

## ARTICLE OPEN



# Climate change impact on hurricane storm surge hazards in New York/New Jersey Coastlines using machine-learning

Mahmoud Ayyad<sup>1</sup>✉, Muhammad R. Hajj<sup>1</sup> and Reza Marsooli<sup>1</sup>

Recent hurricane losses in the New York Metropolitan area demonstrate its vulnerability to flood hazards. Long-term development and planning require predictions of low-probability high-consequence storm surge levels that account for climate change impacts. This requires simulating thousands of synthetic storms under a specific climate change scenario which requires high computational power. To alleviate this burden, we developed a machine learning-based predictive model. The training data set was generated using a high-fidelity hydrodynamic model including the effect of wind-generated waves. The machine learning model is then used to predict and compare storm surges over historical (1980–2000) and future (2080–2100) periods, considering the Representative Concentration Pathway 8.5 scenario. Our analysis encompassed 57 locations along the New York and New Jersey coastlines. The results show an increase along the southern coastline of New Jersey and inside Jamaica, Raritan, and Sandy Hook bays, while a decrease along the Long Island coastline and inland bays.

*npj Climate and Atmospheric Science* (2023)6:88; <https://doi.org/10.1038/s41612-023-00420-4>

## INTRODUCTION

The heavily populated New York metropolitan area covers a region of narrow rivers, estuaries, islands, and sand barriers with elevations that are <5 m above mean sea level<sup>1</sup>. Millions of residences and transportation, energy, and water infrastructure systems in this area are highly prone to storm surge events. Since 2010, tropical cyclones (TCs) Irene in 2011, Sandy in 2012, Isaias in 2020, Fay in 2020, and Ida in 2021 killed tens of people, damaged thousands of houses, and caused interruptions in clean water and electricity supplies<sup>2–7</sup>. ExtraTropical cyclones (ETCs) are equally hazardous. The Great Appalachian Storms in 1950 produced storm surges that were only 13% smaller than that of Hurricane Sandy (after removing the trend in sea level)<sup>8</sup>.

The expectation that climate change will result in an increase in the frequency of very intense (Categories 4 and 5) hurricanes and an eastward shift in their tracks<sup>9,10</sup> points to a potentially significant change in the level of storm surge hazards to this region over the coming decades. These hazards are quantified by assessing the change in the  $N$ -year peak storm surge height (or water level) return periods, defined as the height that has  $1/N$  percent chance of being exceeded in any given year. While Roberts et al.<sup>11</sup> and Lin et al.<sup>12</sup> found that the impact of ETCs' climatology change would not be significant in this region, other studies showed differing impacts of climate change on TC-induced storm surge levels. Lin et al.<sup>13</sup> used the ADvanced CIRCulation (ADCIRC) hydrodynamic model and found that the 100- and 1000-year total water level return periods at the Battery will, respectively, increase by 10% and 5.5% due to climate change. In another study, Lin et al.<sup>14</sup> determined that sea level rise and TC climatology change will result in an 85% increase in the 100-year return periods. Garner et al.<sup>15</sup> used ADCIRC and showed that the 100- and 1000-year storm surge return periods will, respectively, decrease by 9% and 1.5%. Marsooli et al.<sup>16</sup> showed an increase in the 100-year return period in the New York metropolitan area ranging from 4% to 12% for the future period between 2075 and 2095. Marsooli et al.<sup>17</sup> studied the impact of climate change on hurricane flood hazards in Jamaica Bay, NY

using a data set for the future period of 2080–2100. Their results showed that the 100- and 1000-year total water levels, excluding wave effects, will increase by 10% and 14%, respectively. The wave setup contributes to the total water level by transferring the radiation stresses from breaking waves to the water column, which in turn increases the water level by 5–35%, based on the continental shelf size and slope<sup>18–20</sup>. However, none of the previous studies included the effect of wind-generated waves on storm surge heights and return period levels.

Reliable estimates of the  $N$ -year flood levels under different climate change scenarios must be based on data from  $n \times 10 \times N \times f$  storms<sup>21</sup>, where  $n$  is the number of climate models, and  $f$  represents the annual storm frequency. Ayyad et al.<sup>22</sup> estimated that, for the 1000-year return period, 600,000 storm scenarios should be considered when assuming a frequency of five storms per year and an ensemble of six climate models. Given the limited number of historical storms, numerical modeling and simulations of synthetic storms must be used in estimating low-probability high consequence events such as 100 and 1000-year return periods. Performing hydrodynamic simulations for a large number of storm scenarios is a computationally intense task. To reduce the computational burden, simplified hydrodynamic models (e.g. SLOSH) or coarser computational meshes are used, which could compromise and lead to lower confidence in the results. An alternative approach to performing an extensive number of hydrodynamic simulations is to use Artificial Intelligence tools. Lee et al.<sup>23</sup>, Tseng et al.<sup>24</sup>, Lee et al.<sup>25</sup>, De Oliveira et al.<sup>26</sup>, and Kim et al.<sup>27</sup> used artificial neural network (ANN) models to predict surge levels due to synthetic or historical Typhoon data. Hashemi et al.<sup>28</sup> used ANN models to predict maximum water elevation based on synthetic TCs, while Lee et al.<sup>29</sup> predicted the peak storm surge height using tropical cyclone time series. Tiggel et al.<sup>30</sup> predicted hourly surge time series of the global tide and surge model forced using atmospheric variables. Jia et al.<sup>31,32</sup> and Kyrioti et al.<sup>33</sup> used surrogate models to comprehensively assess risk and coastal hazard. Ayyad et al.<sup>22,34</sup> were the first to demonstrate the use of ANN and

<sup>1</sup>Davidson Laboratory, Department of Civil, Environmental and Ocean Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA. ✉email: m-mayyad@outlook.com

machine learning (ML) models to reliably predict the  $N$ -year return periods for the idealized computational domain and real domain, respectively. They used high-fidelity numerical simulations of synthetic TCs to train, validate, and test different ML models that predicted the peak storm surge height. Then, they used the predicted storm surge heights to generate return period curves. The generated curves from one of the ML models showed good agreement with those generated from the hydrodynamic simulations but at a small fraction of the computational time and resources.

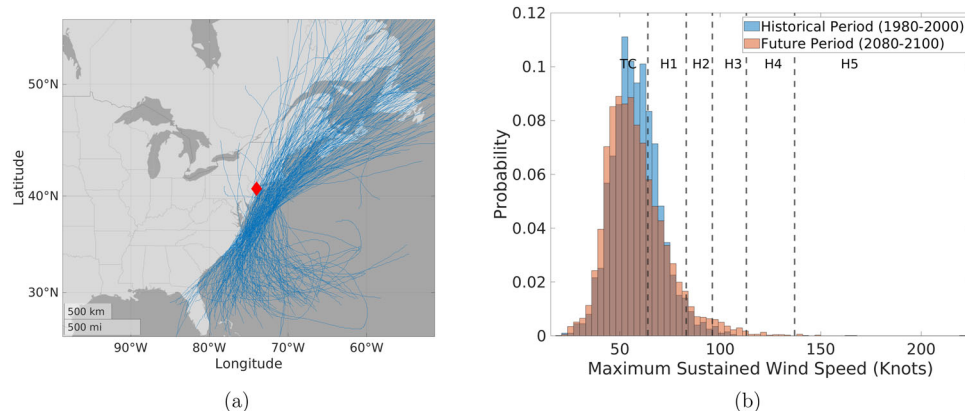
In this study, we implement a machine learning model<sup>34</sup> to assess variations in probabilistic flood hazards from storm surges due to TC climatology change. Furthermore, we include the effects of wind-generated waves that were not considered in previous studies. Because of uncertainties in the TCs timing, we neglect the effect of tides. As demonstrated in recent events, the storm surge could coincide with low tides and cause less hazards as in the case of hurricane Isaias 2020<sup>35</sup>, or with high tides as in the case of hurricane Sandy 2012<sup>36</sup>, which caused significant damage. The peak storm surge heights used to train the ML model are calculated using numerical simulations of the coupled ADvanced CIRCulation and Simulating WAVes Nearshore (ADCIRC + SWAN) model<sup>20</sup>. These simulations are used to train different ML models at 57 sites in the New York metropolitan area. The sites cover a combination of open coast and inner bay sites. The trained ML models are then used to predict the peak storm surge height considering the influence of wind setup, the inverted barometer effect, and wind-generated waves from TC synthetic storms. Data sets for storms covering the historical period of 1980–2000 (late twentieth century) and the projected future period of 2080–2100 (late twenty-first century) are used to determine the impact of climate change. The TCs in the future period are generated under the representative concentration pathways (RCP) 8.5 greenhouse gas concentration scenario. Historical and future TC data sets are based on projections from four global climate models. We also utilize TC data sets generated for the observed climate of the historical period of from the National Centers for Environmental Prediction (NCEP) reanalysis<sup>37</sup> to correct any bias resulting from the climate models. Here, we used the predicted storm surges from the ML models to determine probabilistic flood events up to the 500-year return period to avoid the large statistical uncertainty in higher return periods, which is due to the limited number, between 1800 and 2000, of synthetic TCs in each dataset.

