

# Earth's Future

## RESEARCH ARTICLE

10.1029/2021EF002285

### Key Points:

- A surrogate model was coupled with a stochastic climate model to forecast hypothetical future coastal impacts
- Future changes of characteristics driving flooding and dune impact events are projected contingent on sea level rise
- Relevant time scales of climate variability are incorporated into resilience metrics at Naval Base Coronado, San Diego

### Supporting Information:

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

### Correspondence to:

D. L. Anderson,  
danders5@ncsu.edu

### Citation:

Anderson, D. L., Ruggiero, P., Mendez, F. J., Barnard, P. L., Erikson, L. H., O'Neill, A. C., et al. (2021). Projecting climate dependent coastal flood risk with a hybrid statistical dynamical model. *Earth's Future*, 9, e2021EF002285. <https://doi.org/10.1029/2021EF002285>

Received 1 JUL 2021

Accepted 13 NOV 2021

## Projecting Climate Dependent Coastal Flood Risk With a Hybrid Statistical Dynamical Model

**D. L. Anderson<sup>1</sup> , P. Ruggiero<sup>2</sup> , F. J. Mendez<sup>3</sup> , P. L. Barnard<sup>4</sup> , L. H. Erikson<sup>4</sup> , A. C. O'Neill<sup>4</sup>, M. Merrifield<sup>5</sup> , A. Rueda<sup>3</sup>, L. Cagigal<sup>3,6</sup> , and J. Marra<sup>7</sup>**

<sup>1</sup>College of Engineering, North Carolina State University, Raleigh, NC, USA, <sup>2</sup>College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA, <sup>3</sup>Dpto Ciencias y Tecnicas del Agua y del Medio Ambiente, Universidad de Cantabria, Santander, Spain, <sup>4</sup>Pacific Coastal and Marine Science Center, United States Geological Survey, Santa Cruz, CA, USA, <sup>5</sup>Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA, <sup>6</sup>School of Environment, Faculty of Science, University of Auckland, Auckland, New Zealand, <sup>7</sup>National Oceanic and Atmospheric Administration, Honolulu, HI, USA

**Abstract** Numerical models for tides, storm surge, and wave runup have demonstrated ability to accurately define spatially varying flood surfaces. However these models are typically too computationally expensive to dynamically simulate the full parameter space of future oceanographic, atmospheric, and hydrologic conditions that will constructively compound in the nearshore to cause both extreme event and nuisance flooding during the 21st century. A surrogate modeling framework of waves, winds, and tides is developed in this study to efficiently predict spatially varying nearshore and estuarine water levels contingent on any combination of offshore forcing conditions. The surrogate models are coupled with a time-dependent stochastic climate emulator that provides efficient downscaling for hypothetical iterations of offshore conditions. Together, the hybrid statistical-dynamical framework can assess present day and future coastal flood risk, including the chronological characteristics of individual flood and wave-induced dune overtopping events and their changes into the future. The framework is demonstrated at Naval Base Coronado in San Diego, CA, utilizing the regional Coastal Storm Modeling System (CoSMoS; composed of Delft3D and XBeach) as the dynamic simulator and Gaussian process regression as the surrogate modeling tool. Validation of the framework uses both in-situ tide gauge observations within San Diego Bay, and a nearshore cross-shore array deployment of pressure sensors in the open beach surf zone. The framework reveals the relative influence of large-scale climate variability on future coastal flood resilience metrics relevant to the management of an open coast artificial berm, as well as the stochastic nature of future total water levels.

**Plain Language Summary** Coastal flooding is caused by the combination of winds, waves, tides, and regional changes in water levels due to large-scale climate patterns such as the El Niño Southern Oscillation. Although high fidelity numerical models are our best representation of nearshore and estuarine physical processes, the random nature of climate, weather/storms, and their timing interfering with tides results in too many hypothetical futures to simulate with computationally expensive models. Machine learning techniques are used in this study to efficiently derive coastal water levels throughout the 21st century based on a subset of numerical model simulations. The future simulations provide hourly information allowing for coastal managers to explore resilience metrics that consider the event characteristics (i.e., durations and environmental forcing) of both extreme flood events resulting from large storms and sunny-day flooding due to sea level rise.

## 1. Introduction

The need for robust coastal hazard predictions has grown considerably due to coastal population growth and recent disasters resulting from extreme storm events (Neumann et al., 2015). Observed global sea level rise (SLR) (Nerem et al., 2018) and the consensus expectation of accelerated SLR in the future (e.g., Sweet et al., 2017) further motivate research efforts to quantify spatial and temporal exposure to future coastal flood and erosion hazards. Numerous efforts have focused on the development and application of high-fidelity, but computationally expensive, numerical modeling suites to quantify the impacts of specific storm events (e.g., Barnard et al., 2019; Bilskie et al., 2014; Dietrich et al., 2011; Hsu et al., 2018; Warner et al., 2010; Wolf, 2009) or use output from a limited number of dynamically downscaled multi-decadal general circulation models (GCMs) (e.g., Muis et al., 2020). Other work has investigated statistical approaches to generate 1,000s of multivariate combinations

of waves, sea levels, precipitation, and river flows that could compound to create future extreme events (e.g., Callaghan et al., 2008; Moftakhi et al., 2017; Serafin & Ruggiero, 2014; Wahl & Chambers, 2015), enabling an estimate of the intrinsic variability within climate-dependent processes. Although such statistical approaches can be computationally efficient, they do not account for the nonlinear feedbacks between processes that a numerical model resolves, which can have significant effects on the future water levels and ultimate hazard prediction. Recent efforts have highlighted the benefits of incorporating both lines of research into hybrid dynamical-statistical frameworks that more accurately capture the magnitude of the compound hazard while still providing enough realizations to quantify natural variability (e.g., Moftakhi et al., 2019; Serafin, Ruggiero, Parker, et al., 2019). This study uses surrogate modeling, a form of machine learning (e.g., Goldstein et al., 2019), to develop such a hybrid methodology. The hybrid framework leverages statistically derived hypothetical coastal climates, projected SLR curves, and numerically modeled flood and wave simulations to predict many iterations of possible future water levels.

The expectation of further SLR has led to a large body of literature projecting increasing frequencies of extreme still water levels (ESLs) and the resulting flooding (e.g., Sweet & Park, 2014; Wahl et al., 2017). Coastal flooding is a broad term used to describe flooding that occurs at the land-sea boundary. However, this integrated hazard is a consequence of many constructively and destructively interfering oceanographic, atmospheric, and hydrologic drivers. Physical processes contributing to future flooding include glacial melt processes (slowly occurring over centuries) (e.g., Fox-Kemper et al., 2021), regional sea level anomalies due to large-scale climate oscillations (interannual to inter-decadal variability) (e.g., Barnard et al., 2015, 2017; Merrifield et al., 2012; Wahl & Chambers, 2016), storm events driving surge as well as large wave runup (weather-band timescale) (e.g., Marcos et al., 2019; Vousdoukas et al., 2018), and tidal forcings (hourly to interannual fluctuations) (e.g., Ray & Foster, 2016). Future flooding frequencies, tipping points, and hazard exposures are often quantified with data-driven extreme value analyses relying on the historical still water level provided by tide gauge records (e.g., Moftakhi et al., 2017; Sweet & Park, 2014). Projections derived from such methods fundamentally assume they have observed the full range of variability in underlying forcing. However, historical records are relatively short and have only observed a limited number of meteorologic and climatic extremes that produce compound flood events. Consequently, methods that assume observed distributions of ESLs are a stationary process superimposed on SLR curves can be insufficient for projecting compound extremes (R. E. Kopp et al., 2019).

Limitations associated with relatively short observation records can be addressed by statistically generating hypothetical time series of each individual component contributing to the compounding nearshore hazard (e.g., Serafin & Ruggiero, 2014), or by dynamically downscaling hypothetical futures from GCMs (e.g., Muis et al., 2020). Recently developed statistical frameworks have demonstrated the ability to form explicit links between the climate drivers and contributing water level processes (D. Anderson et al., 2019) while deriving chronological behavior necessary for coupling with time-dependent processes such as geomorphic models (Cagigal et al., 2020). However, limitations are introduced by initially separating compounding processes and assuming that a cumulative hazard can be created by linear superposition during simulation. This assumption is particularly problematic on the open coast where empirical estimates are used for wave runup elevations that simplify beach topography and neglect dune configurations (Gallien, 2016; Stockdon et al., 2014). Such statistical projections are also typically limited to a single point in space (i.e., a tide gauge) and thus cannot provide the alongshore varying hazard predictions possible with numerical models. Decision frameworks for open coast stakeholders are thus in need of new approaches that consider both statistical methods quantifying the stochastic nature of compounding processes and dynamical methods for deriving alongshore varying environmental conditions interacting with varying beach topographies.

Surrogate models, also referred to as metamodels, are predictive models (e.g., neural networks, logistic regressions, decision trees, etc.; Goldstein et al., 2019) that learn the response of a subset of conditions run through a numerical simulator to then predict the response during other conditions with greater computational efficiency relative to the original simulator. In coastal applications, surrogate models have been applied in multiple locations for hurricane storm surge predictions (Jia & Taflanidis, 2013; Jia et al., 2016) and have shown promise for storm-induced geomorphic response of dunes (Santos et al., 2019). Until recently these efforts have been limited in application to emulating storm events (i.e., maximum surge during a storm) with the surrogate model predicting a non-tidal residual due to atmospheric variables (i.e., minimum pressure, storm speed, and storm angle of approach) (Resio et al., 2009). Parker et al. (2019) demonstrated that with the proper tailoring of an

ADCIRC + SWAN simulator, a surrogate model could use offshore wave, water level, and atmospheric observations to efficiently hindcast decades of continuous hourly water levels and wave heights within a large estuary. The advancement from individual storm event emulation to continuous time emulation enables a wide range of questions to be addressed with respect to time-dependent flooding on management time scales, including recurrence intervals, nuisance flooding thresholds, and individual event characteristics such as the relative contributions of specific forcing parameters and their chronological behavior within the event.

In this work, we extend the methodology of Parker et al. (2019) by numerically resolving surf zone dynamics and wave-induced flooding resulting from wave runup and dune overtopping. We also tailor the surrogate model to be forced by the same dimensionality as a stochastic climate model that downscals large-scale climate variability to time series of hypothetical local waves, water levels, and atmospheric conditions (D. Anderson et al., 2019). In this sense, rather than a hindcast making use of observed boundary conditions, we apply the dynamic surrogate model to hypothetical realizations of coastal ocean boundary conditions in order to quantify potential future water levels. The framework is computationally efficient, allowing for multiple SLR scenarios, hundreds of future iterations of large-scale El Niño Southern Oscillations, and thousands of storm events to be simulated while interacting with different tide cycles and different swell signatures originating from distinct wave generation regions.

This study demonstrates a complete framework that includes running several hundred numerical simulations, creating and validating multiple kinds of surrogate model products, and applying stochastic climate time series to project potential future flood risk (Figure 1). Section 2 introduces the study site, San Diego, CA, including both the previously developed numerical modeling framework and the climate model output used to generate southern California coastal conditions. Section 3 reviews the creation of surrogate models by Gaussian Process Regression (GPR) and the design of a simplified numerical simulator to emulate local environmental conditions. Section 4 validates the surrogate models for water levels and wave-induced nearshore dynamics using observations from in-situ deployments. The discussion in Section 5 synthesizes how this novel framework projects the frequency and severity of open coast hazards such that coastal managers can prepare for future conditions.

