled back to the storm surge units. Finally, an Adam solver is used to minimize the root mean-square error (MSE) between ANN predictions and ADCIRC surge training data. We used the Scikit-Learn GridSearchCV package to determine the best set of ANN hyperparameters at each site that minimizes the MSE evaluated on the validation data set, including the number of hidden layers and the number of neurons within each of these and the type of learning rate. The hyperparameters tested are shown in Table S2 in Supporting Information S1 and chosen number of hidden layers and the number of neurons are shown in Figure S1 in Supporting Information S1. As determined by the Scikit-Learn GridSearchCV package, most of the models have 200 neurons per layer, and all models have three hidden layers and constant learning rates (Figure S1 in Supporting Information S1). In the present study, the number of epochs, or maximum training iterations, was set to 1,000, and we hold the L2 regularization constant at 0.0001.

Each ANN model is evaluated using statistical metrics including root-mean-square error (RMSE), MSE, correlation coefficient and R-squared (r^2; Text S2 in Supporting Information S1). Since the ANN models provide an ensemble of surge predictions, following Bruneau et al. (2020) and Tiggeloven et al. (2021) we select the Continuous Ranked Probability Score (CRPS), which generalizes the mean absolute error to the case of probabilistic forecasts, as a metric to evaluate the probabilistic predictions (Text S2 in Supporting Information S1). The input and output data for each ANN model were randomly divided into 70% for training, 10% for validation and 20% for testing. Following Bruneau et al. (2020) and Tiggeloven et al. (2021), we construct an ensemble of 20 models for each location. For each individual model of the ensemble, we randomly sample 50% of the training, validation and testing data. The training data is used during the learning process and is used to fit the parameters (e.g., weights), the validation data set is used to tune the hyperparameters and the test data is used only to assess the performance. While a larger ensemble would have improved our probabilistic forecast, 20 members allows for both variability in the predictions and computational efficiency.

2.3. Metrics to Assess the Sensitivity of Storm Surge to Storm Characteristics

ALE plots, which are one of the most popular approaches for visualizing the effects of the predictors with black box supervised learning models, are used here to assess the impact of each input characteristic on hurricane storm surge prediction (Apley, 2020). ALEs describe how inputs influence the prediction of a machine learning model on average by isolating the change in prediction caused by a change in a single feature. Here, we calculate ALE interactions by isolating the change in storm surge prediction caused by variations in hurricane maximum wind speed, outer size and translation speed. ALE functional forms that deviate markedly from a horizontal line denote larger influence on the storm surge prediction.

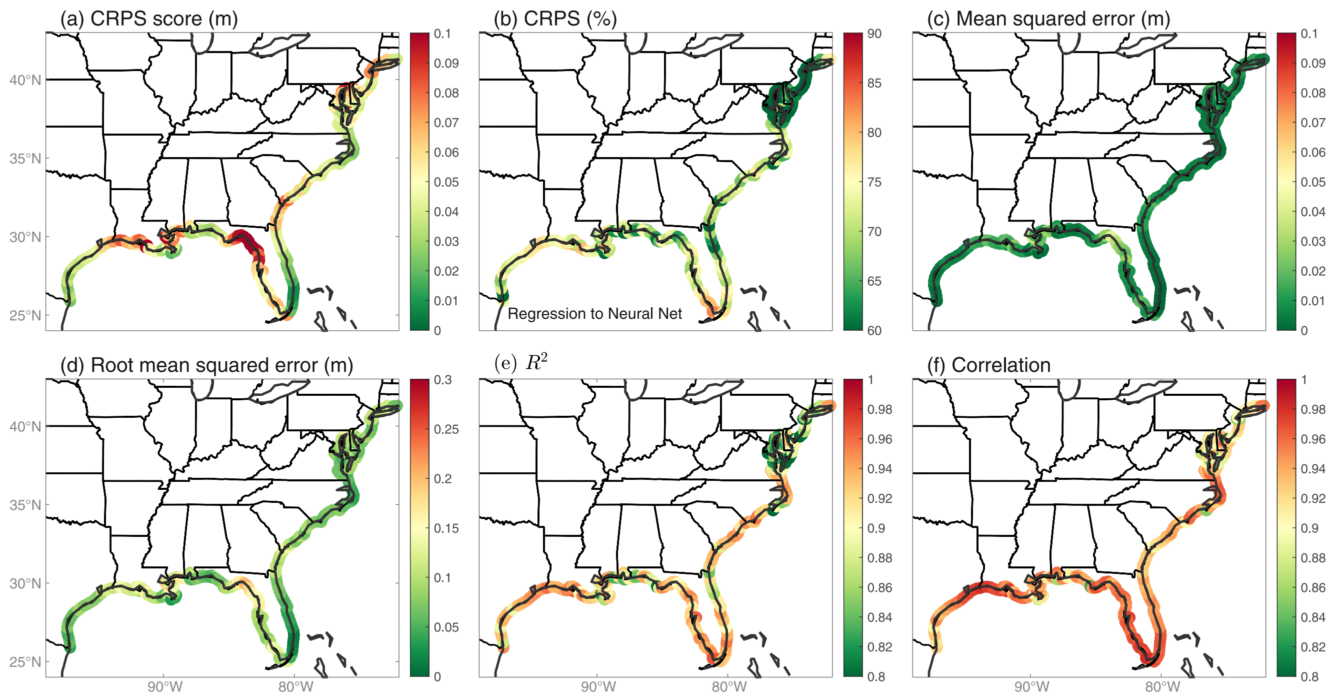

Figure 2. Skill metrics for the ANN: (a) Continuous Ranked Probability Score (CRPS) for each ANN and (b) percentage improvement in CRPS from the multivariate linear regressions to neural network ensemble. (c) Mean squared error, (d) root mean squared error, (e) r-squared (R^2) values, (f) correlations calculated between the neural network prediction and test data at each site.