## RESULTS AND DISCUSSION

### Tropical cyclones climatology

Assessing the impact of tropical cyclone (TC) climatology change on probabilistic flood hazards requires a large number of TCs. Due to the limited number of historical events, synthetic TCs must be used. For the purpose of this study, we use the synthetic TC data sets from Marsooli and Lin<sup>17</sup>. These synthetic TCs were generated using the statistical/deterministic hurricane model of Emanuel et al.<sup>38</sup>. This hurricane model generates synthetic TCs for atmospheric and oceanic conditions determined from either observations or climate models. Synthetic TCs from the observed climate over the historical period of 1980–2000 (late twentieth century) were generated based on the National Centers for Environmental Prediction (NCEP) reanalysis<sup>37</sup>. Synthetic TCs were also generated based on the modeled climates of the same historical period and the future period of 2080–2100 (late twenty-first century). Four global climate models, namely the GFDL5 (Geophysical Fluid Dynamics Laboratory Climate Model, USA)<sup>39,40</sup>; HadGEM5 (Hadley Centre Global Environment Model, U.K. Meteorological Office)<sup>41</sup>; MPI5 (Max-Planck-Institute for Meteorology, Germany)<sup>42</sup>; and MRI5 (Meteorological Research Institute, Japan)<sup>43</sup> were used to generate the synthetic TCs. A total of nine TC data sets were generated. Four climate models were used to generate eight data sets for historical and future periods and the observed climate was used to generate one data set for the historical period. The histograms, density maps, and storm surge return period results of the future period are bias corrected using the NCEP data set and presented as the weighted average over the four climate models, as described in the methods section.

We consider TCs that pass within a 350 km radius of the New York metropolitan area. Tracks of a random sample of the generated synthetic TCs for the historical and future periods are shown in Fig. 1a. About 1800 synthetic TCs from each climate model for the historical and future periods are considered. The number of TCs in the NCEP data set is around 2000. One important characteristic of the TC, used in their classification, is its intensity defined by the maximum sustained wind speed. Because the wind speed varies as the TC travels along its track, we define the TC intensity by its wind speed at the closest location to the Battery. Histograms of TCs' maximum sustained wind speed at the closest location to the Battery for the historical and future periods are shown in Fig. 1b. The historical period histogram shows TCs of the NCEP data set, while the future period histogram shows the weighted average of the TCs from the four future climate models. The histograms show right-skewed distributions. More importantly, the plots show an increase in the number of intense



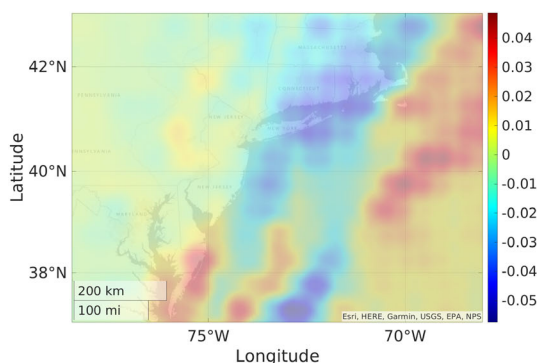
**Fig. 1 Synthetic TCs information from the NCEP and the four climate models data. a** Tracks of a random sample of the synthetic TCs that pass within 350 km of the New York metropolitan area. The red diamond shows the location of the Battery station. **b** Weighted histograms of TCs' maximum sustained wind speed at the closest point to Battery for historical (1980–2000) and future (2080–2100) periods. H1–H5 indicate the five hurricane categories.

hurricanes, i.e. Categories 3 and above, and a decrease in the number of less intense TCs that reach the New York metropolitan area in the future period. The percentages of every category out of the total number of storms in the historical and future periods data sets are presented in Table 1. The TC data sets show that the percentages of tropical storms and Category 1 hurricanes that reach the metropolitan area would decrease in the future period, while that of Category 3 hurricanes would triple in the future. Although the data sets show no storms higher than Category 3 reached the study area in the historical period, they show some high-intensity storms (Categories 4 and 5) reach this area in the future period. This projected increase in TC intensity in the future climate is consistent with most other projections<sup>10</sup>.

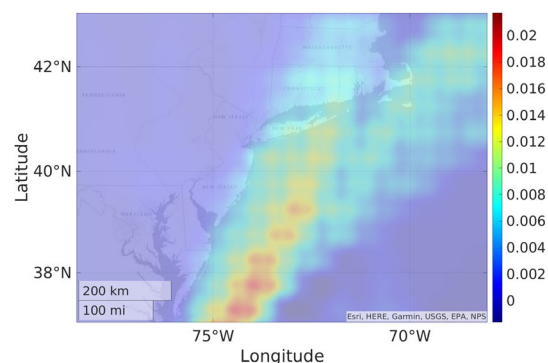
The difference in the weighted density distribution of the TC tracks over our study area is presented in Fig. 2. The density distribution is defined as the number of TCs that cross through a grid box of size  $0.5^\circ \times 0.5^\circ$  and normalized by the area of this grid. The historical period density distribution is calculated using the NCEP data set while that of the future period is the weighted average over the biased corrected density distribution of the four data sets. Then, the difference between the density distributions of both periods is shown in the plot. Figure 2a, which shows the density plot of all the TCs, indicates an eastward shift in storm tracks by the end of the twenty-first century. The TC data sets show more TCs would travel toward Delaware Bay, Cape May, and south of Atlantic City, impacting Delaware and the southern part of the NJ coastline. A lesser number of TCs would impact the western side of Long Island. These findings are consistent with those of Garner et al.<sup>15</sup> who noted an offshore shift of storm tracks. Figure 2b, which shows the density plots of the intense hurricanes only, shows a higher number of intense hurricanes would travel parallel to the NJ coastline in the future. The TC data sets show that a smaller number of these hurricanes would either make landfall on the NJ coastline or on the east side of Long Island. Since the highest wind speeds are encountered on the right side of the hurricane's center, where the total speed is the sum of the hurricane's forward speed and its rotational speed, notable effects on the storm surge height will depend on its track, whether to the right or left of the site of interest.

**Table 1.** The weighted proportions of tropical storms and the five hurricane's categories at the closest location to Battery for the historical (1980–2000) and future (2080–2100) periods (%).

	Tropical storm	Cat. 1	Cat. 2	Cat. 3	Cat. 4	Cat. 5
History	71.3	25.7	2.3	0.7	0	0
Future	68.2	21.9	3.7	2.1	0.5	0.2



(a) All TCs



(b) Intense Hurricanes (Cat. 3–4)

**Fig. 2** The change in the TC track over future (2080–2100) and historical (1980–2000) periods. Differences in the weighted density distribution of **a** all TCs, and **b** intense hurricanes' tracks over future (2080–2100) and historical (1980–2000) periods.