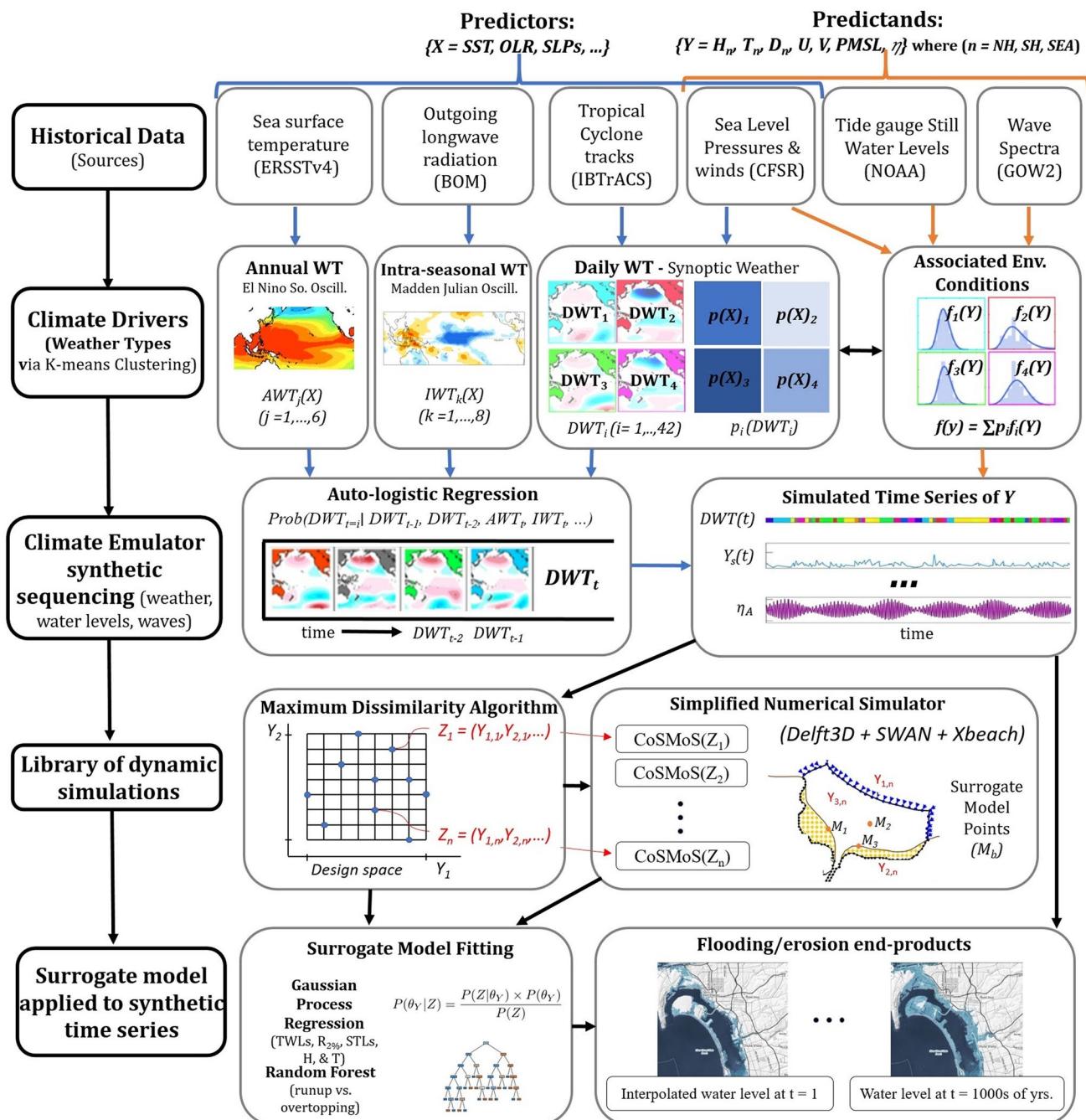
## 2. Study Site, Numerical Model Setup, and Observations

### 2.1. Naval Base Coronado, San Diego, CA

In this study, multiple surrogate models are developed for flooding hazards at Naval installations in San Diego, CA, where erosion has removed the natural dune system and the Navy maintains a constructed beach berm to protect the base from wave-induced flooding. The study area includes the open beach front extending from Coronado Beach south to the Naval Amphibious Base Coronado, as well as back-bay shorelines around the Naval Air Station and into San Diego Bay (Figure 2). The open ocean beaches are composed of median grain size of 0.20 mm and an average beach slope of 0.03 with local variability (Ludka et al., 2015). Annual average wave heights are 2.25 m in deep water but vary seasonally with larger waves in the winter and smaller in the summer (O'Reilly et al., 2016). Nearshore waves often exhibit signatures of multi-modal spectra due to the predominant influence of long period swells from both the southern and northern hemispheres and resulting refraction patterns caused by offshore islands (Hegermiller et al., 2017; Pawka, 1983). However, locally generated wind waves can contribute as much as 40% to the total energy spectrum and the orientation of Coronado spit relative to Point Loma also results in alongshore varying wave exposure due to headland shadowing (Crosby et al., 2016).

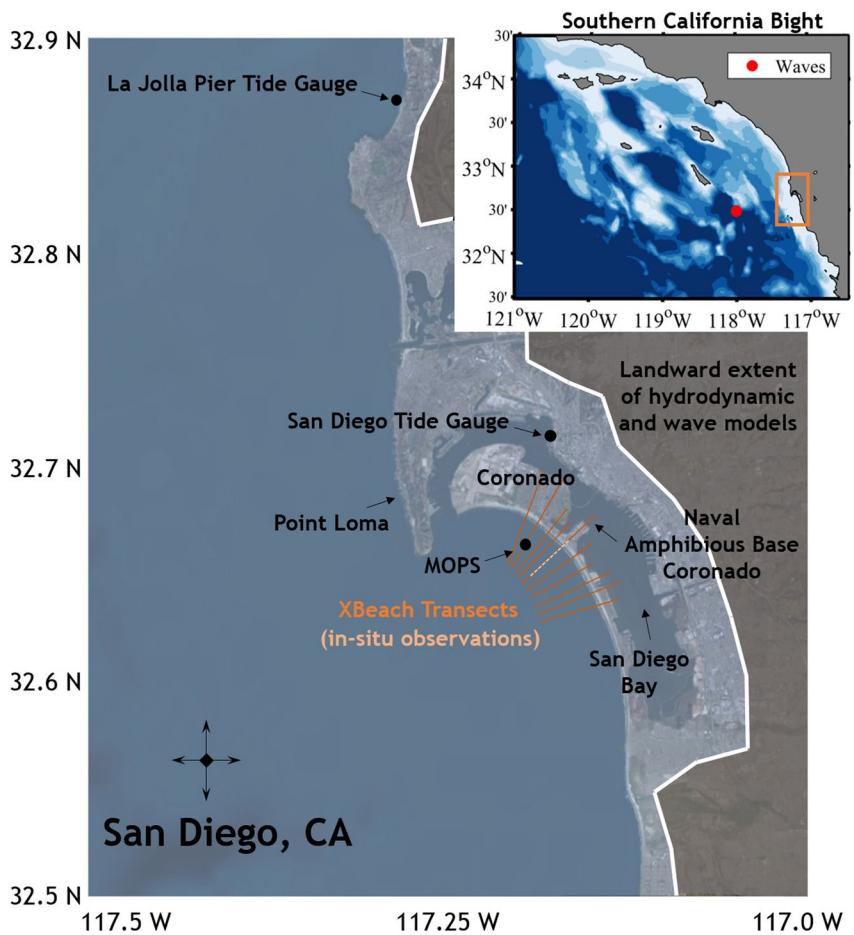
### 2.2. CoSMoS Numerical Model

The Coastal Storm Modeling System (CoSMoS) was developed by the United States Geological Survey (USGS) in collaboration with Deltares for the purpose of projecting 21st century flood hazard maps in southern California (L. Erikson et al., 2018; O'Neill et al., 2018). Regional and locally nested grids of dynamically integrated SWAN (Simulating Waves Nearshore, Delft University of Technology), WaveWatch3, and Delft3D-FLOW models downscale tide, surge, and wave dynamics from deep water to the shoreline. High resolution nearshore grids are used to resolve harbors, bays, estuaries, and overland flow (finest resolution 5 m by 15 m). Cross-shore XBeach (hydrostatic) model (eXtreme Beach: Roelvink et al., 2009) transects at ~100 m alongshore resolution are used to simulate long-wave dominated surf zone processes and thus wave runup and overtopping-induced flooding not produced by Delft3D (Roelvink et al., 2009). XBeach transects begin in ~15 m water depths where wave spectra and water levels are passed in a loosely coupled fashion from Delft3D and SWAN (Figure 2). The original model



**Figure 1.** Conceptual flow diagram of the inputs and methods incorporated into the hybrid framework developed in this study. The top three rows comprise the generation of hypothetical time series of environmental conditions for San Diego, CA. The bottom two rows highlight the creation of a library of CoSMoS simulations optimally chosen from the hypothetical time series to generate surrogate models that can predict water levels at all times.

development integrated GCM output of select storm events intended to represent both rare (the 100-year event) and common (the annual event) return period meteorological storms simulated on multiple SLR scenarios (L. H. Erikson et al., 2018). See Barnard et al. (2014, 2016) and O'Neill et al. (2018) for further details regarding digital terrain model sources, validation, and CoSMoS's numerous products generated throughout California. This study utilizes a subset of the full CoSMoS model, specifically running the Delft3D, SWAN, and XBeach components relevant for San Diego. The offshore edge of the model domain is in deep-water (>1,200 m), with the onshore edge located in elevated backshore terrains not likely to ever experience flooding. Focusing on San



**Figure 2.** Study map of San Diego, CA, including notable landmarks for reference within the text. Tide gauges and wave record locations are denoted by black dots, XBeach transects in orange (light orange for the transect with in-situ validation observations), and the landward extent of the hydrodynamic and wave numerical models (blue patch).

Diego sub-domains drastically reduced the computational power needed for a single simulation, allowing this study to increase the number of event simulations by an order of magnitude compared to previous CoSMoS efforts (from O[10s] to O[100s]). Boundary conditions provided to the model domain are discussed in Section 3.3.

### 2.3. Stochastic Climate Emulator

The time-varying emulator for short and long-term analysis of coastal flooding (TESLA-flood), a stochastic climate emulator introduced by D. Anderson et al. (2019), was applied to derive boundary conditions for CoSMoS. All necessary python codes and documentation can be found at the GitHub page <https://github.com/anderdy/teslaCoSMoS> (<https://doi.org/10.5281/zenodo.5055555>). TESLA uses observations of basic climate drivers such as sea surface temperature (SST), outgoing longwave radiation (OLR), and sea level pressure (SLP) as predictors,  $X$ , for local-scale predictands,  $Y$ , including oceanographic and atmospheric forcings (Figure 1). The framework is based on the fundamental assumptions that synoptic weather is both:

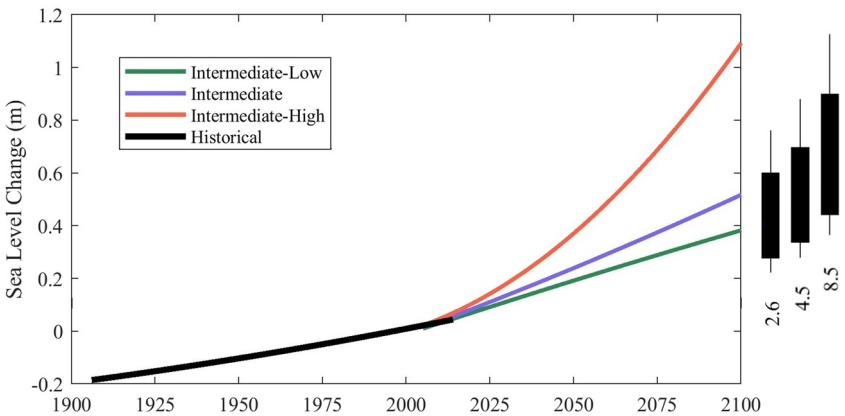
1. A driver of local environmental conditions (e.g., the weather on any given day drives the winds generating ocean waves and the atmospheric pressures producing regional water level anomalies)
2. A localized expression of large-scale climate such that the probability of observing any particular kind of weather system is modulated by processes slowly varying on longer timescales (e.g., storm intensities and tracks are inherently related to ocean temperatures, upper atmosphere wind shear, and seasonality in solar energy input)

Chronological climate behavior is simulated by turning the predictors into categorical proxies representative of large-scale climate oscillations (e.g., El Niño Southern Oscillation, Madden Julian Oscillation, and seasonality) via weather typing. Weather Typing consists of principal component analysis of  $X(lon, lat, time)$  patterns and K-means clustering of resulting empirical orthogonal function's temporal amplitudes (Camus et al., 2014). This is applied on an annual scale to SST patterns in the equatorial Pacific Ocean to identify years with similar ocean temperature structure, defined as six different Annual Weather Types (AWT). Intra-seasonal atmospheric variability affecting daily weather is similarly derived from the last 40 years of observed equatorial OLR to create eight Intra-seasonal Weather Types (IWT). Forward propagation of synthetic AWT and IWT spatial patterns by Markov chains creates temporal variability consistent with the El Niño Southern Oscillation and the Madden Julian Oscillation. However, the synthetic climates contain timing, magnitude, and phase relationships between the two dominant climate oscillations not seen in the relatively short historical record (see D. Anderson et al. (2019) for extensive details regarding climate time series construction).