3. Results

3.1. Performance and Skill of the ANN Models

The skill metrics calculated using the test data for each ANN model are shown in Figure 2. At all sites, CRPS values range from 0.02 to 0.2 m, MSE are values below 0.1 m, and r^2 values and correlation coefficients exceed 0.85, indicating that the ANN models are able to accurately predict storm surge levels. The ANN models are able to capture the full spectrum of storm surge variability, with both positive and negative storm surge levels well predicted by the ANN models (Figure 3). The ANN models along the western coast of Florida show the lowest accuracy in predicting storm surge events compared to other sites, likely resulting from the multiple different directions that the TC may approach the site. For example, a TC may approach from the Gulf of Mexico and produce a large positive surge to its right on the western coast of Florida or from the Atlantic through the east coast of Florida and produce a negative surge at the same site to it's right.

Our ANN models are of comparable performance to Ramos Valle et al. (2021) with MSE values between 0 and 0.1 m and r^2 values exceeding >0.85. In contrast to Ramos Valle et al. (2021), Tiggeloven et al. (2021), Bruneau et al. (2020), and Lee et al. (2021), our ANN models do not under-estimate extreme storm surge levels (Figure 3). This may be due to the fact that we use a much broader set of physically plausible hurricane (including very high intensity storms) and storm surge events to train the ANN models and that we allow the ANN model structure (the number of neurons and layers; see Section 2.2) to vary over each of the 300 sites.

In this study we implement ANN models that process only single data points (in contrast to models such as in Long Short-Term Memory models that capture long-term temporal dependencies), and as a result are not able to learn sequences of data or temporal dynamic behavior. However, we reconstruct the time-series storm surge levels for each ANN models, finding that our ANN models are able to accurately capture the temporal evolution and variability of extreme surge levels along the US East Coast (Figure 4).

While the ANN models generally show high skill in predicting storm surges, it is worth considering how a multivariate linear regression, that does not have nonlinear activation functions, would perform in comparison. Percentage improvement in CRPS from multiple linear regression model to ANN are presented in Figure 2, with consistent improvements across sites of between 60% and 90%. Additionally, the ANN models show significant

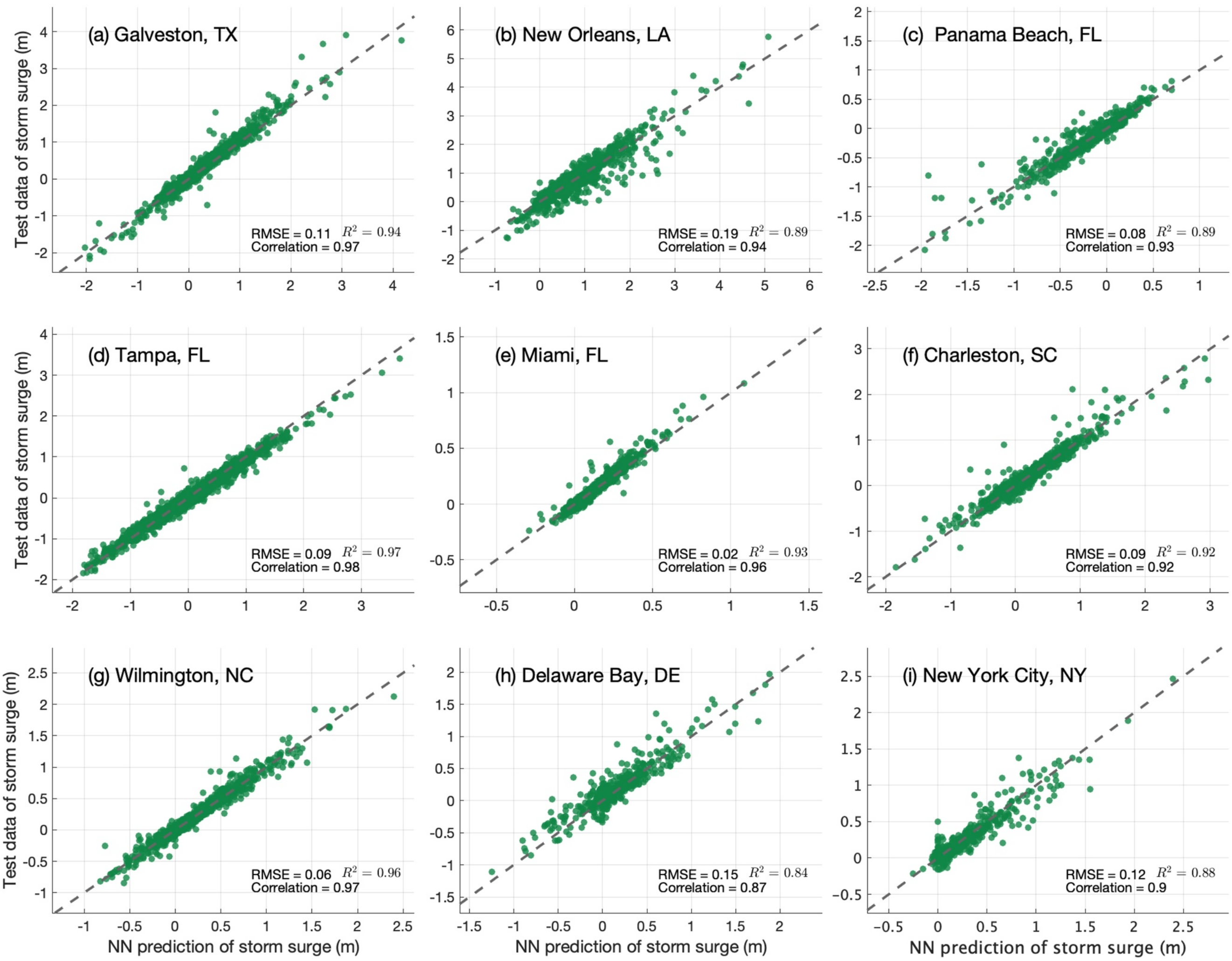

Figure 3. (a–i) Scatter of storm surge values and neural network predictions at select sites for the test data. The data is shown is for the first ensemble only. Figure S2 in Supporting Information S1 shows the same figure but for training data.

improvement over multiple linear regression models in timeseries prediction (Figure 4). The large improvement in prediction from the multiple linear regression to ANN models at each site suggests that the storm surge levels are function of nonlinear combinations of hurricanes characteristics, which can only be captured using nonlinear activation functions.