## Study area

We assess the effect of climate change on storm surge levels at 57 locations along the NJ and NY coastlines, as shown in Fig. 3. These sites are chosen to broadly investigate the spatial variation of the storm surge levels along the NJ and NY coastlines. Specifically, the sites cover five highly populated regions that include the NJ open coastline, Rockaway and Long Beach barrier islands at Long Island, Jamaica bay, upper bay, and lower bay. The locations included 38 equally distributed sites that are 5 km apart along the NJ open coastline between Cape May and Sandy Hook, seven sites along Rockaway and Long Beach barrier islands in Long Island, NY, five sites along the perimeter of Jamaica bay, six sites in Lower, Raritan, and Sandy Hook bays to capture the effect of topography change and the Battery. Of these locations, we choose six representative sites of the five regions for a detailed assessment of the impact of climate change.

## Future return period flood levels

The effect of climate change is assessed by studying the low-probability high consequence flood levels. Because just about 2000 TCs were generated for each climate model, the maximum number of years for a reliable return period is limited to 500 years. Figure 4 shows the Pareto distribution fit of the peak TC storm surge heights return period along with the 90th percentile confidence interval for the six representative study sites for both historical and future intervals. The plots show a negative or no change in the return periods between historical and future periods at Battery and Long Island. In contrast, increased flood levels are noted at the other stations in the future period due to climate change, especially in the 500-year return period levels. The highest impact is in Raritan Bay.

Figure 5 shows the spatial variation along the NJ coastline and in the New York metropolitan area of the 100- and 500-year return periods of the historical period. The plot also shows corresponding percentage changes between the future ( $\eta_f$ ) and historical ( $\eta_h$ ) periods, defined as

$$\text{Percent of change} = \frac{\eta_f - \eta_h}{\eta_h} * 100\%, \quad (1)$$

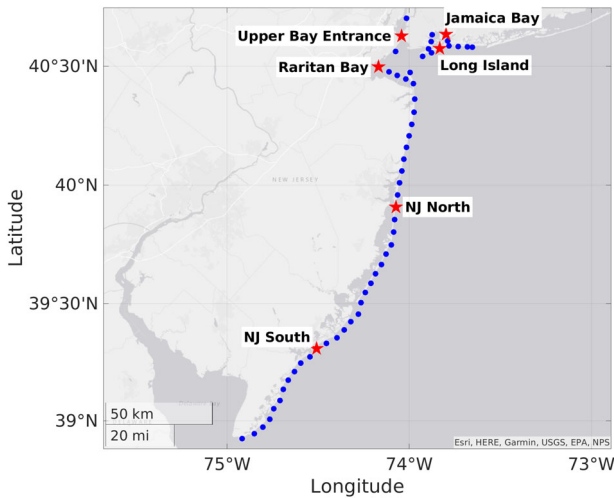
For the historical period, the 100- and 500-year storm surge return periods are noted to have higher flood levels along the southern parts of the NJ coastline (south of Atlantic City) than the northern part of the NJ coastline and the inner bays. This happens because the TCs move parallel and close to the NJ coastline, as discussed above, and due to the topography of the southern part of the NJ coastline. Along the coastline of Long Island, higher flood levels are because this coastline faces the track of most TCs.

Also, the inner bays (Sandy Hook, Raritan, Jamaica, upper, and lower bays) experience relatively large 100- and 500-year peak storm surge return periods because the bays have a concave inlet, which amplifies positive surges through the coastal funneling effect<sup>35,44</sup>.

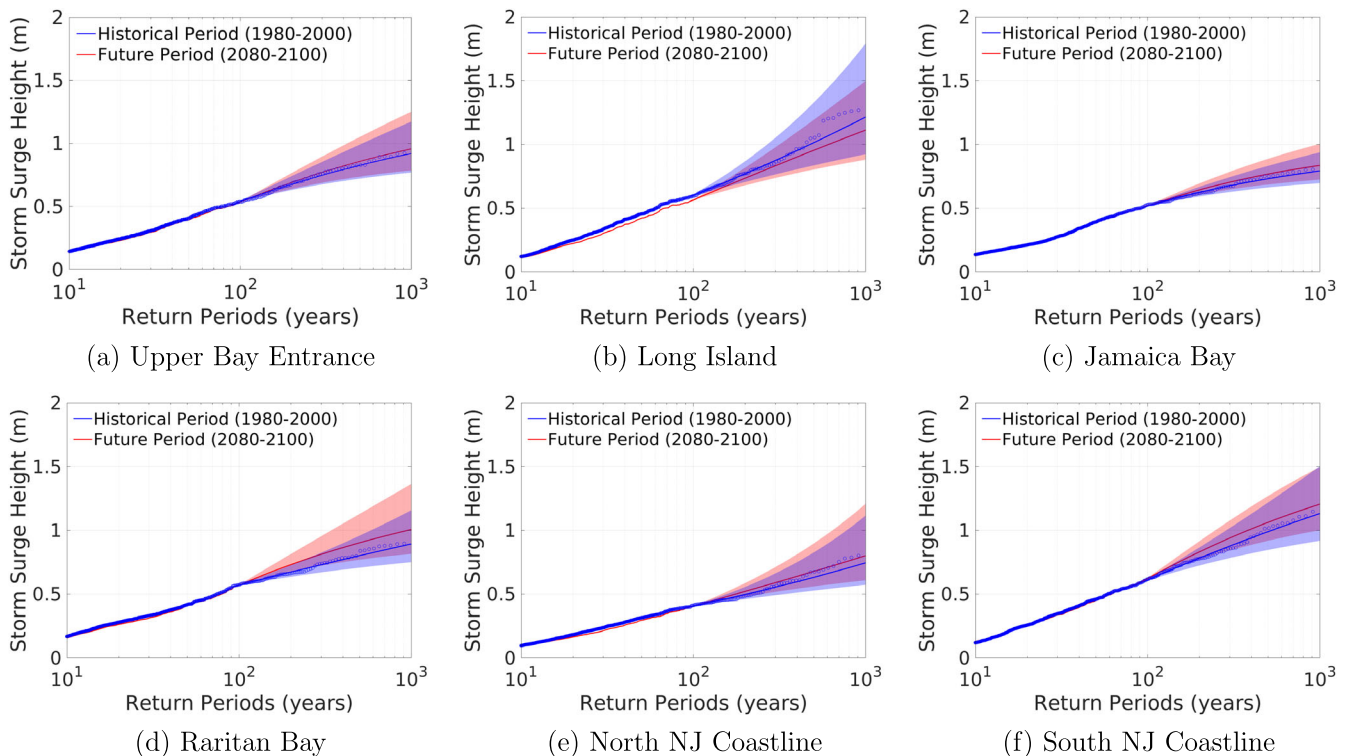
In the future period, the 100-year return period flood levels along the northern part of NJ coastline would decrease slightly by percentages up to 3.5%. In contrast, along the southern part of the NJ coastline, climate change would induce an increase in the 100-year return by up to 4%. This happens because the number of TCs, in the synthetic data sets, traveling toward the Northern area of our domain would decrease, while more TCs could make landfall

or be directed toward the southern part of NJ and Delaware bay, as shown in Fig. 2a. Because more intense hurricanes in the future period could track toward the coastline, as deduced from Fig. 2b, the percent changes of the 500-year return period along the NJ coastline ranges from zero in the northern part to 12% in the southern part. In contrast, negative changes in the 100- and 500-year return periods are noted along the coastline of Long Island as the calculated density of synthetic TCs that pass by this area would decrease in the future period in comparison to that in the historical period. Also, most of the intense hurricanes in the data set would not reach the Long Island coastline and deviate eastwards, such that this coastline will be on the west side of hurricanes, which is less prone to high storm surges. On the other hand, the changes in the 100-year return period are between -3.5% at the Battery and 1.5% in Jamaica bay because the density of the TCs that pass by this area would decrease. However, a 12% increase in the 500-year return periods would occur at Sandy Hook and in Raritan Bay, while up to 5% increase would occur in the upper- and lower bays and in Jamaica bay. This is because the storm data sets show that in the future period, more intense hurricanes would approach the southern part of the bays' entrance, as shown in Fig. 2b.