Representative daily synoptic weather patterns for the entire extent of the Pacific Ocean are derived from the 1979 to 2019 SLP record provided by the Climate Forecast System Reanalysis (CFSR) from the National Center for Environmental Prediction (NCEP). Historical tropical cyclones in the eastern Pacific Ocean were identified in the International Best Track Archive (IBTrACS) and classified into categorical variables depending on Saffir-Simpson scale. All remaining SLP patterns were converted to categorical daily weather types (*DWT*) via weather typing to create 42 total synoptic weather patterns. Chronological prediction of synthetic *DWT*( $t$ ) time series is performed by an auto-regressive logistic (ALR) model contingent on synthetic time series of AWTs, IWTs, and seasonality (Guanche et al., 2014). Each *DWT* is associated with unique probability distributions of  $Y$  consisting of generalized extreme value distributions fit to historical observations  $f_i(Y)$ . The complete dimensionality of  $Y$  includes the East-West and North-South components of wind ( $U, V$ ), atmospheric pressure at mean sea level (*PMSL*), monthly mean sea level anomalies ( $\eta_{MMSLA}$ ), and offshore wave heights, periods and directions for swell partitions from the northern hemisphere (*NH*), southern hemisphere (*SH*), and locally generated winds (*SEA*). The joint probabilities of  $Y$  are defined by copulas specific to each  $DWT_i$ . Copulas identify correlation after transforming each marginal distribution to a uniform distribution (Rasmussen & Williams, 2006). A wide range of copula families exist, each providing different advantages for capturing dependence structures and tail relationships. Gaussian copulas were shown to ensure realistic co-occurrence for wave and surge conditions in San Diego in D. Anderson et al. (2019) and previous works (e.g., Rueda et al., 2016). Gaussian copulas are ideal for high dimensional problems; however they are elliptically symmetric about the diagonal and asymptotically independent and thus not perfect representations of the correlations in  $Y$ .

During simulations, hypothetical *DWT* chronologies are filled with plausible environmental parameters by randomly sampling from the appropriate *DWT*'s copula and interpolating to hourly conditions consistent with historically observed storm hydrographs (wave ramp-up and ramp-down during a storm). The complete output from TESLA includes those variables in the copulas plus tides and monthly mean sea levels due to SST changes associated with ENSO (Figure 4). Specifically, the tidal amplitudes are composed of the leading eight constituents and a modifying 18.6 year period oscillation forward-predicted for the 21st century with the T-Tides toolbox (Pawlowicz et al., 2002), and thus neglecting potential changes in tidal amplitudes and propagation characteristics resulting from SLR (Ray & Foster, 2016). The monthly mean sea level contains both seasonal cycles and variability induced by ENSO SST anomalies. The efficient statistical modeling approach can directly simulate compound events not observed in the historical record, better constraining the magnitude of extremes that can occur at a location which experiences multivariate flooding. TESLA also chronologically orders those conditions on an hourly time-step such that return periods of flood magnitudes can be quantified for the present-day and for future climates considering SLR and tidal variability. The stochastic nature of the approach results in many unique realizations of coastal conditions that can occur within the same large-scale climate, which inherently quantifies the variability and thus uncertainty in extremes.

TESLA was intentionally developed to use SLPs and SSTs as fundamental drivers because they are both observable in the real world as well as output products simulated by GCMs. GCMs are constantly improving with respect to the inclusion of more process-based physics and their ability to hindcast historical climates, suggesting frameworks that make direct use of their output will continue to be relevant throughout the 21st century. However, the weather and climate pattern projections presently produced by GCMs contain considerable variability between models (e.g., disagreements concerning the affect that different model configurations have on the magnitude and



**Figure 3.** Tide gauge-derived historical sea level rise (black) and three projected sea level rise curves from Sweet et al. (2017) for San Diego, CA. Projected SLR in 2100 derived from the IPCC sixth Assessment are provided to the right hand side for SSPs-2.6, -4.5, and -8.5 with boxes signifying the 17th to 83rd percentiles and whiskers representing 5th to 95th percentiles (Fox-Kemper et al., 2021).

frequency of future ENSO cycles; Marjani et al., 2019; Stevenson et al., 2021). Rather than folding this high degree of uncertainty into this study, results hereafter contain climate variability statistically consistent with historically observed oscillations. Larger magnitude El Niños and more extreme storm conditions than historically observed are still generated by sampling from GEVs fit to historical conditions. To this end, the potential changes in storminess patterns associated with varying degrees of global warming are not included in this study, but the approach has been developed such that SLP outputs from GCMs can be incorporated into future work. The effect of different warming scenarios are instead indirectly accounted for through the inclusion of several regional SLR scenarios generated by an inter-agency task force to inform coastal risk management tools at regional scales for the entire U.S. coastline (Sweet et al., 2017). Sweet et al. (2017) account for processes driving vertical land motion (subsidence, sediment compaction, and groundwater extraction) in the San Diego region to create five SLR scenarios with varying probabilities of occurrence for each Shared Socio-Economic Pathway (SSP). Three scenarios are considered in this work, referred to in Figure 3 by their scenario name in the original report. Medium-confidence sea level projections in 2100 derived from the 2021 IPCC Sixth Assessment are provided for reference (Fox-Kemper et al., 2021; Garner et al., 2021). Coastal forcing used in this study thus accounts for both climate change induced SLR and chronological synoptic weather conditions that preserve time-dependent intra-seasonal to interannual climate variability.

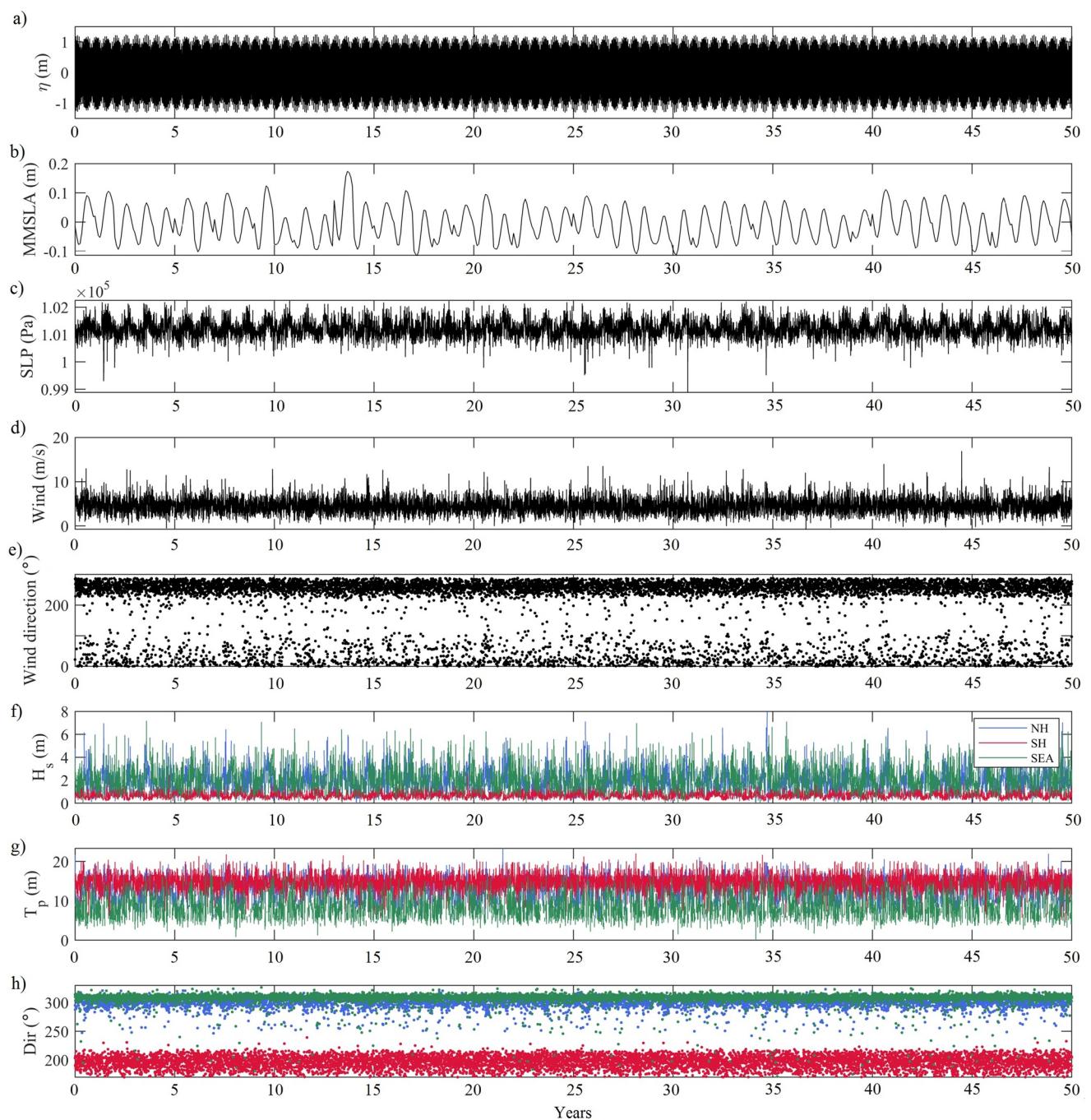
#### 2.4. Observational Data

This study requires multiple data sets for the purposes of developing CoSMoS boundary conditions and for local-scale model validation via instrument deployments. Input to the climate emulator described in Section 2.3 includes both ocean and atmospheric reanalysis products. Wave forcing from 1979 to 2015 was derived from the GOW2 historical reanalysis of Perez et al. (2017) and was separated into one local wind sea (SEA) and two swell partitions generated in the Northern (NH) and Southern (SH) Hemispheres using the approach developed by (Rueda et al., 2017):

$$S(f, \phi) = S_{SEA}(f, \phi) + S_{NH}(f, \phi) + S_{SH}(f, \phi) \quad (1)$$

where  $S_{NH}(f, \phi)$  and  $S_{SH}(f, \phi)$  are calculated by aggregating all wave trains between  $240^\circ < D < 360^\circ$  and  $140^\circ < D < 240^\circ$ , respectively where  $D$  is the mean wave direction. Regional winds and atmospheric sea level pressures from 1979 to 2018 were provided by CFSR. Water level observations for both the outer coast and within San Diego Bay were provided by two tide gauges operated by the National Ocean and Atmospheric Administration (NOAA) (station 9410230 located on a pier in La Jolla, CA, and station 941017 located in the northeast corner of San Diego Bay). The still water level, SWL, recorded at the gauges was separated into its constitutive components following (Serafin & Ruggiero, 2014):

$$SWL = \eta_{MSL} + \eta_A + \eta_{SE} + \eta_{MMSLA} + \eta_{SS} \quad (2)$$



**Figure 4.** Example time series of CoSMoS boundary conditions provided by TESLA. (a) astronomical tides, (b) regional monthly sea level anomalies, (c) atmospheric pressure at sea level, (d) nearshore winds magnitudes, (e) wind directions, (f) wave heights, (g) wave periods, and (h) wave directions. Wave parameters in (f), (g), and (h) are presented as components from the Northern Hemisphere (NH), Southern Hemisphere (SH), and local wind sea (SEA).

where  $\eta_{MSL}$  is the mean sea level relative to some reference datum,  $\eta_A$  is the astronomical tide,  $\eta_{SE}$  is an intra-annual seasonal water level variation,  $\eta_{MMSLA}$  is the monthly mean sea level variability, and  $\eta_{SS}$  is the storm surge signal resulting from atmospheric pressure variability and wind setup.

Validation of the XBeach surrogate models in the nearshore was performed with both a locally downscaled wave hindcast (MONitoring and Prediction System (MOPS) (O'Reilly et al., 2016)) and in-situ measurements from six pressure sensors deployed in the surf zone fronting the NABC between January and April 2018. MOPS is

specifically tailored to account for the local sea states and swells interfering in the southern California bight by assimilating and transforming a network of coastal buoys to provide wave information in 10 meters of water since 2000 (data from the Coastal Data Information Program (CDIP) at Scripps Institution of Oceanography). Sensor locations during the nearshore deployment included one instrument approximately 350 m offshore at the edge of the surf zone and -8.0 m relative to mean sea level (MSL). Five more sensors spanned the inner surf zone and swash zone in 10 m cross-shore horizontal increments at vertical heights of -1.12, -0.55, -0.45, 0.15, and 0.75 m, respectively. Pressure recordings were made with buried paros sensors sampling at 2 Hz.