Climate change is expected to alter hurricane characteristics into the future, including their translation speeds and maximum wind speeds (Knutson et al., 2020). This may prove to be a challenge for predictions made by machine learning models, as models that are trained on one population (e.g., historical data) generally have poor accuracy on different temporal and spatial populations (e.g., future climates) (Chattopadhyay et al., 2020). To assess the ability of the ANN models (trained with historical NCEP data) in predicting hurricane storm surges in a future environment, we also test the ANN models with hurricanes downscaled from SSP5-8.5. We find that the ANN models are generally able to predict hurricane storm surge variability with satisfactory accuracy with RMSE below 0.7 m and correlation coefficients exceed 0.7 (Figure 5 and Figure S3 in Supporting Information S1). The ANN models, however, predict storm surge variability with lower accuracy compared to the historical NCEP synthetic hurricanes (Figure 3).

Our analysis of TC translation speed and intensity (maximum wind speed) reveals an increase in the number of slow-moving and stronger TCs along the US East Coast (Figure 6). Hurricane storm surge is expected to increase approximately with the square of tropical cyclone wind speed (Woodruff et al., 2013). Given this relationship, we

10.1029/2022JD037617

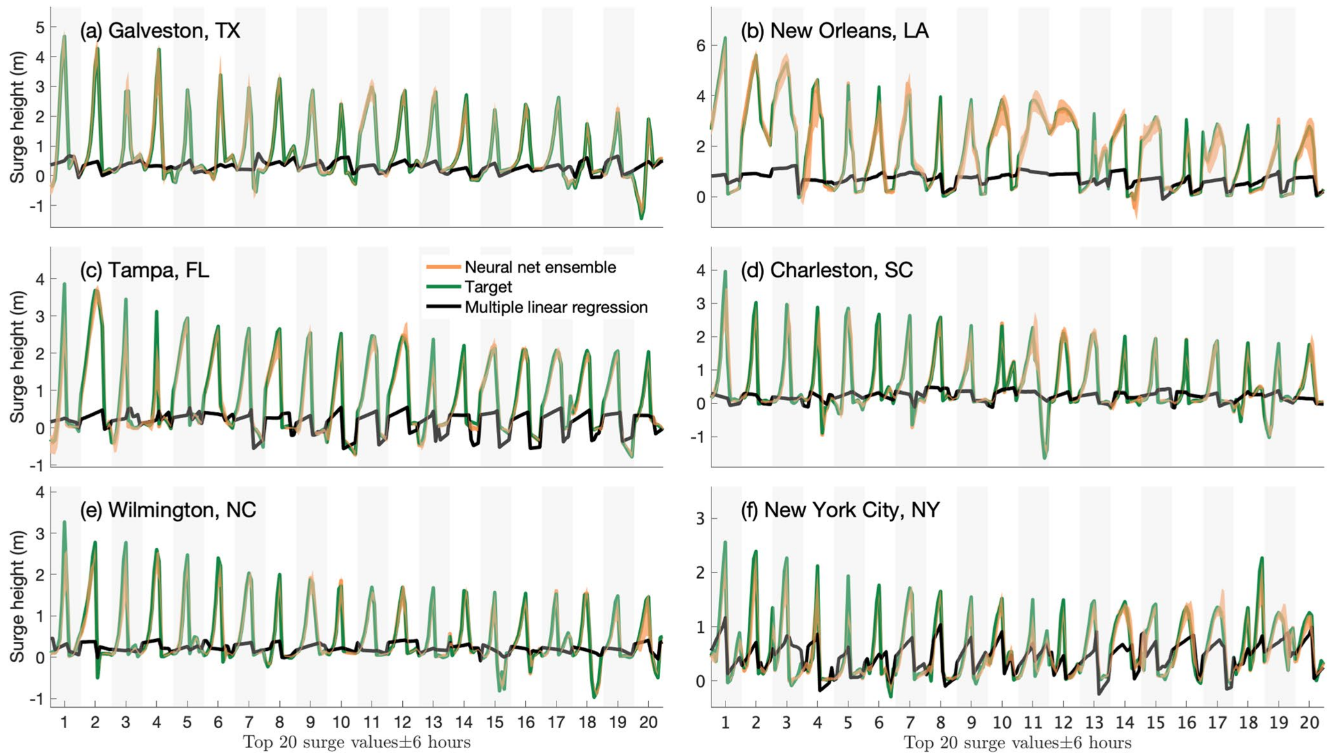

Figure 4. (a–f) Six hours before and after the 20 events with the largest surge heights at each sites. Green lines shows the target timeseries and orange colors denote the ANN models ensemble average (line) and range (shading) of predictions. The black lines show the multiple linear regression ensemble calculated using the same data input to the ANN models.

may expect the ANN models to be unable to model the storm surge values associated with stronger storms as they will fall outside of the NCEP data used to train the models. However, as there is a reduction in accuracy across all storm surge levels (Figure 5), it is likely that the reduced accuracy of the ANN models in predicting storm surge values under SSP5-8.5 result also from changes in the combination of hurricane characteristics (e.g., slower and stronger storms) that fall outside of the NCEP data used to train the ANN models.

3.2. Sensitivity of Storm Surge to Hurricane Characteristics and Geophysical Features

We next calculate the accumulated local affect (ALE) plots for each ANN model using the test data, to assess how the storm surge prediction changes with a single input variable. We isolate the change in storm surge prediction caused by changes in hurricane maximum wind speed (V_{max}), outer size (Ro) and translation speed (Figures 7 and 8). ALE functional forms that deviate markedly from a horizontal line denote larger influence on the prediction, with positive slopes indicating that the increase of the values of that input variable will increase the predicted storm surge levels. At sites along the US East and Gulf Coasts, the ALE plots show that increases in the maximum wind speed and outer radius independently lead to increases in storm surge levels (Figure 7).