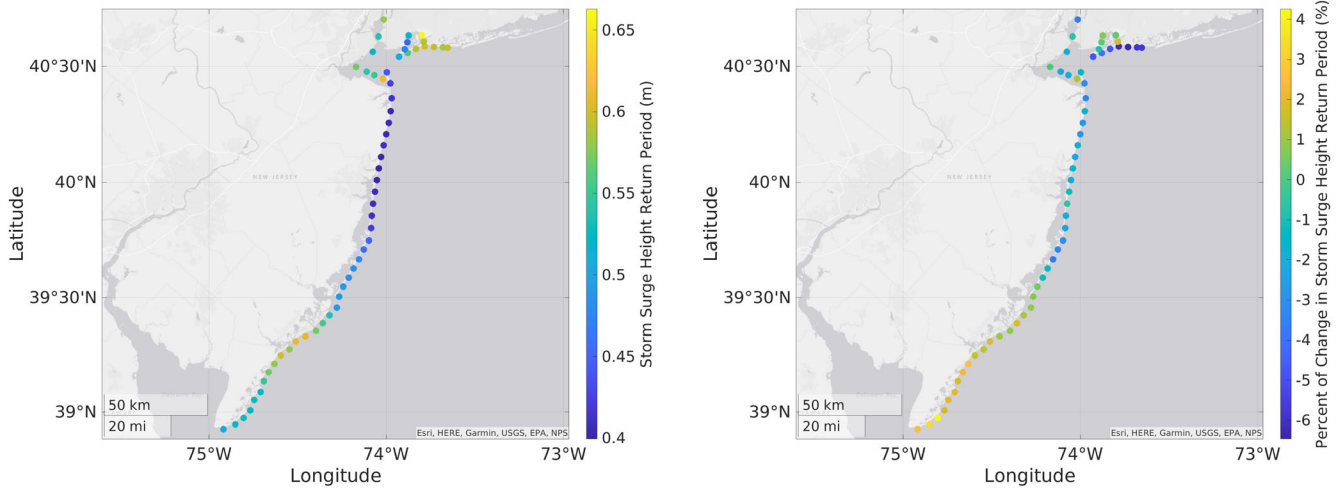
Predicting low-probability events while considering uncertainties in climate change and using different approaches is challenging. As such, it is important to compare the above results to those that have been previously published under different assumptions. Garner et al.<sup>15</sup> predicted a negative change in the storm surge height return periods at the Battery, which matches our findings. The reason for the difference in the percentage may be because they used different data sets and different domains. Also, their future projections were not bias-corrected and their study did not account for the effect of waves. Another study by Marsooli et al.<sup>16</sup> showed an average increase of 6% in the 100-year return period at the southern part of the New Jersey coastline



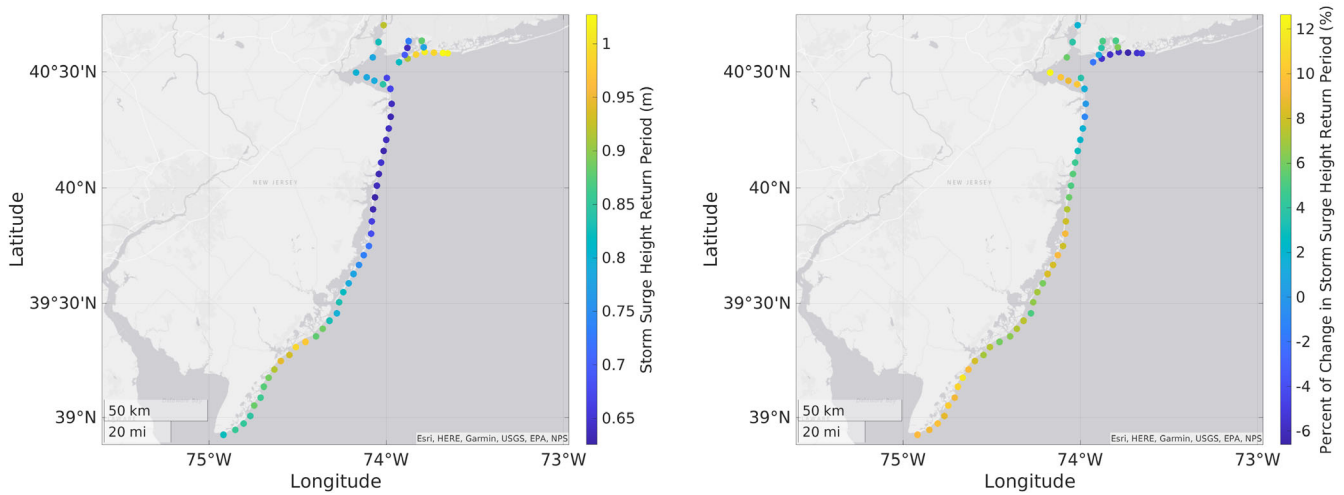
**Fig. 3** The map showing the study sites in blue circles, and red stars for representative sites.



**Fig. 4** TC storm surge height return period curves for the six representative stations. The lines show the Pareto distribution fits of both historical and future periods, while the shaded areas show the corresponding 90th percentile confidence interval at **a** Upper Bay entrance, **b** Long Island, **c** Jamaica Bay, **d** Raritan Bay, **e** North NJ, and **f** South of NJ study sites.



(a) 100-year return period



(b) 500-year return period

**Fig. 5** TC storm surge height return period results over the 57 study sites. The **a** 100-year, and **b** 500-year return period curves over the 57 study sites. The left panels show the storm surge height return periods, in meters, for the historical period, while the right panels are the percent of change between historical and future return periods based on the RCP 8.5 scenario.

which is slightly larger than our 4% prediction. Also, Lin et al.<sup>14</sup>, and Marsooli et al.<sup>16</sup> showed positive changes in the Battery which contradicts our findings and those of Garner et al.<sup>15</sup>. This discrepancy might be because Lin et al.<sup>14</sup> and Marsooli et al.<sup>16</sup> considered the effects of tides. They also used a different computational mesh. The same reason applies to the predictions of percentage change in Jamaica bay by Marsooli et al.<sup>17</sup> who showed a positive change but with slightly higher magnitudes than current predictions.

When considering the above results, it is important to recognize the sources of uncertainties associated with the projections. One uncertainty results from the nonlinear interaction between tides and storm surge, which could impact total water level, especially in the case of low-tides<sup>45,46</sup>. However, adding tides to the surge adds large uncertainties in predicted water levels because of the high uncertainty in the hurricanes' timing. Another source of uncertainty is the small variation in the size of the synthetic TC, represented by the radius of maximum sustained wind speed. The considered radii range varied between 40 and 60 km restricting the ML model to this range. Moreover, using the ML model to predict the storm surge heights of more severe TCs than those

provided in the training data set would include higher uncertainties. These variations can be incorporated into future studies when more data on synthetic storms becomes available.