### 3. Surrogate Modeling Techniques

Surrogate modeling was introduced in the 1970s (Blanning, 1975) and is now widely used across the physical sciences, medical industries, economics, and for automation of machinery or computer decisions (Razavi et al., 2012). The basic concept is to choose a set of representative scenarios as inputs to a simulator, and subsequently fit a statistical function to the observed relations between the simulator's inputs and outputs. With respect to coastal applications, the simulator is often taken as the inputs and outputs from a dynamic numerical model, with many different statistical methods used as the function mimicking dynamical behavior. Nearshore wave predictions from ensembles of SWAN simulations have been emulated with look-up tables (LUT) and linear interpolation schemes (Allan et al., 2015), radial basis functions (Camus et al., 2011; Gouldby et al., 2014), Gaussian process regressions (Pullen et al., 2018), and artificial neural networks (James et al., 2018). Similarly, storm surge predictions have been emulated with surge response functions (SRFs) (Irish et al., 2011), Gaussian process regression (Jia et al., 2016; Parker et al., 2019), artificial neural networks (Kim et al., 2015), and random forests (Tadesse et al., 2020). More recently, hybrid frameworks considering multiple loosely-coupled numerical models have used surrogate models to investigate the compound flooding resulting from tropical cyclone induced surge coupled with rainfall runoff (Bass & Bedient, 2018) and estuary water levels coupled with river discharges (Serafin, Ruggerio, Parker, et al., 2019). This study uses Gaussian process regression (GPR), also known as kriging, to emulate water levels from Delft3D as well as loosely coupled XBeach simulations of the surf zone (explained further in Section 3.2).

#### 3.1. Selection of Training Cases

All surrogate models are effectively interpolation/extrapolation schemes. Based on model results from a subset of input scenarios, they infer likely outcomes for new input scenarios. Many space filling strategies exist for choosing the optimal subset of conditions to improve interpolation (i.e., stratified sampling, cluster sampling, equal probability sampling, etc.) (Camus et al., 2011). Ideally, the selected training cases (commonly referred to as design points) span the entire range of variability in potential inputs, allowing for more constrained interpolation as opposed to extrapolation. This study has the advantage of a priori knowing 1,000s of different combinations of offshore conditions produced by the climate emulator (TESLA), which provides the full variability of conditions we would like to provide to the surrogate models. A Maximum Dissimilarity Algorithm (MDA) was applied to 100,000 years of simulated conditions created by simulating a hundred years of hourly conditions 250 times each for the three different SLR projections described above (Sweet et al., 2017) as well as 250 simulations with no sea level rise applied. The MDA algorithm identified 600 design points, defined in this framework as  $Z_n$  where  $n = (1, \dots, 600)$ . MDA uses an iterative approach to identify the next point in the data set that is most dissimilar (calculating Euclidean distances in multi-dimensional space) from the already selected points (Camus et al., 2011). It tends to initially identify outliers and progressively adds design points located within the center of the multi-dimensional space. The sampling technique orders the cases chosen by importance for filling the dimensional space, which allows for assessments of the output sensitivity to the size of the library of dynamical runs (addressed in Supporting Information S1) while also concentrating sampling points in multi-dimensional space where there is a higher density of intended simulation points.

#### 3.2. Gaussian Process Regression

Gaussian process regression is a supervised machine learning approach that is non-parametric and highly generalizable such that many other surrogate model formulations collapse to a GPR formulation under specific circumstances (i.e., neural networks and radial basis functions; Anjyo and Lewis (2011)). GPRs are commonly referred

to as a distribution over functions, which governs how every point in the input space is related (Rasmussen & Williams, 2006). The key assumption in GPRs is that all data provided for training can be represented as a sample of a multivariate Gaussian distribution defined by a mean,  $\mu$ , and covariance function,  $\sigma$ , usually referred to as a kernel. The infinite number of functions that could define the input space are constrained by the data through Bayesian inference,

$$P(\theta_Y | Z) = \frac{P(Z|\theta_Y) \times P(\theta_Y)}{P(Z)} \quad (3)$$

where  $P(\theta_Y | Z)$  is the posterior function,  $P(Z|\theta_Y)$  is the likelihood distribution,  $P(Z)$  is the prior, and  $\theta_Y$  encompasses the parameter values of a Gaussian ( $\mu$ ,  $\sigma$ ) for each input dimension  $Y$ . The final posterior is the mean of all functions which pass through the constraining points (and thus has an associated uncertainty related to the variability of the functions).

Mean functions can cover a wide range of complexities, from constant zeros to high order polynomials individually tailored to each input dimension. Kernels can similarly exhibit a wide range of complexities, with the most common being exponential, squared exponential, and Matern functions (Rasmussen & Williams, 2006). The Matern 5/2 kernel formulation produced the smallest errors in each surrogate model developed in this study,

$$\sigma_{5/2}(d) = \sigma^2 \left( 1 + \frac{d\sqrt{5}}{l} + \frac{5d^2}{3l^2} \right) \exp\left(-\frac{d\sqrt{5}}{l}\right) \quad (4)$$

where  $d$  is the Euclidean distance, and  $l$  is a length scale. The function is effectively a product of an exponential and a polynomial, and is twice differentiable. The Matern kernel was also previously demonstrated to perform well in coastal water level predictions with a similar number of dimensions (Parker et al., 2019). Specifics regarding optimization of hyperparameters, mean functions, and feature relevance are provided in Supporting Information S1.

### 3.3. Simulator Design

Surrogate models,  $M$ , map the full dimensionality of input parameters to desired output. If the training library includes a high degree of complexity and variability across each case's inputs, then more dimensions are necessary in the surrogate model. As more dimensions are added, the library of cases must become large enough to include the necessary range of variability across each input variable. Minimizing the number of dimensions and overall complexity in the simulator is the first step in creating a quality surrogate model. Common assumptions include uniform offshore wave spectra boundary forcing (Gouldby et al., 2014), uniform and stationary water levels neglecting tides (Jia et al., 2016), and steady state forcing to allow simulations to reach equilibrium with the forcing (Parker et al., 2019).

In this study we simplify the number of parameters that vary to exclusively be those controlling water levels in the model domain,  $M(Y)$ . All default model parameters (in Delft3D-FLOW, SWAN, and XBeach) are used (i.e., surface drag and bottom friction coefficients) and the same bathymetry configuration is applied to every case. We follow the methodology proposed and validated by Parker et al. (2019), whereby time series of evolving conditions are approximated by a series of steady-state simulations. The design points  $Z$  derived from applying the MDA algorithm are assumed spatially and temporally constant during three-day simulations. This allows for a dynamic equilibrium to be reached in the model (i.e., storm surge resulting from atmospheric pressure anomaly and wind shear reaches a fully developed dynamic equilibrium). Delft3D-FLOW and SWAN are dynamically coupled such that wave-induced currents and water levels are communicated between the two models every 20 min in simulated time.

The total dimensionality of  $M(Y)$  is 21 variables, with three atmospheric variables, nine wave variables, and nine water level variables. Atmospheric variables of  $PMSL$ ,  $U$ , and  $V$  are assumed spatially and temporally constant during each simulation. Wave variables ( $H_{NH}$ ,  $T_{NH}$ ,  $D_{NH}$ ,  $H_{SH}$ ,  $T_{SH}$ ,  $D_{SH}$ ,  $H_{SEA}$ ,  $T_{SEA}$ ,  $D_{SEA}$ ) are also assumed to form a constant multi-modal spectrum along the offshore boundary of the SWAN domain. The nine variables are combined into a single wave spectrum,  $S(f, \phi)$  composed of three separate JONSWAP peaks for each of the directional components. One of the water level variables is a mean water level composed of regional non-tidal residuals such as MMSLA, SLR, and the 18.6-year nodal tide cycle. The tide is then represented by the phase of

the leading eight constituents ( $M_2$ ,  $S_2$ ,  $N_2$ ,  $K_2$ ,  $K_1$ ,  $O_1$ ,  $P_1$ ,  $Q_1$ ). The water level at the end of a simulation is thus dynamic with respect to the tides. It is spatially varying and includes the nonlinear interactions between the mean water level, the surge induced water level, and the momentum introduced by the tides. See Parker et al. (2019) for further details regarding the error introduced by each of these simplifications for a similarly designed simulator using ADCIRC-SWAN in Grays Harbor, Washington.

The water level and complete directional wave spectra at the offshore end of each XBeach transect were extracted in the last instance of the 3-day Delft3D-SWAN simulations. XBeach simulations were then run in hydrostatic-mode for a single hour of waves propagating over a stationary water level. Transects begin offshore in approximately 15 m water depth and continue across the barrier into San Diego Bay. XBeach in hydrostatic-mode solves short-wave amplitude variations separately from the long waves and currents created by radiation stress gradients. The model's formulism saves considerable computational time and provides a useful dynamical representation of the cross-shore variation in wave-induced water levels and the long-wave dominated runup on the beach face. Maximum runup on the beach, maximum flood depth and cross-shore distance of water behind the dune, as well as mean cross-shore water level are saved from each simulation.

#### 4. Emulator Results and Validation

GPR emulators were created to predict still water levels (Section 4.1), bulk wave parameters (Section 4.2), and dynamic surf zone water levels containing wave setup (Section 4.3). K-fold validations and error metrics for each of the surrogate models can be found in Supporting Information S1 (Arlot & Celisse, 2010). Direct comparison of surrogate model hindcasts to observed tide gauge and wave records are provided in this section to quantify the cumulative error of the coupled simplified simulator and GPR design. Section 4.4 demonstrates a surrogate modeling configuration for predicting the binary decision of runup versus wave overtopping and the ultimate total water level at the artificial berm.