In contrast to maximum wind speed and outer storm radius, the ALE plots for hurricane translation speeds are found to strongly differ across sites (Figure 8). A recent empirical study found that increases (decreases) in hurricane speed amplify storm surge levels along open (semi-enclosed) coastlines (Islam & Takagi, 2021). Following Islam and Takagi (2021), we average the ALE plots for translation speed across semi-enclosed and open ocean regions (Figure 8). Here, we use a morphometric definition of bays and estuaries (semi-enclosed regions) following Healy and Harada (1991) as regions where the ratio of length of the major axis within the enclosed sea to entrance width exceeds 4, with at least one entrance to the open ocean (schematic illustration shown in Figure S4 in Supporting Information S1).

We find that on average, slower moving hurricanes produce higher storm surges in semi-enclosed sites (Figure 8), independent of changes in other hurricane characteristics. In contrast, in open ocean regions, the storm surge levels

10.1029/2022JD037617

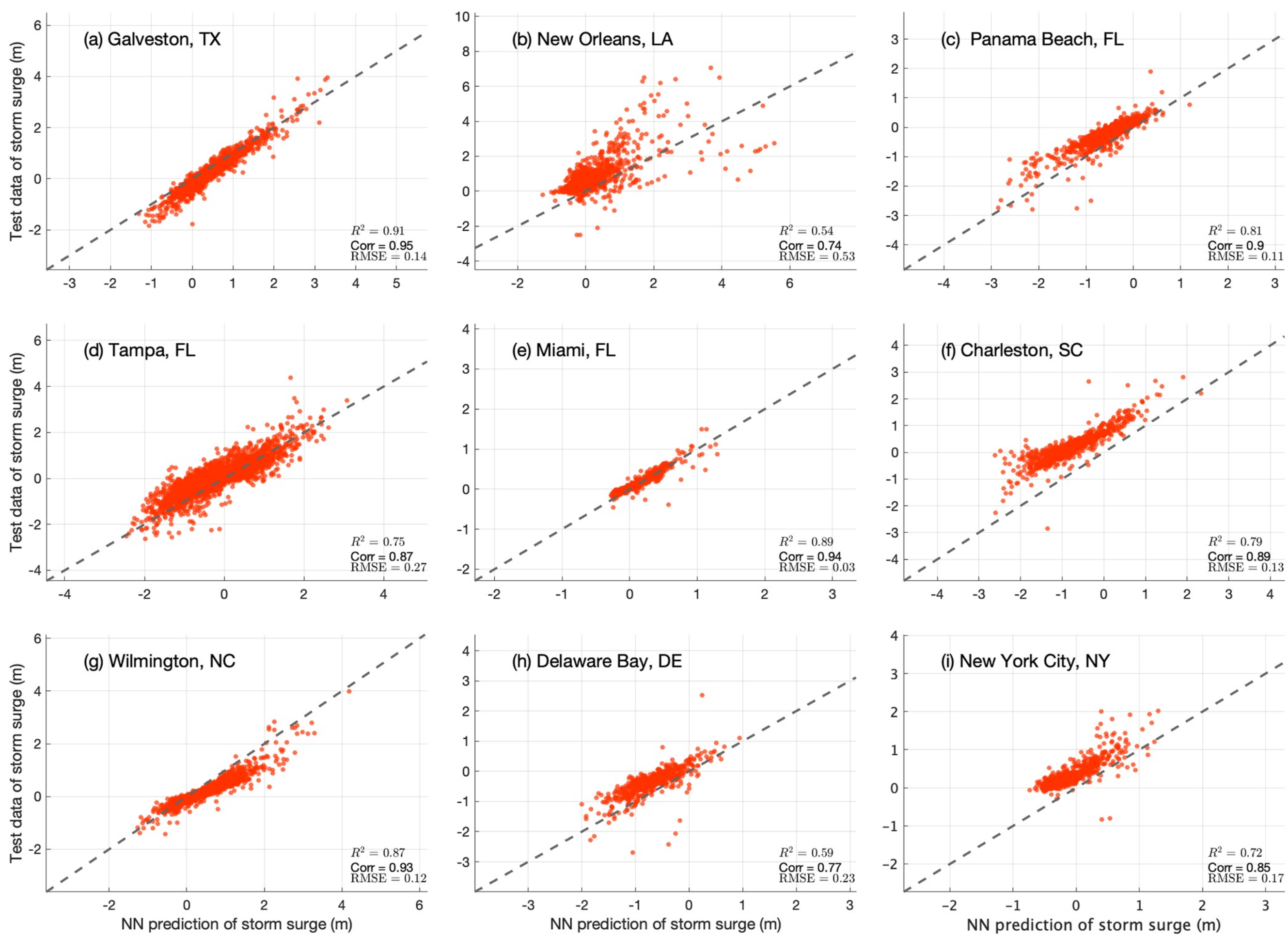

Figure 5. (a–i) Scatter of storm surge target values and neural network predictions at select sites for the end of century SSP5-8.5.

increase as the hurricane translation speed increases. This distinction between open ocean and semi-enclosed coastlines may be related to the local hydrodynamic response to storm characteristics. In enclosed sites, slower moving hurricanes will have more time to push water into the bays leading to an increase in local storm surge levels. In open coastlines, water would be able to flow unconstrained away due to limited laterally confinement by coastlines (Zhang & Li, 2019). Also, it is plausible that in open coastlines fast-moving hurricanes would energize a coastal shelf wave and cause an increase in storm surge levels because the hurricane's translation speed will coincide with the long-wave shelf propagation speed ($\sqrt{gh}$; h is water depth) as predicted by Proudman's linear theory (Islam & Takagi, 2021; Proudman, 1953).