The study of the effect of greenhouse warming on the vulnerability of coastal areas to flood hazards is important for long-term development and planning. Such studies require a large number of TCs, which are not available in the historical data. We use thousands of synthetic TCs to determine the impact of climate change on storm surges along the shores of New York and New Jersey. Previously, high-fidelity hydrodynamic models were used to predict storm surge heights from these TCs, which is computationally expensive. To reduce the computational burden and consider a broader TC distribution, we develop machine learning models to predict the peak storm surge height from the TCs. The modeling and analysis were applied at 57 different locations along NJ, Long Island coastlines, and the Raritan, Sandy Hook, Jamaica, lower, and upper bays. The ML model was trained, validated, and tested using a data set generated from the ADvanced CIRCulation and Simulating WAVes Nearshore (ADCIRC + SWAN) coupled model, including the effects of waves on storm surge levels. TC parameters including, maximum

sustained wind speed, upper- and lower-latitudinal, right- and left-longitudinal distances at three different moments, and the minimum distance between the study site and TC eye, constituted the input features to the ML model. To discern the effects of TC climatology change, we applied the ML model to synthetic storms over historical (1980–2000), and future (2080–2100) periods. Analysis of these storms shows that while climate change would cause more high-intense hurricanes to reach higher latitudes including the study area, most of them would shift further to the east. The analysis also shows a higher probability that storms would make landfall in the southern part of the NJ coastline in the future period. The ML-predicted storm surge heights were used to generate return period curves. The results indicate a decrease in the 100- and 500-year storm surge return periods at the Battery and Long Island coastline, in agreement with some previous results. In contrast, an increase in the storm surge levels is noted over the southern parts of NJ coastline and in the inner bays, also in agreement with previous results. These results demonstrate the capability of using ML models, at a fraction of the computational cost of high-fidelity simulations, when performing risk-informed coastal planning, development, or management that require consideration of uncertainties associated with climate change.

## METHODS

### Training data set, TC parameters, and feature selection

The training data set of the ML storm surge model was generated from the coupled ADCIRC + SWAN simulations of storm surge resulting from 10,300 synthetic TCs that were used in Ayyad et al.<sup>34</sup> but for different stations except for Battery and stations close to Atlantic City. Thus, the used data sets have the same characteristics when compared to that used in Ayyad et al. (2022). The used data set spans all five hurricane categories and tropical storms. The data set was imbalanced such that the majority of the simulated TCs generate a peak storm surge height of more than 0.5 m. Training an ML model using this imbalanced data set generates a surrogate model that underestimates the predicted storm surge heights. Thus, we followed the procedure of Ayyad et al.<sup>34</sup> by dividing the data set into two smaller ones. The two data sets include TCs that pass within and outside a radius of 100 km from the location of interest, referred to as DS-1 and DS-2, respectively. Following the findings of Ayyad et al.<sup>34</sup>, the ML model input features include six parameters that identify the TC characteristics, namely the maximum sustained wind speed, radius of maximum wind, upper and lower latitudinal distances, and right and left longitudinal distances, at three-time steps, namely 6-h post, 0-, and 6-hours prior to the time of the closest TC location to the study site. Also, the minimum distance between the TC location to the study site is added to the ML model features which makes a total number of 19 features used. Feature selection was conducted by Ayyad et al.<sup>34</sup> using the correlation coefficient and mutual information values between the input feature and corresponding simulated storm surge height of the training part of the two data sets separately. Only 13 different features for each of the two data sets were chosen. The radius of maximum wind speed, left longitudinal and upper latitudinal distances were removed from both data sets, while the lower latitudinal and right longitudinal distances were removed from DS-1 and DS-2, respectively.

### Machine learning model and performance metrics

For each of the 57 study sites, we generate two ML models, one for the DS-1 data set and the other for the DS-2 data set. Thus, a total of 114 ML models are trained. Different ML algorithms can be used to train the surrogate model. Ayyad et al.<sup>34</sup> found that Adaptive Boost (AdaBoost) algorithm with support vector regressor (SVR) as the base estimator has the best performance

among seven different ML algorithms. Tuning the hyper-parameters is crucial to avoid generating either an under- or over-fitted model. To tune the model, we firstly divided the DS-1 and DS-2 models to 60% for training, 20% for validation, and the rest for testing the model's performance. The training was performed using the scikit-learn library on Python<sup>47</sup>. The used hyper-parameters were tuned using the cross-validation grid search method, by trying all possible combinations of the hyper-parameters and getting the best-performing configuration for training. Ayyad et al.<sup>34</sup> found that the tuned hyper-parameters of DS-1 data set differ from those of the DS-2 data set. For DS-1, the learning rate is set to 0.09, the number of weak learners is set to 15, and the regularization and epsilon-insensitive loss parameters of SVR are set to 90 and 0.09, respectively. For the DS-2 data set, the learning rate is set to 0.05, the number of weak learners is set to 50, and the regularization and epsilon-insensitive loss parameters of SVR are set to 65 and 0.03, respectively. Exponential loss function is used for the two data sets.

The predicted peak storm surge height from the trained machine learning model ( $\eta_p$ ) should be evaluated against those calculated from ADCIRC + SWAN model ( $\eta_a$ ). In this study, we used the correlation coefficient ( $R$ ), defined as

$$R = \frac{\text{cov}(\eta_a, \eta_p)}{\sigma_{\eta_a} \sigma_{\eta_p}}, \quad (2)$$

coefficient of determination ( $R^2$ ), defined as

$$R^2 = 1 - \frac{\sum_{i=1}^N (\eta_{ai} - \eta_{pi})^2}{\sum_{i=1}^N (\eta_{ai} - \bar{\eta}_{ai})^2}, \quad \text{and} \quad (3)$$

Root mean square error (RMSE), defined as

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\eta_{ai} - \eta_{pi})^2} \quad (4)$$

to evaluate the ML model's performance, where  $\text{cov}(\dots)$  and  $\sigma$ , respectively represent the covariance and standard deviation,  $N$  is the size of the data set, and the overline represents the average value. A perfect match is considered when  $R$  and  $R^2$  are equal to one, while RMSE is equal to zero.

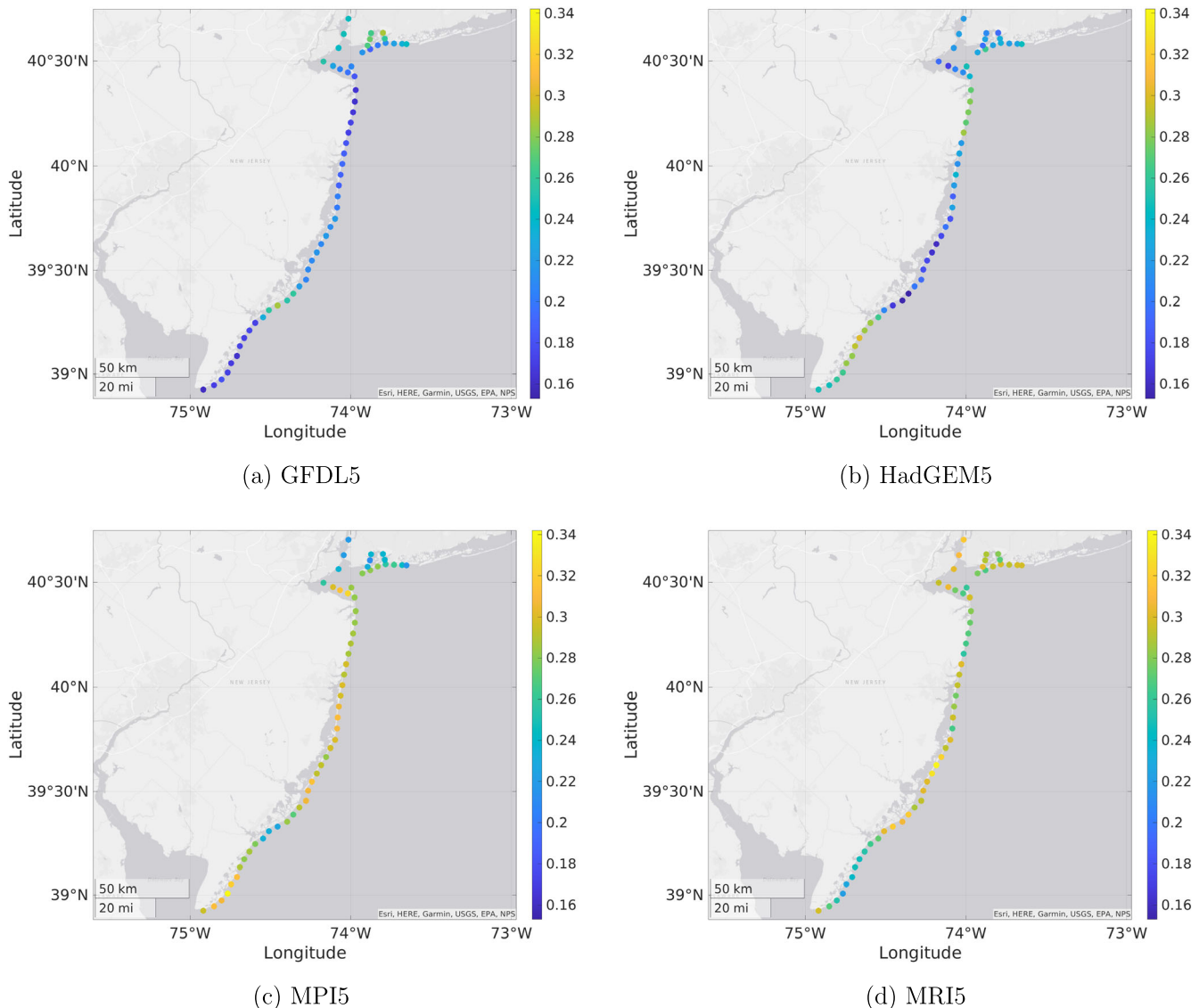
### Statistical model and bias correction