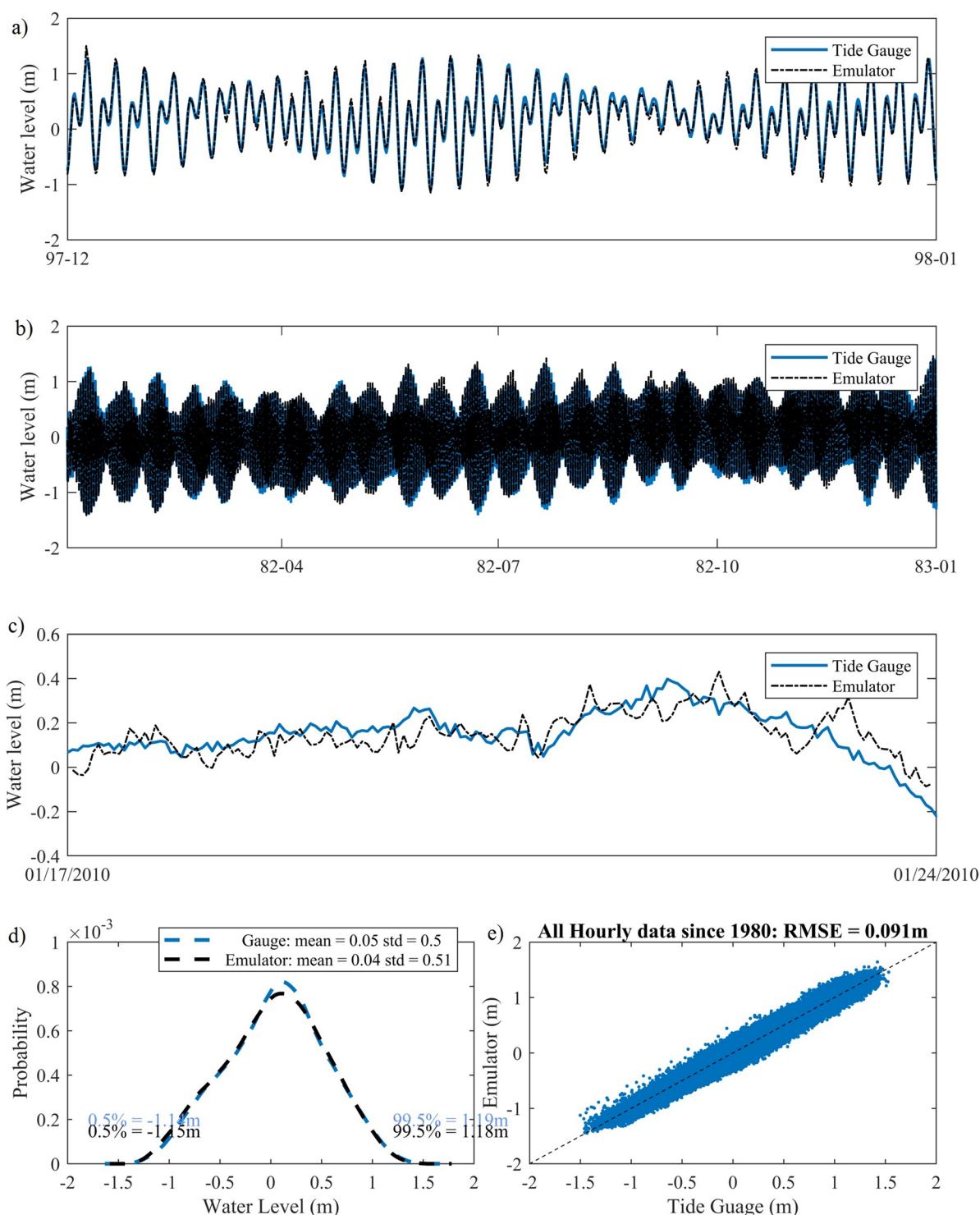
##### 4.1. Emulating Still Water Levels

Surrogate models can be created between the offshore conditions and the simulated water levels at any node within the numerical model domain. The nodes located at the La Jolla pier tide gauge and the San Diego Bay tide gauge (shown in Figure 2) are used for validation of the ability of the surrogate model to reproduce tide and surge dynamics (see Text S1 in Supporting Information S1 for quantified errors introduced by only the use of GPR). A surrogate model hindcast of 34 years of water level observations at the La Jolla tide gauge produces an RMSE of 9 cm, with mean depths and standard deviations differing by less than 1 cm (Figure 5d). Similar results are obtained at the San Diego Bay tide gauge (not shown), where RMSE is 10 cm, suggesting the framework can capture tidal propagation dynamics between the open ocean and back bay despite using a subset of only eight constituents and training on 600 interference patterns of those constituents (infinite number of possible combinations).

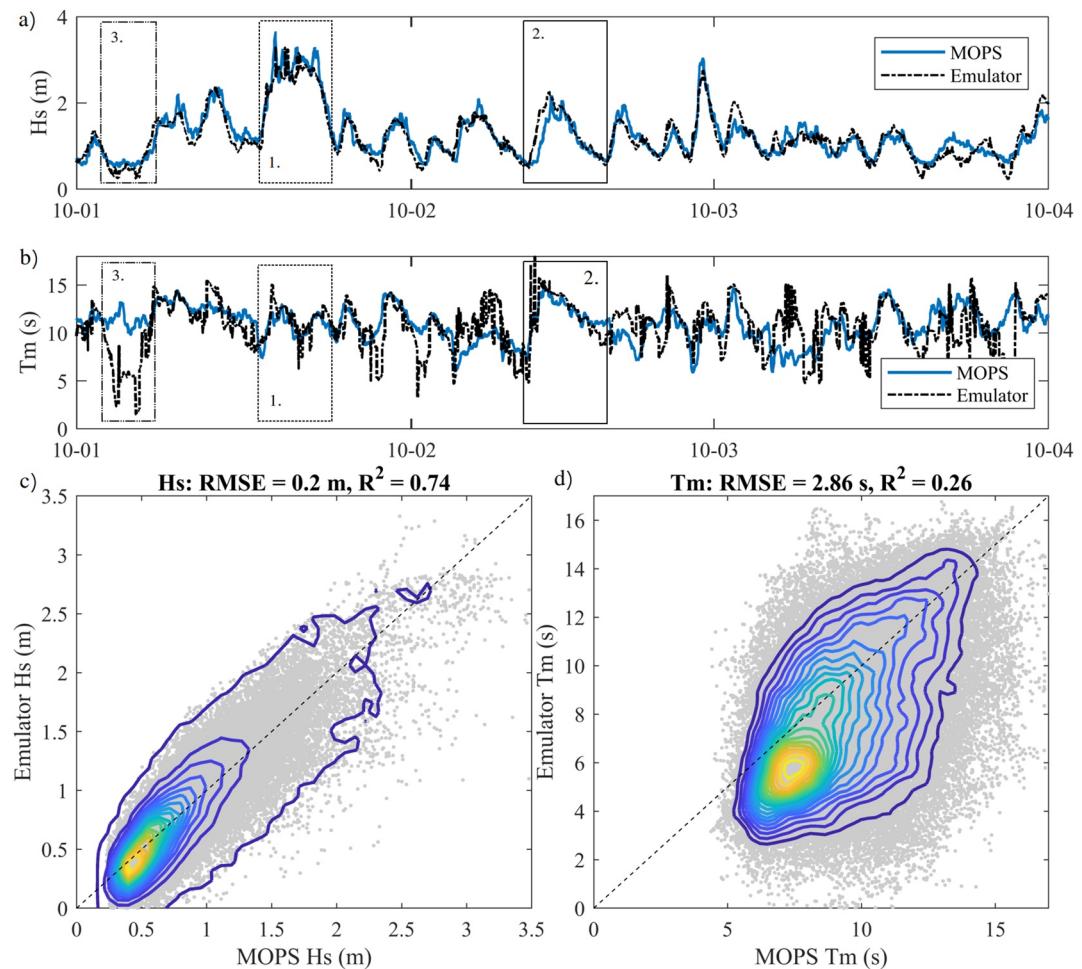
The surrogate model additionally emulates non-tidal dynamics driving water level variability about the tidal signal. Figure 5c provides a temporal comparison of observed NTRs and emulated NTRs for the week containing the largest storm surge on record. NTRs from both the observed tide gauge and emulated tide gauge were identified using the approach of Serafin and Ruggiero (2014), which removes frequencies relevant to tidal dynamics. Figure 5e extends this analysis to NTRs from all hours in the 34-year hindcast, with RMSE of 7.9 cm and  $R^2$  of 0.64. It should be noted that the intended application of the framework is to consider water levels during future elevated sea levels. Of the 600 cases training the surrogate model, 280 cases simulate regional water levels higher than those observed in the historical record. The validation in Figure 5 is thus concentrated in multi-dimensional space such that GPR interpolation is between a subset of the training cases as opposed to being optimized for the validation space.

##### 4.2. Emulating Nearshore Waves

Nearshore wave characteristics in southern California are variable in the alongshore due to complicated wave refraction around the Channel Islands, complex bathymetry, and the confluence of waves from the north Pacific, south Pacific, and local winds. The emulator in this study accounts for this variability by including nine parameters defining the spectrum as well as two parameters defining wind and allowing SWAN to generate waves within



**Figure 5.** Predicted water levels by a 21-to-1 dimensional GPR model at the same location as the La Jolla pier tide gauge. (a) Tide gauge records and emulated water level for December 1997, one of the highest MMSLA months on record. (b) Tide gauge records and emulated water levels for all of 1982, capturing spring-neap cycles as well as a 12-month seasonal signal. (c) Non-tidal residuals observed at the tide gauge records and emulated NTRs for the largest storm surge on record. 1-to-1 comparison of (d) emulated water levels and (e) emulated non-tidal residuals compared to observations for all hours from 1980 to 2013. Gray points denote hourly values and density contours are provided to highlight concentration of points along 1-to-1.



**Figure 6.** Validation of emulated nearshore wave conditions by time series of (a) wave height and (b) wave period for a 3 month period during the El Niño Modoki in winter 2010 (contained the largest wave event in the record), and 1-to-1 of (c) wave height and (d) wave period for all hours between 2000 and 2013 presented as gray points and density contours highlighting concentrations within the point cloud.

the domain. Validation of the emulator's ability to capture wave propagation and transformation from deep water into shallow water is performed by comparison to the CDIP MOPS product at a location in 10 m of water offshore of Coronado Beach. Wave characteristics were extracted from the appropriate XBeach model after transformation, and emulators for significant wave height and average wave period composed of separate 21-to-1 GPRs were directly compared to the 14 years of MOPS waves (Figure 6). Three months of temporal variability from the winter of the 2009–2010 El Niño are provided to highlight notable behavior within the emulator (Figures 6a and 6b). The prolonged large wave event in January is the largest event in the 14 years and was dynamically driven by a regional storm event (highlighted by dotted box labeled 1). The magnitude and duration is accurately captured by the emulator, including period variations during the event. Dynamically the emulator also captures the arrival of a distant swell event in the middle of February (solid line box labeled 2). Note the jump to long wave periods that travel faster and thus arrive first, followed by steady decrease as slower, shorter wave periods arrive later in the event.

The three month subset also includes an example of some of the largest errors in the emulator, in particular with respect to wave period in early January (dashed box labeled 3). Although wave height is accurately predicted to be very small, the period is incorrect by 10 s (some of the gray dots lying furthest from the 1-to-1 line in Figure 6d). This highlights how errors in periods are largest for low wave energy events, which has a net effect of biasing the emulator low compared to MOPS, and produces a relatively poor  $R^2 = 0.26$  despite the temporal variability in Figure 6b appearing realistic. It should be noted that the MOPS spectra are pieced together with certain

frequencies from deeper water buoys, other frequencies from shallow water buoys, and a number of smoothing assumptions (see O'Reilly et al., 2016), while the emulator pieces together three JONSWAPs to a single spectrum that is further transformed by both SWAN and XBeach. However, despite such spectra manipulation, the emulator captures magnitudes and temporal dynamics associated with both distant and locally-generated large swells that contribute to open-coast total water levels.

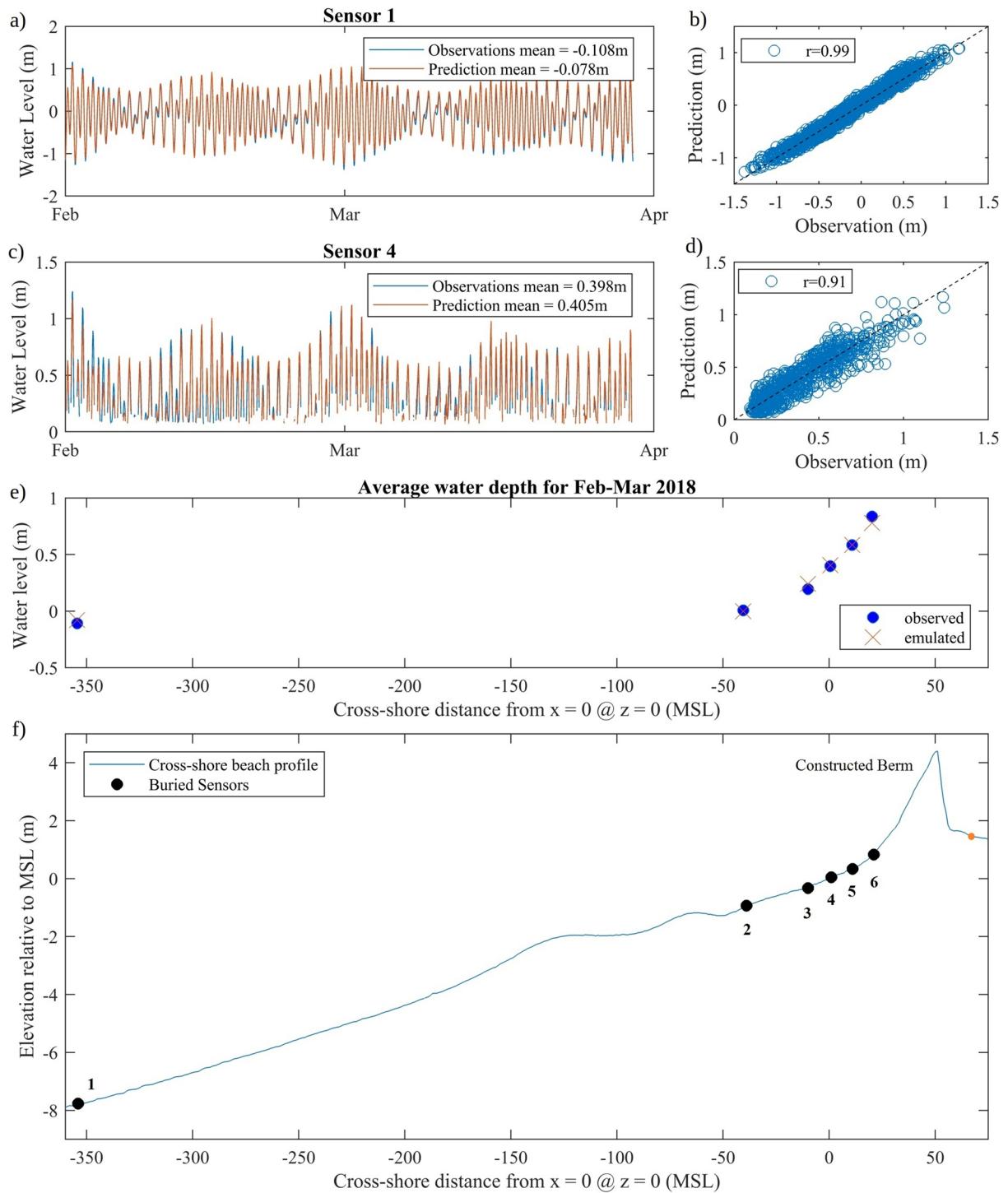
#### 4.3. Emulating Surf Zone Water Levels