4. Discussion

The ANN models trained in this study show significant skill in predicting peak storm surge levels and surge time series. In contrast to prior data-driven studies that use course resolution reanalysis (Bruneau et al., 2020; Tiggeloven et al., 2021) or that train ANN models with a limited number of hurricane scenarios (Lee et al., 2021; Ramos Valle et al., 2021), our ANN models do not under-predict extreme storm surge levels. This may be due to the fact that we use a much broader set of physically plausible hurricane and storm surge events to train the ANN models and that we allow the structure of the ANN models (e.g., the number of hidden layers) to vary between for each site.

Analyses of historical hurricane events and physics-based hurricane surge modeling studies have revealed that wind intensity is not the only storm parameter that can markedly influence storm surge levels (Irish et al., 2008; Peng et al., 2004). In particular, a number of studies found that storm surge levels can increase or decrease with

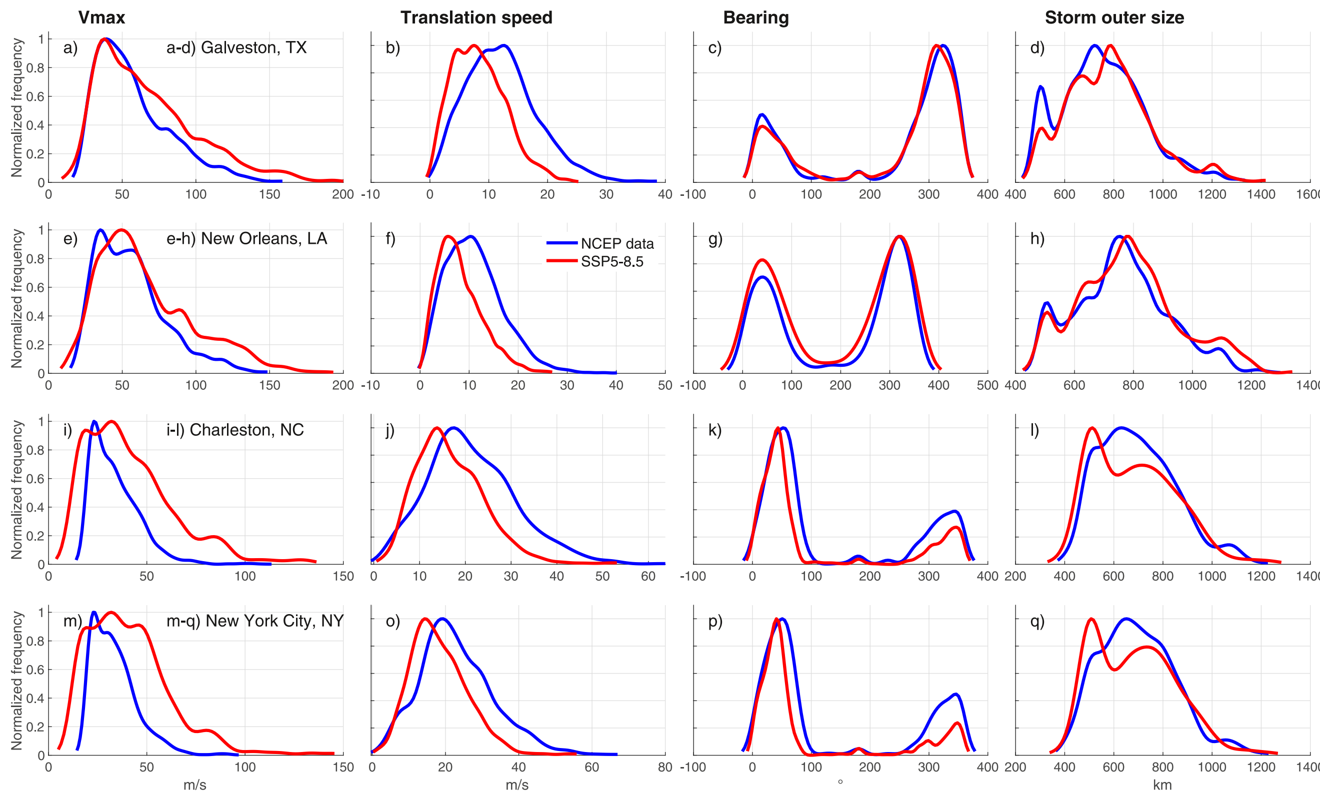

Figure 6. (a–i) Changes in hurricane characteristics between the NCEP (blue) and future (SSP5-8.5) time periods for select sites along the US East and Gulf Coasts. The summary statistics for each site can be found in Figure S3 in Supporting Information S1.

storm translation speed (see Table S3 in Supporting Information S1 for the details and results of each study). Using a limited number of locations and storm events (detailed in the introduction), Islam and Takagi (2021), Thomas et al. (2019) suggested that the different responses of storm surges to hurricane translation speed may be related to coastal configuration.

We use the recently developed accumulated local affects (ALE) plots to explore the sensitivity of storm surge levels to hurricane characteristics and local geophysical features. We find that coastline geometry can impact local storm surge levels, with semi-enclosed regions producing the largest surge levels during slow moving hurricanes. In open ocean shorelines, on the other hand, storm surge levels are lower during slower moving hurricanes