The probabilistic flood hazard assessment, shown in Figs. 4 and 5, is presented in terms of the peak storm surge height at the 57 study sites. The return period ( $T$ ) of storm surge height ( $\eta_{TC}$ ) that exceeds a given threshold ( $h$ ) is given by<sup>14</sup>

$$T = \frac{1}{F(1 - P[\eta_{TC} \leq h])} \quad (5)$$

where the equation assumes a stationary Poisson distribution for the storm's arrival.  $F$  is the TC annual frequency and  $P[\eta_{TC} \leq h]$  is the cumulative density function of the TC storm surge heights where its long tail is modeled using the generalized Pareto distribution (GPD). The GPD threshold value is selected by trial and error so that the modeled CDF well represents the data points. The GPD controlling parameters are estimated using the maximum-likelihood method.

Global climate models (GCMs) are designed to model the climate globally and, as such, they may be biased. We used the approach adopted by Lin et al.<sup>14</sup>, and Marsooli et al.<sup>16,17</sup> to correct the bias in the TC frequency, density, and return period curve estimates for each GCM. The bias is calculated by subtracting the model-based estimates, i.e. results based on the GCM data set, for the historical period from the NCEP-based estimates, i.e. results based on the NCEP data set. Then, this bias is removed from the model-based estimates for the future period.



**Fig. 6** The weighting factors of the four GCMs. The weighting factors of **a** GFDL5, **b** HadGEM5, **c** MPI5, and **d** MRI5 global climate models over the 57 study sites.

The presented results in this study are the weighted average of the results of the four used GCMs. A weighting factor ( $W_i$ ) is allocated for each GCM ( $i$ ) which is given by

$$W_i = \frac{S_i}{\sum_{i=1}^4 S_i} \quad (6)$$

where  $S_i$  is the Willmott skill score of GCM  $i$ . This score is determined by comparing the NCEP-based storm surge height return period with the one projected by the GCMs for the historical period, which is defined as<sup>48,49</sup>

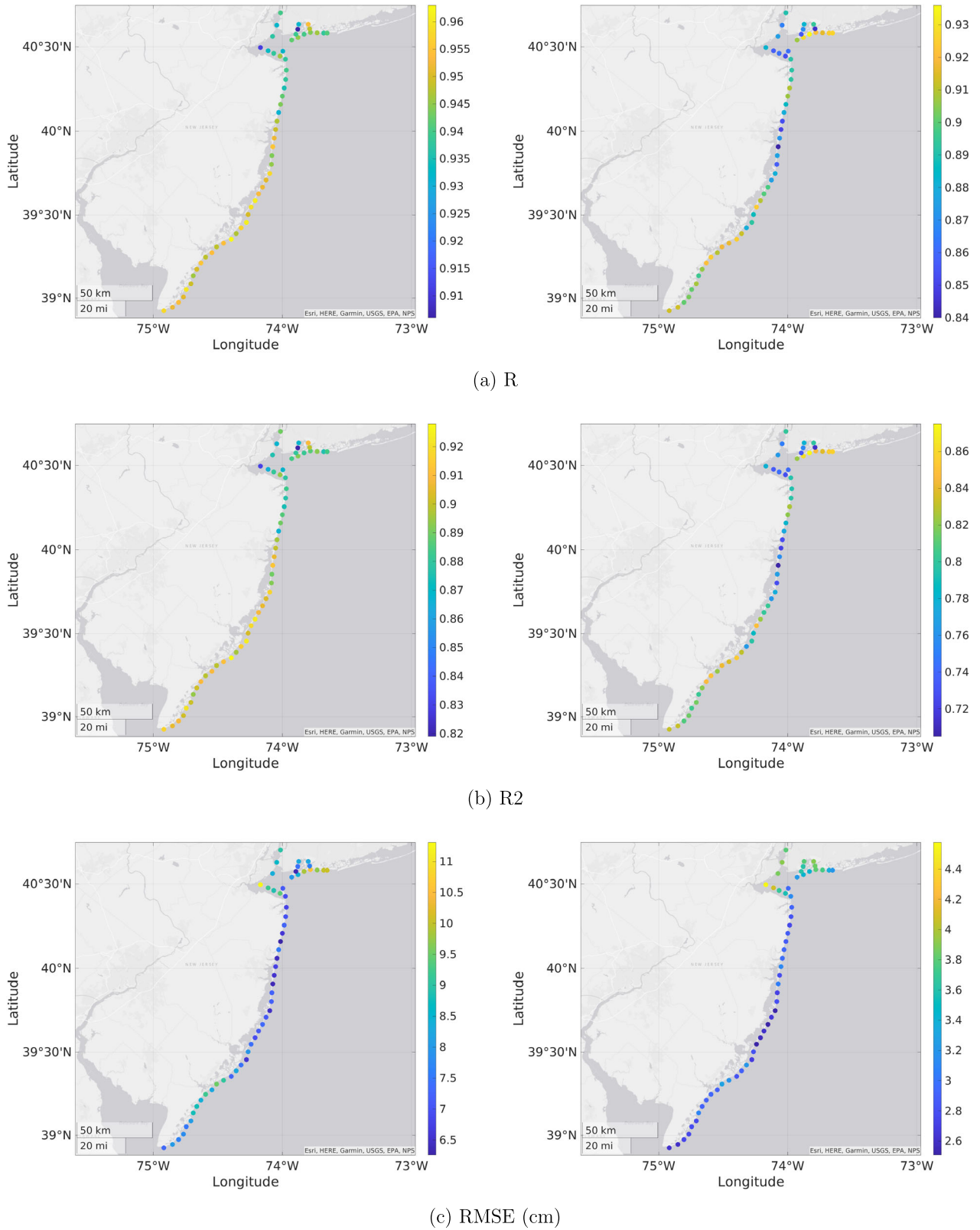
$$S_i = 1 - \frac{\sum_{j=1}^N X_{j\text{GCM}_i} - X_{j\text{NCEP}}^2}{\sum_{j=1}^N \left( X_{j\text{GCM}_i} - \bar{X}_{j\text{NCEP}} + X_{j\text{NCEP}} - \bar{X}_{j\text{NCEP}} \right)^2} \quad (7)$$

where,  $X_j$  is the  $j$ th-year storm surge return period which is summed over the  $N$  years, and  $\bar{X}_j$  is the mean value of  $X_j$ . The skill value ranges between zero and one with a perfect agreement at one. The weights of the four climate stations are presented in Fig. 6. The weights range between 0.16 and 0.34. The weights show that the historical data of MPI5 and MRI5 GCMs matches those of the NCEP data along the NJ coastline, while the weights of all GCMs are

almost the same over the other regions. This approach was used previously by Lin et al.<sup>12</sup>, and Marsooli et al.<sup>16,17</sup>.