The XBeach simulations provide a single realization of the surf zone contingent on the spectra and water levels derived from each SWAN and Delft3D simulation. The spectra are sampled in a random manner such that none of the XBeach simulations are replicates of the exact wave by wave forcing observed during the field study described in Section 2.4. Further, none of the 600 cases modeled for surrogate model development are forced with 21 dimensions that exactly match the observed atmospheric and oceanic conditions during the field study. The coupling of XBeach is intended to provide information regarding the runup elevations and potential for dune overwash not resolved by Delft3D. Default hydrodynamic, wave dissipation, and bottom friction parameters were used in XBeach to avoid a biased calibration to the specific bathymetric profile and the limited range of relatively mild conditions observed during the field study. Specific observations of runup and dune overtopping were unavailable precluding a validation with respect to these parameters. However, assessing the ability to emulate wave induced processes is necessary to determine if the wave-induced component of open beach water levels is captured by the proposed framework. This was evaluated with respect to wave-averaged observations such as wave induced set-down and setup, and the cross-shore variability observed during the field study. The mean water depth simulated within each XBeach case was used to create a water level surrogate model at each pressure sensor location in the cross-shore of the instrumented transect immediately offshore of the NABC (Figure 7f). The surrogate models were created with the same approach as those already demonstrated, relating the 21 dimensions of forcing provided at the offshore boundaries to the XBeach mean water level predicted at a single location. If a sensor location did not get wet during a simulation due to low tides and/or small wave forcing, it was left out of the GPR training data set. This data was removed from training because the prediction of wet versus dry, or bed level versus water level, is effectively a step-function rather than a continuum.

Figures 7a and 7c provide examples of hourly mean water level observations between February and March 2018, directly compared to the emulated prediction at two different cross-shore locations. In these emulations, when the surrogate model predicts a water level at or beneath the bed level it is assumed that water did not reach that beach elevation. Figure 7e provides the mean observed and mean emulated water levels for all hours during the two month field study when all six sensors were wet. The cross-shore variability in the observed levels indicates that wave-induced setup within the surf zone is captured by the sensors. Previous studies have suggested that XBeach can simulate wave setup and infragravity swash on dissipative beaches with saturated surf zones but note that predictions are less accurate on steep beaches (Palmsten & Splinter, 2016; Stockdon et al., 2014). On Coronado Beach, the XBeach surrogate models reasonably resolve a similar cross-shore variation, indicating that the wave-induced component of open beach water levels is captured by the framework with realistic magnitudes. Similar to earlier validations (Sections 4.1 and 4.2), the conditions observed in the field study are from a narrow subset of the full range covered by the 600 cases in the training library. Calibrated XBeach parameters and more design cases representative of conditions during the two month experiment would likely improve the comparisons.

#### 4.4. Emulating Runup and Dune Overtopping

The mean water level validation provides confidence that XBeach surrogate models can resolve parameters relevant to wave-induced flooding, such as wave runup elevations and dune overtopping volumes, although no observations exist for validation of such parameters. Runup was tracked throughout time within each simulation and is defined here as the maximum elevation and landward reach within the hour-long Xbeach simulation. Unlike the water level simulations, where surrogate model nodes are both stationary and wet in all cases, the occurrence of dune overtopping is a binary problem. In other words, overwash occurs in only a subset of the simulations and is thus a true/false classification result as opposed to a continuum. We have accounted for this behavior by including a binary predictive model to classify every point in emulated time as either wave runup on the beachfront or dune overtopping. This decision triggers either a surrogate model for runup elevation on the berm face or a surrogate



**Figure 7.** In-situ wave-averaged water depth observations during the nearshore deployment of February and March 2018 compared to a surrogate model's predictions. (a) At the offshore sensor with respect to time and (b) a 1-to-1 comparison. (c) At a mid-surf zone sensor with respect to time and (d) a 1-to-1 comparison. (e) The average observed and average emulated water level at all six sensors. (f) The cross-shore beach profile used in XBeach simulations with sensor locations marked by black circles. An orange circle also denotes the location of the flood-depth surrogate model.

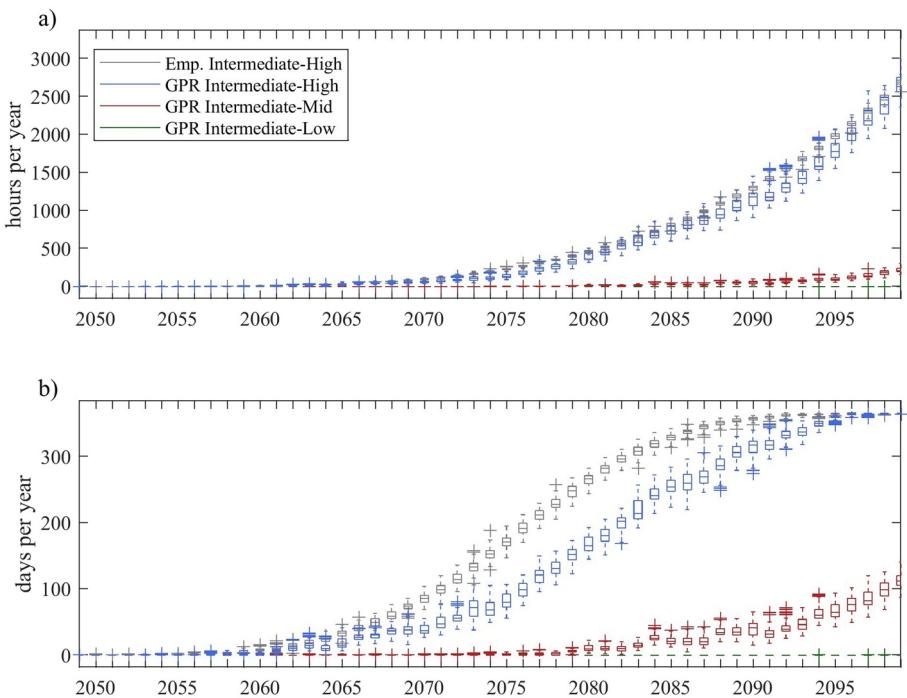
model for water depth behind the berm using a GPR model trained on only the relevant subset of simulations. Predictive models for classification problems, or true/false problems, include logistic regression, support vector machines, neural networks, and random forests (Goldstein et al., 2019). Each have advantages and disadvantages for this particular application, but the methods provided here are a demonstration of utility as opposed to an exhaustive investigation of the skill of different machine learning algorithms. Random forests were chosen because they are a classifier with the ability to become more general with more observations, as opposed to more tuned and potentially over-fit (Breiman, 2001) (see Text S3 in Supporting Information S1 for more details regarding random forest and accounting for an unbalanced data set).

Overwash was identified as any XBeach simulation where runup exceeded the dune crest. The depth of the water immediately behind the berm (orange dot in Figure 7) was extracted at the end of each simulation and defined as the backshore flood depth. Separate surrogate models for runup elevation and flood depth were then created using the entire data set. Five K-fold cross validations each withholding 20% of the training library from GRP calibration resulted in root-mean square errors for the flood depth and runup elevation surrogates of 0.91 and 0.38 m, respectively. While the errors are considerably larger than those observed at the tide gauges, both models are attempting to predict higher order physical processes with large ranges (~3 and ~6 m, respectively). Both are affected by the details of the bathymetry and topography, but the flood depth in particular includes a topography constrained maximum. Wave-transported water fills a low point, eventually exceeding the elevation of a parking lot behind the berm, and any additional volume thus continues to spill over into more distant backshore locations without increasing the flood depth behind the berm.

## 5. Stochastic Emulation of Hypothetical Futures

The complete framework described herein produces hypothetical hourly forcing conditions that downscale to time-dependent local oceanographic parameters. To demonstrate the power of the fully stochastic coupled framework, 500 different iterations of large-scale climate, and synoptic weather from 2020 to 2100 were simulated with the three different SLR scenarios discussed in Section 2.3. The time-explicit hourly products can be used to answer a broad range of questions useful for coastal planning and risk assessment. Previous studies have commonly focused on water levels that exceed elevation thresholds relevant to flooding, which is an ideal peak-over-threshold query for probabilistic frameworks focused on extreme values (e.g., Ghanbari et al., 2019; Sweet & Park, 2014). Such analysis pertains to when in the future specific structures will regularly become impacted, which may indicate a need for management plans to account for downtime, moving activities altogether, or implementing/strengthening flood defenses. Output from the hybrid framework presented in this work can be applied to the same questions with similar results. Figure 8 provides a forecast for the number of impact hours per year that total water levels exceed a use elevation threshold for a pier subjected to the Intermediate-High SLR curve on the bay-side northeast corner of the NABC. The yearly box plots signify the variability in yearly hours across all 500 simulations. The pier is not far from the San Diego tide gauge (Figure 1), allowing for a comparison to other approaches. The gray boxplots in Figure 8 are derived by shifting the historically observed residual SWL distribution by the Intermediate-High SLR curve and randomly sampling at an hourly frequency to create 500 time series analogous to the hybrid output. The impact hours per year from both methods follow the same trajectory (Figure 8a), demonstrating that from the perspective of a cumulative metric, the hybrid approach produces results consistent with traditional approaches that assume the historical distribution of still water levels will remain stationary (e.g., Ghanbari et al., 2019).