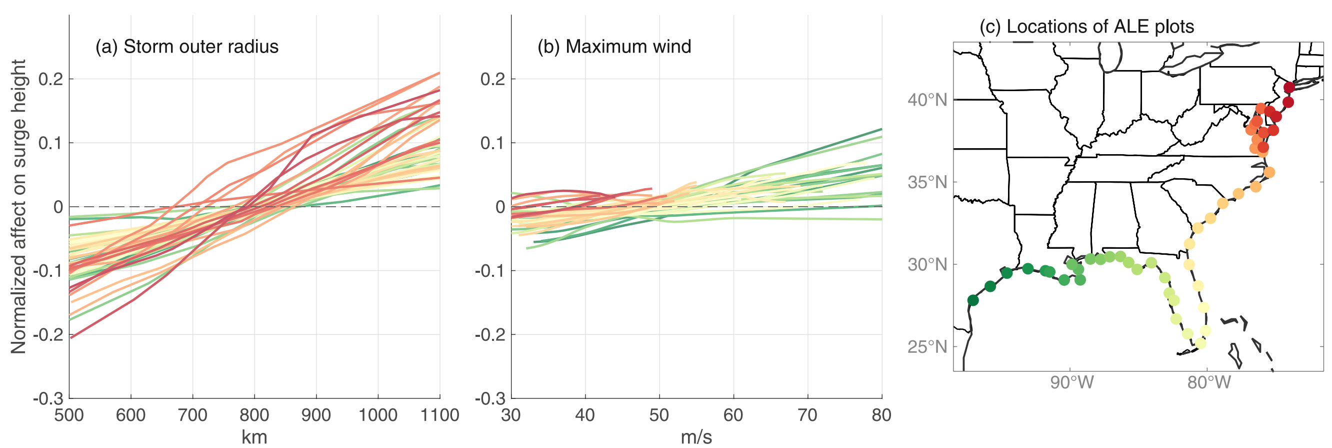

Figure 7. Accumulated local effect plots of (a) radius of outer storm size and (b) maximum wind speed for select locations (c). Color of curves indicates select locations shown on (c). Positive y values show that the value positively impacts surge prediction.

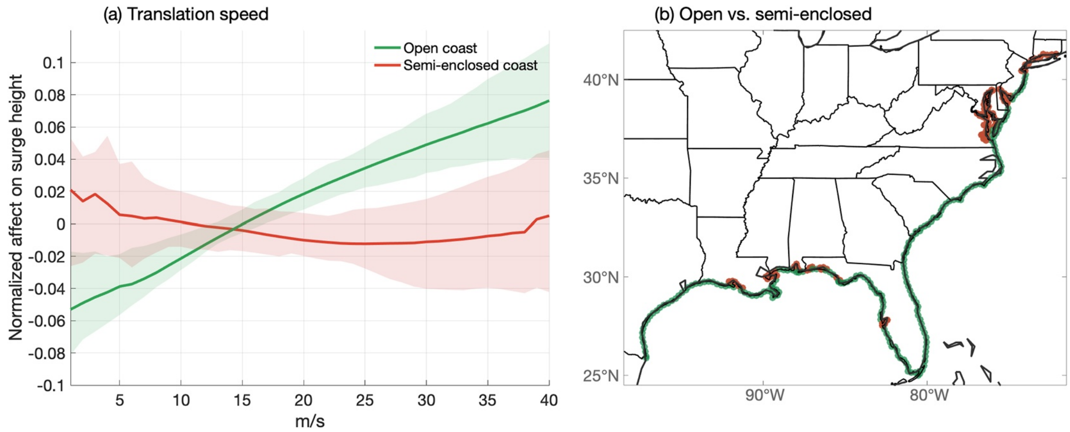

Figure 8. (a) Accumulated local effects for hurricane translation speed, averaged over open (green) and semi-enclosed regions (red) (b). Shadings denote the 20%–80% range across the semi-enclosed and open ocean regions. Semi-enclosed regions are defined following Healy and Harada (1991) as regions where the ratio of length of the major axis within the enclosed sea to entrance width exceeds 4, with at least one entrance to the open ocean.

when compared to faster storms. Several recent studies (Knutson et al., 2020; Kossin, 2018; Sun et al., 2021) have suggested that the translation speed of hurricanes has and will continue to decrease significantly both at the global and regional scales. Therefore, semi-enclosed sites (e.g., New York City and New Orleans) may be at greater hazard to any further slow down in hurricane speed with climate change compared to open ocean regions, independent of changes in other hurricane characteristics.

Empirical analyses of historical hurricanes by Needham and Keim (2014) revealed a positive correlation between storm surge levels and hurricane radius of 34, 50, and 64 knot winds. Using ADCIRC, Irish et al. (2008) showed that size as measured by the radius of maximum wind speed plays an important role in surge generation particularly on mildly sloping regions; with increases in size enhancing storm surge levels. In agreement with these prior studies, we find that peak surge increases as storm size (Ro) increases.

While the ANN model structures are not directly derived from physical mechanisms governing storm surges, future work can make use of physics-informed machine learning to improve the predictive ability (Karniadakis et al., 2021). We note that this paper is not an attempt to comprehensively find the deep learning model type with the highest accuracy at each site. Although our model is able to predict timeseries and storm surge variability with good accuracy, the investigation of other deep learning models (e.g., LTSM, RNN, CNN) that can improve the predictions is subject to future work.

5. Conclusion

Data-driven neural networks are increasingly used to extract patterns and insights from the ever-increasing stream of geospatial data, and as fast and accurate way of predicting geophysical hazards (Reichstein et al., 2019). In this study, we use models and a large set of physically plausible hurricanes to predict hurricane storm surge levels along the US East and Gulf Coasts. These ANN models show significant skill in predicting variability of storm surges levels and storm surge time series. Our ANN models are also able to predict the variability of storm surge levels for hurricanes downscaled from SSP5-8.5. When using ANN models to predict storm surge in future environments, however, the training set may be expanded to include storms covering future parameters space to improve accuracy.

We also use the ALE plots to assess the sensitivity of storm surge levels to variations in hurricane characteristics and local geophysical features. Progressively stronger maximum wind speeds and larger outer radius sizes independently increase storm surge levels along the full US East and Gulf Coast. The response of storm surge levels to changes in hurricane translation speed, however, is found to be sensitive to coastal configuration, with increases in hurricane translation speed amplifying (reducing) storm surge levels in open ocean (semi-enclosed) regions.

Although we focus here on storm surge levels, we note that the ANN models employed in this study could be expanded to incorporate tides (Bruneau et al., 2020), waves (Ramos Valle et al., 2021) and pluvial and fluvial processes, to produce comprehensive and fast probabilistic coastal flood hazard predictions. The model framework developed in this study can accommodate fast predictions for the time-series surge evolution and thus it may also be used to predict flood area and extent, and compound flooding (Gori et al., 2022). Our ANN models may also be able to predict storm surge levels for coastal communities in future environments and support stakeholders and communities through efficient risk assessment and emergency response management operations, in near term forecasting and adaptation planning.