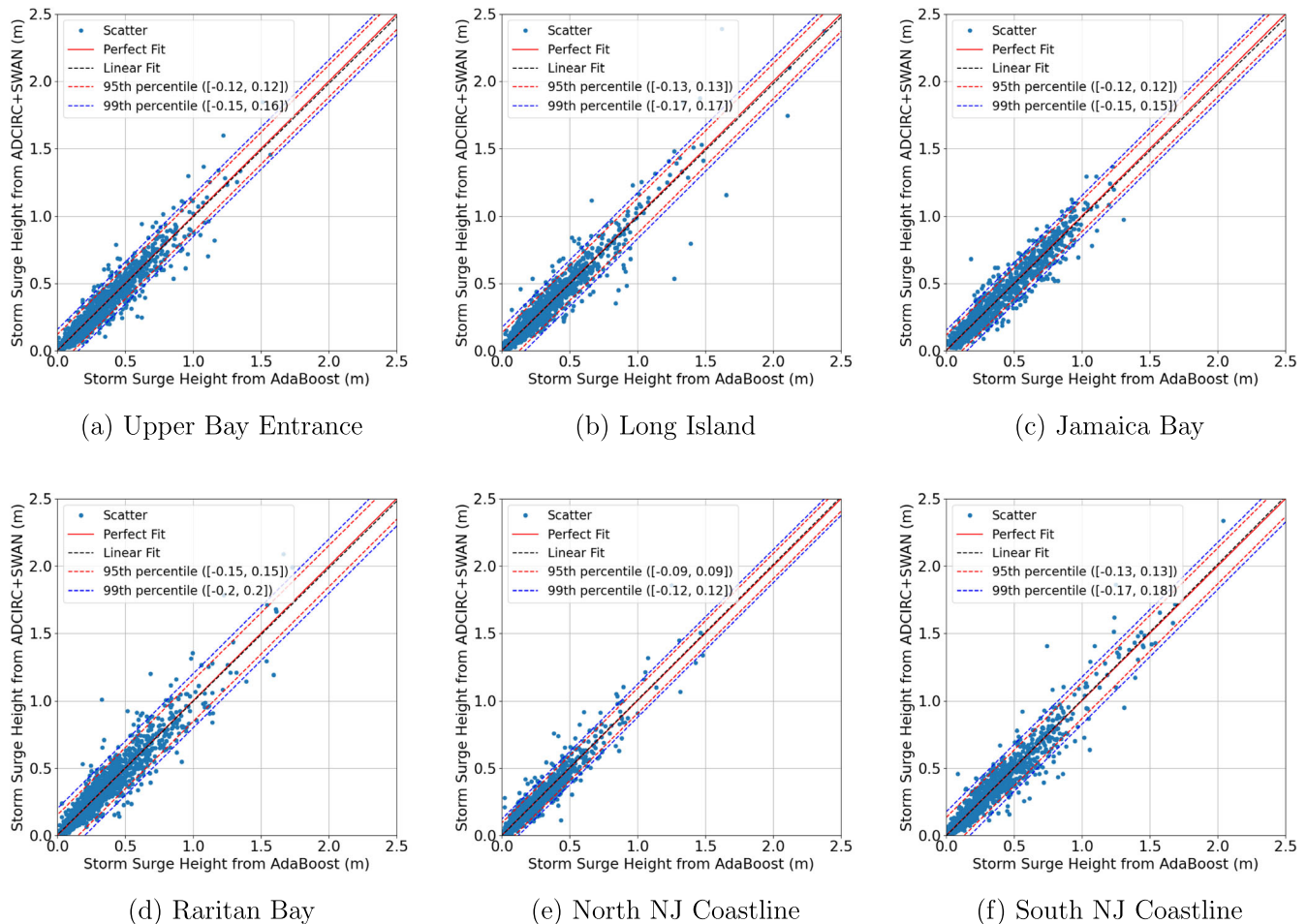
### Model validation

To establish the goodness of the model fit, we compare peak storm surge heights as predicted from the ML model and those calculated from ADCIRC + SWAN model for the 57 study sites. The trained ML models are tested using the DS-1 and DS-2 test data sets. The performance metrics, namely the correlation and determination coefficients and the RMSE, are presented in Fig. 7. The plots show that the minimum values of correlation coefficient ( $R$ ) and coefficient of determination ( $R^2$ ) of the two data sets are 0.84, and 0.71, respectively. The RMSE ranges between 6.5 and 11 cm for data set DS-1, which has a maximum storm surge height value of 2.2 m, and between 2.5 and 4.5 cm for data set DS-2, which has a maximum storm surge height value of 90 cm. Given that peak surge heights are respectively up to 2.5 and 0.9 m for data sets DS-1 and DS-2, the relatively low RMSE values assert the goodness of the ML predictions. The scatter plots of Fig. 8, showing peak storm surge heights calculated using ADCIRC + SWAN model with those predicted from the ML models, confirm



**Fig. 7** Performance metrics of the ML models at the 57 study sites using the test data set. The **a** correlation coefficient (R), **b** coefficient of determination (R<sup>2</sup>), and **c** Root Mean Square Error (RMSE) at the 57 study sites. The left and right panels show the validation of data sets DS-1 and DS-2, respectively.





**Fig. 8 ML model validation results at the six representative sites.** Scatter plots of the peak storm surge height calculated using the ADCIRC + SWAN simulations and those predicted from ML models at **a** Upper Bay entrance, **b** Long Island, **c** Jamaica Bay, **d** Raritan Bay, **e** North NJ, and **f** South of NJ study sites.

further the goodness of the ML model. The plots are presented for the six representative stations. For a perfect fit, the scatter points should fit a diagonal line having a slope of 1. The slopes of the linear fit for the representative sites range between 0.985 and 1.001. The mean, and standard deviation of the error, defined by the difference between the storm surge height calculated using ADCIRC + SWAN and predicted from ML models for the two test data sets, at the six stations are, respectively, near zero and 7.8 cm. The reference lines, represented by dashed lines, in Fig. 8 indicate the 95th and 99th percentiles calculated from the normal distribution fits of the errors, on average 53 out of 2072 storms are outside the 99th percentile range and 107 are outside the 95th percentile range. These results show that, to a great extent, the peak storm surge heights predicted from the ML models match those simulated by the ADCIRC + SWAN model.

#### DATA AVAILABILITY

The trained machine learning models are available in the GitHub repository [https://github.com/m-ayyad/ML\\_Hurricane\\_Surge\\_NY\\_NJ.git](https://github.com/m-ayyad/ML_Hurricane_Surge_NY_NJ.git). While the training and climate models data sets are available upon request.

#### CODE AVAILABILITY

A Jupyter notebook example on how to use the machine learning models is available in the GitHub repository [https://github.com/m-ayyad/ML\\_Hurricane\\_Surge\\_NY\\_NJ.git](https://github.com/m-ayyad/ML_Hurricane_Surge_NY_NJ.git).

Received: 6 December 2022; Accepted: 5 July 2023;  
Published online: 14 July 2023

#### REFERENCES

- Colle, B. A. et al. New York city's vulnerability to coastal flooding: storm surge modeling of past cyclones. *Bull. Am. Meteorol. Soc.* **89**, 829–842 (2008).
- Avila, L. A. & Cangialosi, J. *Tropical Cyclone Report: Hurricane Irene (al092011)* (National Hurricane Center, 2011).
- Grinsted, A., Moore, J. C. & Jevrejeva, S. Homogeneous record of Atlantic hurricane surge threat since 1923. *Proc. Natl Acad. Sci. USA* **109**, 19601–19605 (2012).
- Blake, E.S. et al. *Tropical Cyclone Report: Hurricane Sandy (al182012)* (National Hurricane Center, 2013).
- Latto, A., Hagen, A. & Berg, R. *National Hurricane Center Tropical Cyclone Report: Hurricane Ida (al092020)* (National Hurricane Center, 2021).
- Beven II, J. L. & Berg, R. *National Hurricane Center Tropical Cyclone Report: Tropical Storm Fay (al062020)* (National Hurricane Center, 2021).
- Beven II, J. L., Hagen, A. & Berg, R. *National Hurricane Center Tropical Cyclone Report: Hurricane Ida (al092021)* (National Hurricane Center, 2022).
- Catalano, A. J. & Broccoli, A. J. Synoptic characteristics of surge-producing extratropical cyclones along the northeast coast of the United States. *J. Appl. Meteorol. Climatol.* **57**, 171–184 (2018).
- Grinsted, A., Moore, J. C. & Jevrejeva, S. Projected Atlantic hurricane surge threat from rising temperatures. *Proc. Natl Acad. Sci. USA* **110**, 5369–5373 (2013).
- Knutson, T. et al. Tropical cyclones and climate change assessment: Part II: Projected response to anthropogenic warming. *Bull. Am. Meteorol. Soc.* **101**, 303–322 (2020).