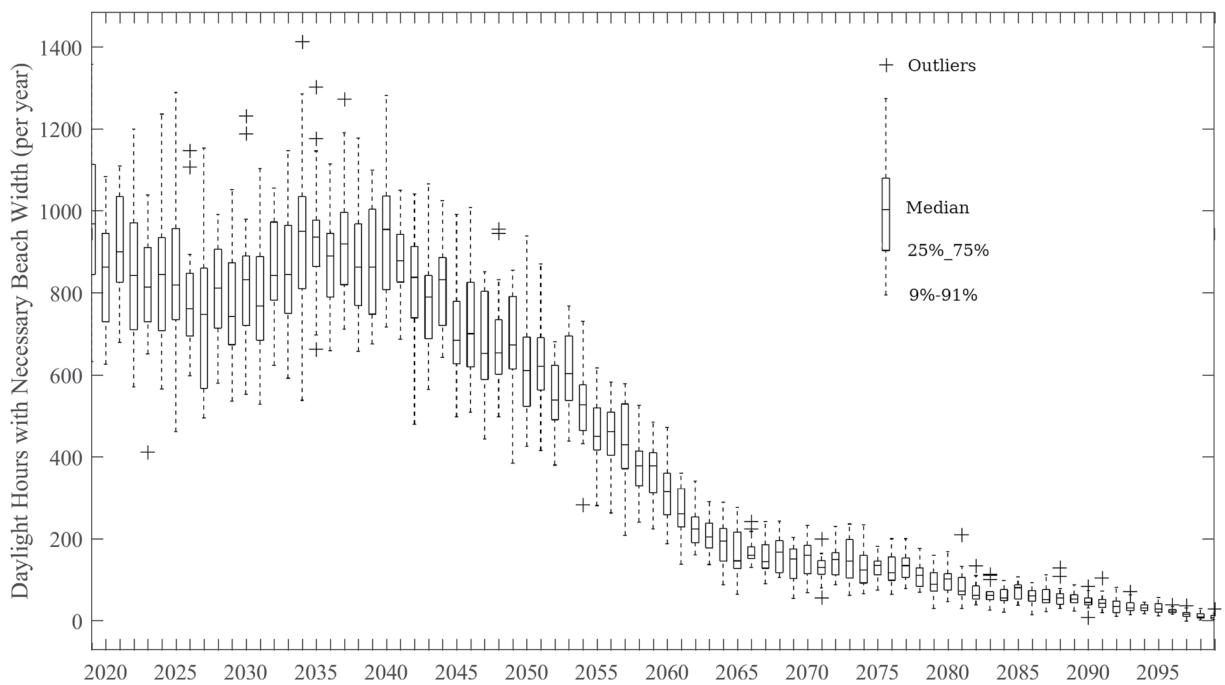
However, the number of days experiencing an impact, and thus disrupting naval base activities is potentially more relevant for understanding whether the exceedance is a nuisance or debilitating. The same time series in Figure 8a are assessed by the number of independent days per year exhibiting flooding in Figure 8b. Note that the probabilistic and hybrid Intermediate-High scenarios diverge within this particular metric. This is a consequence of the hybrid framework containing chronological behavior that results in multiple impact hours occurring on the same day, while the simpler approach of randomly sampling from the historical probabilities ultimately distributes those hours over more days such that days per year is biased higher (gray boxplots in Figure 8b). This is traditionally overcome by performing completely separate extreme value analyses for events with specific characteristics or for metrics at different temporal scales (i.e., hourly, daily, annual) (e.g., Ghanbari et al., 2019). For example, Sweet and Park (2014) used an empirical quadratic equation fit to observed water levels to project future nuisance flooding in San Diego at a daily scale. They defined a tipping point as being the first year with



**Figure 8.** Yearly box plots from 500 simulations of the (a) hours per year and (b) days per year that water level exceeded a threshold for pier usability.

at least 30 days of water levels above 0.5 m mean high high-water (MHHW), and found that 60% of simulations with SLR curves for RCPs 2.6, 4.5, and 8.5 resulted in a tipping point occurring between 2030 and 2040 (SLR curves from R. W. Kopp et al. [2014]). All three SLR curves in that work project water levels by 2100 that fall between the Intermediate-Low and Intermediate-Mid scenario in this work. The hybrid framework predicts the Sweet and Park (2014) tipping point to occur on average in 2029 for the Intermediate-Mid SLR curve and 2041 for the Intermediate-Low curve. This suggests that the hybrid framework projects future water levels at the daily timescale that are in line with traditionally accepted methods. More notably, by starting with the underlying climate as the principal driver of the hybrid framework, the conditional dependence of such tipping points on large-scale climate can be assessed. In this scenario, years exhibiting either El Niño or El Niño Modoki SST anomalies accounted for 47% of the Intermediate-Mid tipping points and 60% of the Intermediate-Low tipping points despite such climate phenomena occurring in only 24% of years in the total simulation record. This demonstrates that metrics across multiple temporal scales can be derived from the same framework without needing separate independent forecasting methods.

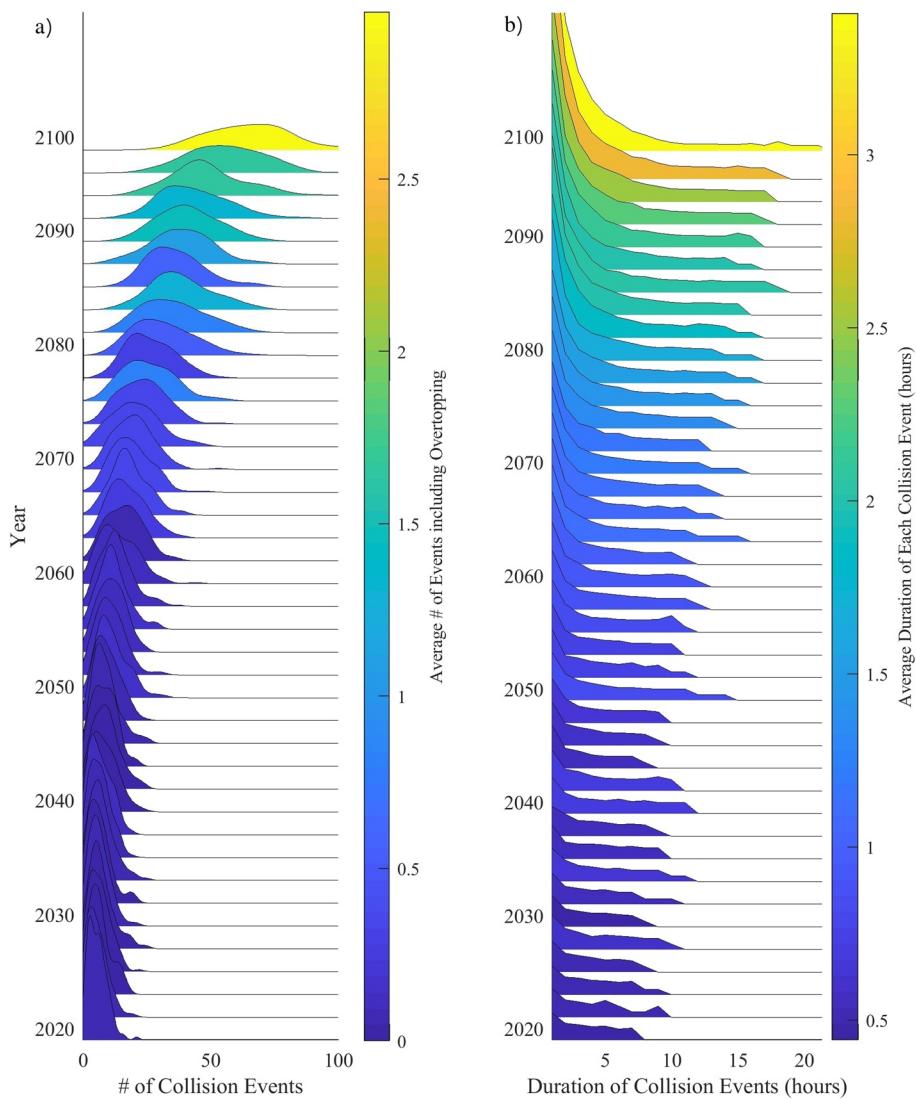
An additional advantage of the presented framework is that surrogate models can be created for any location within the model domain, potentially aiding with naval base planning even in locations with no historical observations. Consider Figure 9, which assumes the beach profile is steady into the future to identify the number of daylight hours (between 6 a.m. and 6 p.m.) with an Intermediate-Mid SLR curve when the beach would be wide enough to conduct a particular training activity (e.g., at least 50 m between runup extent and dune toe). No historical observations of beach width exist at this location, preventing traditional data-driven forecasting approaches. Further, this beach usability metric is time-dependent, necessitating frameworks containing chronological behavior. Figure 9 highlights the insight that can be derived with respect to interannual variability and future tipping points. The interannual variability within this metric is dominated by the 18.6-year nodal oscillation until mid-century, but a tipping point is eventually reached as the Intermediate-Mid SLR parabola overcomes climate variability and the number of usable daylight hours rapidly declines for multiple decades. A transition to a much slower decline then occurs in the late 2060s and persists until the end of the century. This behavior is due to the majority of the water level distribution exceeding the elevation contour associated with a 50-m-wide beach such that the necessary beach width for training is only experienced during low water extremes. Yearly variability due to the wave climate also consistently declines once the SLR scenario forces a mid-century drop in daylight



**Figure 9.** Identifying daylight hours between 6 AM and 6 PM when the beach width exceeds 50 m for Naval training activities. Yearly box plots and whiskers define the variability across 500 simulations with a 1.0 m SLR scenario imposed on all simulations.

training hours. This example highlights the interplay between wave climate variability, tidal variability, and a particular SLR scenario. Constraints on when the tipping point occurs are most sensitive to the SLR scenario, but the variability across the 500 simulations indicates that cumulative variability resulting from waves, winds, and sea level anomalies will continue to be of an order of magnitude that is relevant to naval base management decisions throughout the century.

Incorporating the hybrid downscaling enables more detailed queries and unique metrics relevant to coastal management on the open coast. For example, the long term viability of the artificial berm fronting the NABC is not necessarily a question of the number of impact hours per year, but rather how those hours occur. Berm impacts resulting from runup are commonly assessed by the Storm Impact Scale (Sallenger Jr., 2000), which separates the maximum runup elevations into swash, collision, overwash, and flood regimes. Allocating resources to re-build the artificial berm does not need to occur every time there is dune collision (runup exceeding the dune toe). Such action becomes necessary after several consecutive hours of collision cumulatively undermines dune integrity, or after overtopping reduces the height of the dune top. Consecutive hours of collision can be identified in the runup emulator and grouped into a single event with an associated duration. Figure 10a demonstrates how the yearly total number of collision events evolves through time when considering the Intermediate-Mid SLR curve. The hybrid framework also distinguishes which of those events includes any amount of overtopping during the event (designated by the color of the distribution). Figure 10b provides yearly distributions for the corresponding durations of those collision events. The early half of the century is predominantly composed of several events per year with longer durations (the extreme storms). As time progresses the number of events per year increases, and the envelope of variability across the 500 simulations widens. The average duration also increases with time as longer duration events occur in the later half of the century. However the number of relatively short events increases considerably by 2100, skewing the distribution more and more with time. Local coastal managers could use this information, coupled with their experience of historical berm performance, to develop site-specific tipping points beyond which the artificial berm will become unmanageable. Consider a scenario in which berm reconstruction becomes unsustainable once there are at least six events with either overtopping or more than three consecutive hours of collision in a given year. The 500 simulations in Figure 10 would indicate an average tipping point across all climate simulations of 2071. However, 25% of the simulations reach the tipping point by 2063, and 10% by as early as 2051. Alternatively, another 25% do not reach the tipping point until 2078, and 10% are



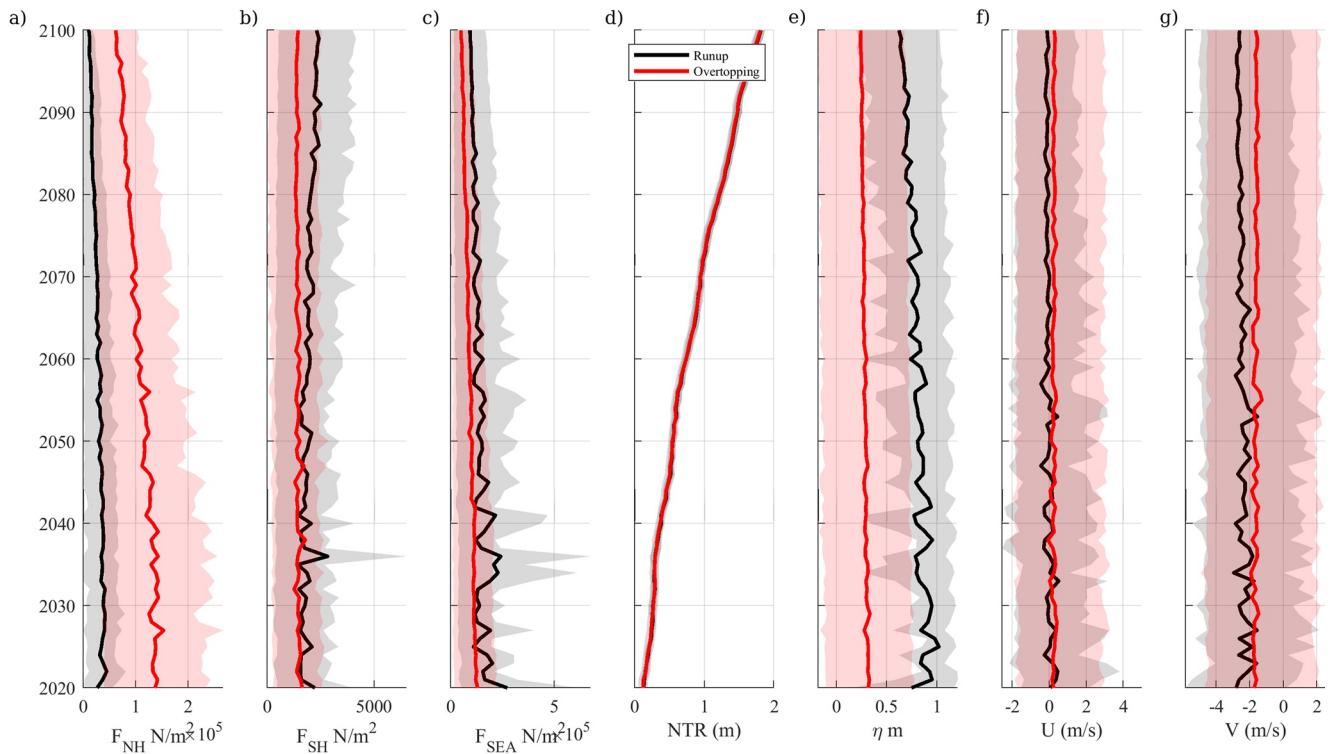
**Figure 10.** (a) Distributions for the yearly number of artificial berm collision events derived from the runup surrogate model coupled with 500 simulations. Distribution are provided every two years to reduce visual clutter. Color of the distribution depends on the number of overtopping events as derived from a separate overtopping surrogate model. (b) The durations of the same events portrayed in (a) are provided with color indicating the average duration.

not until after 2083. This highlights the wide range of potential futures that could result as a consequence of the stochastic nature of storm activity.