Data Availability Statement

Code used in this analysis can be obtained at https://doi.org/10.5281/zenodo.7261531. The original synthetic tropical cyclone datasets used in this study are freely available from Kerry Emanuel for research purposes. For the details and availability of the synthetic datasets, please refer to Emanuel (2021) (https://doi.org/10.1175/JCLI-D-20-0367.1) [Dataset].

Acknowledgments
JWL was supported by the Princeton University Fellowship in Natural Sciences and Engineering. The study was also supported by the U.S. National Science Foundation Grant 2103754 as part of the Megalopolitan Coastal Transformation Hub.

References

Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B*, *82*(4), 1059–1086. https://doi.org/10.1111/rssb.12377

Ayyad, M., Hajj, M. R., & Marsooli, R. (2021). Spatial variation in sensitivity of hurricane surge characteristics to hurricane parameters. *Journal of Engineering Mechanics*, *147*(10), 04021070. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001984

Bloemendaal, N., Muis, S., Haarsma, R., Verlaan, M., Irazoqui, M., Moel, H., et al. (2019). Global modeling of tropical cyclone storm surges using high-resolution forecasts. *Climate Dynamics*, *52*(7–8), 5031–5044. https://doi.org/10.1007/s00382-018-4430-x

Bruneau, N., Polton, J., Williams, J., & Holt, J. (2020). Estimation of global coastal sea level extremes using neural networks. *Environmental Research Letters*, *15*(7), 074030. https://doi.org/10.1088/1748-9326/ab89d6

Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti, R. (2020). Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics*, *11*(4), 995–1012. https://doi.org/10.5194/esd-11-995-2020

Chattopadhyay, A., Hassanzadeh, P., & Pasha, S. (2020). Predicting clustered weather patterns: A test case for applications of convolutional neural networks to spatio-temporal climate data. *Scientific Reports*, *10*(1), 1317. https://doi.org/10.1038/s41598-020-57897-9

Chavas, D., Lin, N., & Emanuel, K. (2015). A model for the complete radial structure of the tropical cyclone wind field. Part I: Comparison with observed structure. *Journal of the Atmospheric Sciences*, *72*(9), 3647–3662. https://doi.org/10.1175/JAS-D-15-0014.1

Chavas, D., Lin, N., & Emanuel, K. (2016). A model for the complete radial structure of the tropical cyclone wind field. Part II: Wind field variability. *Journal of the Atmospheric Sciences*, *72*(9), 3093–3113. https://doi.org/10.1175/JAS-D-15-0185.1

Chen, K., Kuang, C., Wang, L., Chen, K., Han, X., & Fan, J. (2022). Storm surge prediction based on long short-term memory neural network in the East China Sea. *Applied Sciences*, *12*(1), 181. https://doi.org/10.3390/app12010181

Emanuel, K. (2021). Response of global tropical cyclone activity to increasing co2: Results from downscaling CMIP6 models. *Journal of Climate*, *34*(1), 57–70. https://doi.org/10.1175/jcli-d-20-0367.1

Emanuel, K., Sundararajan, R., & Williams, J. (2008). Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bulletin of the American Meteorological Society*, *89*(3), 347–368. https://doi.org/10.1175/BAMS-89-3-347

Fox-Kemper, B., Hewitt, H. T., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., et al. (2021). Ocean, cryosphere and sea level change. In *Climate change 2021: The physical science basis. Contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.

Gori, A., Lin, N., Xi, D., & Emanuel, K. (2022). Tropical cyclone climatology change greatly exacerbates us extreme rainfall–surge hazard. *Nature Climate Change*, *12*(2), 1–8. https://doi.org/10.1038/s41558-021-01272-7

Gregory, J. M., Griffies, S. M., Hughes, C. W., Lowe, J. A., Church, J. A., Fukimori, I., et al. (2019). Concepts and terminology for sea level: Mean, variability, and change, Both Local and Global. *Surveys in Geophysics*, *40*(6), 1251–1289. https://doi.org/10.1007/s10712-019-09525-z

Guillou, N., & Chapalain, G. (2021). Machine learning methods applied to sea level predictions in the upper part of a tidal estuary. *Oceanologia*, *63*(4), 531–544. https://doi.org/10.1016/j.oceano.2021.07.003

Healy, T., & Harada, K. (1991). Editorial: Enclosed and semi-enclosed coastal seas. *Journal of Coastal Research*, *7*(1), i–v.

Holland, G. J. (1980). An analytic model of the wind and pressure profiles in hurricanes. *Monthly Weather Review*, *108*(8), 1212–1218. https://doi.org/10.1175/1520-0493(1980)108〈1212:AAMOTW〉2.0.CO;2

Horsburgh, K., De Vries, H., Etala, P., Murty, T., Seo, J., Dube, S., et al. (2011). Guide to storm surge forecasting.