11. Roberts, K. J., Colle, B. A. & Korfe, N. Impact of simulated twenty-first century changes in extratropical cyclones on coastal flooding at the battery, New York City. *J. Appl. Meteorol. Climatol.* **56**, 415–432 (2017).
12. Lin, N., Marsooli, R. & Colle, B. A. Storm surge return levels induced by mid-to-late-twenty-first-century extratropical cyclones in the northeastern United States. *Nat. Clim. Chang.* **15**, 143–158 (2019).
13. Lin, N., Emanuel, K., Oppenheimer, M. & Vanmarcke, E. Physically based assessment of hurricane surge threat under climate change. *Clim. Change* **2**, 462–467 (2012).
14. Lin, N., Kopp, R. E., Horton, B. P. & Donnelly, J. P. Hurricane Sandy's flood frequency increasing from year 1800 to 2100. *Proc. Natl Acad. Sci. USA* **113**, 12071–12075 (2016).
15. Garner, A. J. et al. Impact of climate change on new york city's coastal flood hazard: increasing flood heights from the preindustrial to 2300 CE. *Proc. Natl Acad. Sci. USA* **114**, 11861–11866 (2017).
16. Marsooli, R., Lin, N., Emanuel, K. & Feng, K. Climate change exacerbates hurricane flood hazards along us Atlantic and Gulf coasts in spatially varying patterns. *Nat. Commun.* **10**, 1–9 (2019).
17. Marsooli, R. & Lin, N. Impacts of climate change on hurricane flood hazards in Jamaica bay, New York. *Clim. Change* **163**, 2153–2171 (2020).
18. Funakoshi, Y., Hagen, S. C. & Bacopoulos, P. Coupling of hydrodynamic and wave models: case study for Hurricane Floyd (1999) hindcast. *J. Waterw. Port Coast. Ocean Eng.* **134**, 321–335 (2008).
19. Sheng, Y. P., Alymov, V. & Paramygin, V. A. Simulation of storm surge, wave, currents, and inundation in the Outer Banks and Chesapeake Bay during Hurricane Isabel in 2003: the importance of waves. *J. Geophys. Res. Oceans* **115**, C4 (2010).
20. Dietrich, J. et al. Modeling hurricane waves and storm surge using integrally-coupled, scalable computations. *Coast. Eng.* **58**, 45–65 (2011).
21. Wallingford, H. et al. *The Joint Probability of Waves and Water Levels: Join-sea, a Rigorous But Practical New Approach* Vol. 1. HR Wallingford Report (HR Wallingford, 2000).
22. Ayyad, M., Hajj, M. R. & Marsooli, R. Artificial intelligence for hurricane storm surge hazard assessment. *Ocean Eng.* **245**, 110435 (2022).
23. Lee, T. L. Neural network prediction of a storm surge. *Ocean Eng.* **33**, 483–494 (2006).
24. Tseng, C. M., Jan, C. D., Wang, J. S. & Wang, C. Application of artificial neural networks in typhoon surge forecasting. *Ocean Eng.* **34**, 1757–1768 (2007).
25. Lee, T. L. Back-propagation neural network for the prediction of the short-term storm surge in Taichung harbor, Taiwan. *Eng. Appl. Artif. Intell.* **21**, 63–72 (2008).
26. De Oliveira, M. M., 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).
27. 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).
28. 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).
29. 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).
30. Tiggeloven, T., Couasnon, A., van Straaten, C., Muis, S. & Ward, P. J. Exploring deep learning capabilities for surge predictions in coastal areas. *Sci. Rep.* **11**, 1–15 (2021).
31. 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**, 24–38 (2013).
32. 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).
33. Kyriotti, A. P., Taflanidis, A. A., Nadal-Caraballo, N. C. & Campbell, M. Storm hazard analysis over extended geospatial grids utilizing surrogate models. *Coast. Eng.* **168**, 103855 (2021).
34. Ayyad, M., Hajj, M. R. & Marsooli, R. Machine learning-based assessment of storm surge in the New York metropolitan area. *Sci. Rep.* **12**, 1–12 (2022).
35. Ayyad, M., Orton, P. M., El Safty, H., Chen, Z. & Hajj, M. R. Ensemble forecast for storm tide and resurgence from tropical cyclone Isaias. *Weather Clim. Extrem.* **38**, 100504 (2022).
36. Brandon, C. M., Woodruff, J. D., Donnelly, J. P. & Sullivan, R. M. How unique was hurricane sandy? Sedimentary reconstructions of extreme flooding from New York harbor. *Sci. Rep.* **4**, 1–9 (2014).
37. Kalnay, E. et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–472 (1996).
38. Emanuel, K., Sundararajan, R. & Williams, J. Hurricanes and global warming: results from downscaling IPCC ar4 simulations. *Bull. Am. Meteorol. Soc.* **89**, 347–368 (2008).
39. Delworth, T. L. et al. Gfdl's cm2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Clim.* **19**, 643–674 (2006).
40. Gnanadesikan, A. et al. Gfdl's cm2 global coupled climate models. Part II: The baseline ocean simulation. *J. Clim.* **19**, 675–697 (2006).
41. Bellouin, N. et al. The hadgem2 family of met office unified model climate configurations. *Geosci. Model Dev.* **4**, 723–757 (2011).
42. Giorgetta, M. A. et al. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the coupled model intercomparison project phase 5. *J. Adv. Model. Earth Syst.* **5**, 572–597 (2013).
43. Yukimoto, S. et al. *Meteorological Research Institute-Earth System Model Version 1 (MRI-ESM1)—Model Description* (Meteorological Research Institute, 2011).
44. Gurumurthy, P., Orton, P. M., Talke, S. A., Georgas, N. & Booth, J. F. Mechanics and historical evolution of sea level blowouts in New York harbor. *J. Mar. Sci. Eng.* **7**, 160 (2019).
45. Rego, J. L. & Li, C. Nonlinear terms in storm surge predictions: effect of tide and shelf geometry with case study from Hurricane Rita. *J. Geophys. Res. Oceans* **115**, C6 (2010).
46. Poullose, J., Rao, A. & Bhaskaran, P. K. Role of continental shelf on non-linear interaction of storm surges, tides and wind waves: an idealized study representing the west coast of India. *Estuar. Coast. Shelf Sci.* **207**, 457–470 (2018).
47. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
48. Willmott, C. J. On the validation of models. *Phys. Geogr.* **2**, 184–194 (1981).
49. Marsooli, R. & Lin, N. Numerical modeling of historical storm tides and waves and their interactions along the us east and Gulf coasts. *J. Geophys. Res. Oceans* **123**, 3844–3874 (2018).

## AUTHOR CONTRIBUTIONS

M.A. conceived the study, conducted the analysis, curated the data, created the ML model, and wrote the first draft; M.R.H. led the research, initiated the study; R.M. collected the data, designed the statistical methodology; M.R.H. and R.M. reviewed and edited the paper; M.A., M.R.H., and R.M. interpreted and discussed the findings.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Mahmoud Ayyad.

**Reprints and permission 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 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023