Irish, J., Resio, D., & Ratcliff, J. (2008). The influence of storm size on hurricane surge. *Journal of Physical Oceanography*, *38*(9), 2003–2013. https://doi.org/10.1175/2008JPO3727.1

Islam, R., & Takagi, H. (2021). Statistical significance of tropical cyclone forward speed on storm surge generation: Retrospective analysis of best track and tidal data in Japan. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, *15*(3), 247–257. https://doi.org/10.1080/17499518.2020.1756345

Karniadakis, G., Kevrekidis, Y., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning (pp. 1–19). https://doi.org/10.1038/s42254-021-00314-5

Knutson, T., Camargo, S. J., Chan, J. C. L., Emanuel, K., Ho, C.-H., Kossin, J., et al. (2020). Tropical cyclones and climate change assessment: Part II: Projected response to anthropogenic warming. *Bulletin of the American Meteorological Society*, *101*(3), E303–E322. https://doi.org/10.1175/BAMS-D-18-0194.1

Kossin, J. P. (2018). A global slowdown of tropical-cyclone translation speed. *Nature*, *558*(7708), 104–107. https://doi.org/10.1038/s41586-018-0158-3

Lee, J.-W., Irish, J. L., Bensi, M. T., & Marcy, D. C. (2021). Rapid prediction of peak storm surge from tropical cyclone track time series using machine learning. *Coastal Engineering*, *170*, 104024. https://doi.org/10.1016/j.coastaleng.2021.104024

Lin, N., Emanuel, K. A., Smith, J. A., & Vanmarcke, E. (2010). Risk assessment of hurricane storm surge for New York city. *Journal of Geophysical Research*, *115*(D18), D18121. https://doi.org/10.1029/2009JD013630

Liu, Y., Asher, T. G., & Irish, J. L. (2019). Physical drivers of changes in probabilistic surge hazard under sea level rise. *Earth's Future*, *7*(7), 819–832. https://doi.org/10.1029/2019EF001216

Luettich, R., Jr., Westerink, J., & Scheffner, N. (1992). ADCIRC: An advanced three-dimensional circulation model for shelves, coasts, and estuaries. Report 1. Theory and methodology of ADCIRC-2DDI and ADCIRC-3DL Rep. DRP-92-6, 143. *Dredging Research Program Tech*.

Marsooli, R., & Lin, N. (2018). Numerical modeling of historical storm tides and waves and their interactions along the U.S. East and Gulf Coasts. *Journal of Geophysical Research: Oceans*, *123*(5), 3844–3874. https://doi.org/10.1029/2017JC013434

Muis, S., Lin, N., Verlaan, M., Winsemius, H. C., Ward, P. J., & Aerts, J. C. (2019). Spatiotemporal patterns of extreme sea levels along the Western North-Atlantic coasts. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-40157-w

Needham, H. F., & Keim, B. D. (2014). An empirical analysis on the relationship between tropical cyclone size and storm surge heights along the U.S. gulf coast. *Earth Interactions*, *18*(8), 1–15. https://doi.org/10.1175/2013EI000558.1

Nielsen, P. (2009). How storm size matters for surge height. *Coastal Engineering*, *56*(9), 1002–1004. https://doi.org/10.1016/j.coastaleng.2009.02.006

Pandey, S., & Rao, A. D. (2019). Impact of approach angle of an impinging cyclone on generation of storm surges and its interaction with tides and wind waves. *Journal of Geophysical Research: Oceans*, *124*(11), 7643–7660. https://doi.org/10.1029/2019JC015433

Peng, M., Xie, L., & Pietrafesa, L. J. (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. https://doi.org/10.1016/j.ecss.2003.07.010

Proudman, J. (1953). Dynamical oceanography.

Ramos Valle, A., Curchitser, E., Bruyère, C., & McOwen, S. (2021). Implementation of an artificial neural network for storm surge forecasting. *Journal of Geophysical Research: Atmospheres*, *126*(13). https://doi.org/10.1029/2020JD033266

Rego, J., & Li, C. (2009). On the importance of the forward speed of hurricanes in storm surge forecasting: A numerical study. *Geophysical Research Letters*, *36*(7). https://doi.org/10.1029/2008GL036953

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Sun, Y., Zhong, Z., Li, T., Yi, L., & Shen, Y. (2021). The slowdown tends to be greater for stronger tropical cyclones. *Journal of Climate*, *34*(14), 5741–5751. https://doi.org/10.1175/JCLI-D-20-0449.1

Tadesse, M., Wahl, T., & Cid, A. (2020). Data-driven modeling of global storm surges. *Frontiers in Marine Science*, *7*, 260. https://doi.org/10.3389/fmars.2020.00260

Thomas, A., Dietrich, J., Asher, T., Bell, M., Blanton, B., Copeland, J., et al. (2019). Influence of storm timing and forward speed on tides and storm surge during hurricane matthew. *Ocean Modelling*, *137*, 1–19. https://doi.org/10.1016/j.ocemod.2019.03.004

Tiggeloven, T., Couasnon, A., Straaten, C., Muis, S., & Ward, P. (2021). Exploring deep learning capabilities for surge predictions in coastal areas. *Scientific Reports*, *11*(1), 17224. https://doi.org/10.1038/s41598-021-96674-0

Weisberg, R., & Zheng, L. (2006). Hurricane storm surge simulation for tampa bay. *Estuaries and Coasts*, *29*(6), 899–913. https://doi.org/10.1007/BF02798649

Woodruff, J. D., Irish, J. L., & Camargo, S. J. (2013). Coastal flooding by tropical cyclones and sea-level rise. *Nature*, *504*(7478), 44–52. https://doi.org/10.1038/nature12855

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. https://doi.org/10.1007/s13131-018-1081-z

References From the Supporting Information

Hashemi, M. R., Spaulding, M. L., Shaw, A., Farhadi, H., & Lewis, M. (2016). An efficient artificial intelligence model for prediction of tropical storm surge. *Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards*, *82*(1), 471–491. https://doi.org/10.1007/s11069-016-2193-4

Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, *15*(5), 559–570. https://doi.org/10.1175/1520-0434(2000)015⟨0559:DOTCRP⟩2.0.CO;2

Jia, G., & Taflanidis, A. A. (2013). Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment. *Computer Methods in Applied Mechanics and Engineering*, *261–262*, 24–38. https://doi.org/10.1016/j.cma.2013.03.012

Kajbaf, A., & Bensi, M. (2020). Application of surrogate models in estimation of storm surge: A comparative assessment. *Applied Soft Computing*, *91*, 106184. https://doi.org/10.1016/j.asoc.2020.106184

Song, Y. K., Irish, J., & Udoh, I. (2012). Regional attributes of hurricane surge response functions for hazard assessment. *Natural Hazards*, *64*(2), 1475–1490. https://doi.org/10.1007/s11069-012-0309-z