An additional advantage of the hybrid framework is the ability to explicitly link detailed forcing conditions to the ultimate hazard being considered. Figure 11 summarizes the environmental forcing causing all hours of either berm collision or overtopping identified in the events in Figure 10. Average collision conditions for each year are defined by the black line and average overtopping conditions by the red line. Shaded patches define the 16th and 84th percentiles of the distribution of conditions within each year. Wave energy flux from the north ( $F_{NH}$ ), south ( $F_{SH}$ ), and local wind seas ( $F_{SEA}$ ) are calculated as:

$$F = \frac{\rho g^2 H^2 T}{64\pi} \quad (5)$$

where  $\rho = 1,025 \text{ kg/m}^3$ . All eight tidal phases are summed into a single variable,  $\eta$ , while the SLR, MMS-LA, and 18.6-year nodal cycle are also summed to a single variable  $NTR$ . As the sea level rises during the



**Figure 11.** The forcing conditions responsible for all collision and overtopping events identified in Figure 10. (a) Wave energy flux from the northern hemisphere,  $F_{NH}$ , (b) wave energy flux from the southern hemisphere,  $F_{SH}$ , (c) wave energy flux from the local wind sea,  $F_{SEA}$ , (d) non-tidal residuals, NTR, as a summation of SLR, MMSLA, and the 18.6-year nodal tide cycle, (e) tide height at the peak of the event, (f) East-West wind component, U, and (g) North-South wind component, V. Average conditions causing berm collision/overtopping are denoted by the black/red lines, and transparent shaded patches represent 16th and 84th percentiles of the distribution.

Intermediate-Mid scenario, the average tidal elevation during dune collision lowers and the distribution of tidal elevations widens (Figure 11e). Similarly, the necessary  $F_{NH}$  and  $F_{SEA}$  to produce both collision and overtopping reduces as sea levels rise (Figures 11a and 11c). Conversely, the average  $F_{SH}$  driving collision increases with time (Figure 11b). This dynamic is a consequence of all three sea/swell states contributing to any one runup prediction. Large  $F_{NH}$  and  $F_{SEA}$  accompanied by relatively small  $F_{SH}$  dominate sea states inducing collision or overtopping in the early half of the century, resulting in a low average  $F_{SH}$ . However sea states with low  $F_{NH}$  and  $F_{SEA}$  accompanied by larger  $F_{SH}$  begin to affect the berm more by the later half of the century due to elevated still water levels, ultimately increasing the average  $F_{SH}$  contributing to collision. Taken as a whole, Figures 10 and 11 are representative of when nuisance berm collision begins to occur in the future. The hybrid framework identifies the changing wave forces responsible for potential future local impacts despite only accounting for climate change projections through SLR curves. However, climate change projections for other dimensions, such as wave climate change (e.g., L. H. Erikson et al., 2015), could be implemented in future work because the framework is fundamentally built on distributions specific to each forcing parameter.

Ultimately, the coupling of stochastic climate conditions with surrogate models enables the analyses demonstrated in Figures 10 and 11. The projections provide considerably more context for coastal managers compared to traditional exceedance approaches (e.g., Figure 8). Computational effort is required on the front-end for developing the library (the 600 CoSMoS simulations completed in 10 days on a cluster with 24 2.1 GHz processor cores), but applying output from the climate emulator to the surrogate models is computationally fast. Extracting water levels from 80 years of hourly environmental conditions (21 forcing dimensions) requires only about 4 s on a laptop with an Intel i5 2.6 GHz processor. This enables uncertainties and extremes to be identified akin to data-driven methods while retaining the benefits of dynamical models (information at locations with no historical data, nonlinear dynamics, alongshore variability due to bathymetric variations). The projections are also readily adaptable to future changes in SLR projections (Nicholls et al., 2021), as only a single forcing dimension needs to be updated and the same library with associated surrogates quickly produces new results.

### 5.1. Limitations and Future Improvements

The methods of this work aim to remove subjective decisions from the hybrid statistical dynamical modeling framework, but are not exhaustive with respect to the many techniques currently being developed within the burgeoning field of machine learning derived surrogate models. Detailed sensitivities with respect to library size (Loeppky et al., 2009), design point selection algorithms (Camus et al., 2011), and accounting for unbalanced data sets are provided in Supporting Information S1. Perhaps the most significant limitation is that the bathymetry within Delft3D and beach morphology in XBeach were held static in all dynamical simulations (Serafin, Ruggerio, Barnard, et al., 2019). This simplification allowed the GPRs to be developed on environmental forcing, reducing the total number of dimensions to 21. Accounting for bathymetric evolution would necessitate adding an additional dimension for every node where bathymetry varied. The static assumption is significant considering that there is seasonal variability in the beach slope, and even the potential for slope changes within individual weather events as large waves can alter the beach slope over the course of a storm. Long-term changes to the lower beach profile and the berm profile are also neglected but can be relevant to long-term projections of runup and overwash. However it should be noted that the study area is a highly engineered shoreline, with the berm profile routinely reconstructed to pre-storm profiles. Accounting for future engineering decisions is beyond the scope of this paper, and the current abilities of surrogate modeling, but the derived output could aide such future decisions as this modeling framework provides projections for runup and overwash that could occur if the engineered shoreline is maintained as currently managed.

The surrogate models presented in this work were created for specific locations within the domain. Creating a surrogate model for every node where water levels and waves are computed in the numerical model would be computationally limiting (fitting tens of thousands of individual GPRs). However, increasing the number of output dimensions to include more locations reduces the quality of the prediction at every location (i.e., GPR of 21 input dimensions to N output dimensions as opposed to 21 to 1). A promising technique for reducing the dimensionality of the problem is to make use of the correlation between neighboring water levels. As an example, principal component analysis can be applied to the spatial water level output of the training library and a reduced number of surrogate models can be created to predict the magnitude of the dominant spatial modes (Jia et al., 2016). Such an analysis is the subject of ongoing work (T. Anderson, 2019), which could have the potential to derive spatially varying water levels to readily predict flooding along the entire extent of the bay shoreline.

## 6. Summary

The hybrid statistical-dynamical framework presented in this work demonstrates an approach where multiple burgeoning risk assessment techniques are coupled to assess potential coastal hazards for a large number of hypothetical future iterations of synoptic weather. An efficient stochastic climate emulator was coupled with surrogate models for coastal water levels, wave-induced flooding, and wave runup to assess the likelihood of potential future flooding events at an artificial berm fronting San Diego, CA. The surrogate models were built using 600 different cases of a dynamically coupled hydrodynamic-wave numerical model (Delft3D-FLOW and SWAN) and a loosely coupled nearshore surf zone model (XBeach). Such a framework is necessary for locations where extremes are dominated by compound events with varying contributions from wind-induced, wave-induced, river-induced, and even temperature-induced water level anomalies.

The output accounts for multiple time scales of variability in the climate, while also containing nonlinear feedbacks between wave, surge, and tidal processes learned from a library of dynamic simulations. Gaussian process regression applied to a simplified, yet high-fidelity, numerical simulator was able to predict hourly historic water levels with 10 cm RMSE. Application of the surrogate models to multiple climate simulations naturally leads to a wide range of future resilience metrics with quantified uncertainty due to variability in the timing and magnitude of environmental forcings, which are inherently due to variability in climate and synoptic weather. The resilience metrics can provide greater context for both managers and researchers, with ability to track chronological behavior while explicitly relating the local impact to the responsible offshore conditions and climate forcing. The efficient framework addresses multiple limitations in current hazard frameworks. Although the upfront computation cost is not negligible, the library of cases requires multiple orders of magnitude less computing power than fully dynamic simulations of many hypothetical future climates. The results are also easily updated if and when future advances in constrained SLR scenarios are available. The complete approach can readily explore sensitivities to a

wide range of impact metrics, while capturing the uncertainty that exist in those metrics as a result of stochastic climate processes.

## Data Availability Statement

Open-source codes used in this study to synthesize climate predictors and generate local time series of coastal conditions can be accessed at <http://github.com/anderdyl/teslaCoSMoS> (<https://doi.org/10.5281/zenodo.5055555>). Open-source codes for the Delft3D modeling framework can be downloaded at <https://oss.deltares.nl/web/delft3d/downloads>, and XBeach can be downloaded at <https://oss.deltares.nl/web/xbeach/downloads>. The NCEP Climate Forecast System Reanalysis (CFSR) is freely available at <https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/climate-forecast-system-version2-cfsv2> and Extended Reconstructed Sea Surface Temperature Version 4 (ERSSTv4) data are freely available through the National Oceanic and Atmospheric Administration (NOAA) from their website at <https://data.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00884>. Tide gauge records are also freely available from NOAA for San Diego Bay, CA at <https://tidesandcurrents.noaa.gov/stationhome.html?id=9410170> and the open ocean pier in La Jolla, CA at <https://tidesandcurrents.noaa.gov/stationhome.html?id=9410230>.

## Acknowledgments

This work was funded by the Strategic Environmental Research Development Program (DOD/SERDP RC-2644). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. F. J. Mendez, A. Rueda, and L. Cagigal acknowledge the partial funding from the Spanish Ministry of Science and Innovation, project Beach4cast PID2019-107053RB-I00. The authors thank the Scripps Center for Coastal Studies for their efforts to deploy, recover, and process surf zone pressure sensor data used as validation in this study. The authors thank Melisa Menendez for sharing GOW2 hindcast data for Southern California. The authors thank the sea-level rise projection authors for developing and making the sea-level rise projections available, multiple funding agencies for supporting the development of the projections, and the NASA Sea-Level Change Team for developing and hosting the IPCC AR6 Sea-Level Projection Tool.

## References

- Allan, J. C., Ruggiero, P., Garcia-Medina, G., O'Brien, F., Stimely, L., & Roberts, J. (2015). *Coastal flood hazard study, Tillamook County Oregon (Technical Report)*. Oregon Department of Geology and Mineral Industries.
- Anderson, D., Rueda, A., Cagigal, L., Antolinez, J. A., Mendez, F. J., & Ruggiero, P. (2019). Time varying emulator for short and long-term analysis of coastal flood hazard potential. *Journal of Geophysical Research: Oceans*. <https://doi.org/10.1029/2019JC015312>
- Anderson, T. (2019). *Evaluating climate change induced coastal flood impacts to San Diego Bay Naval Infrastructure-a hybrid statistical dynamical modeling approach* (pp. 1–58). Master's Thesis at Oregon State University.
- Anjyo, K., & Lewis, J. P. (2011). RBF interpolation and Gaussian process regression through an RKHS formulation. *Journal of Math-for-Industry*, 3(6), 63–71.
- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40–79. <https://doi.org/10.1214/09-SS054>
- Barnard, P. L., Erikson, L. H., Foxgrover, A. C., Hart, J. A., Limber, P., O'Neill, A. C., & Jones, J. M. (2019). Dynamic flood modeling essential to assess the coastal impacts of climate change. *Scientific Reports*, 9(1), 1–13. <https://doi.org/10.1038/s41598-019-40742-z>
- Barnard, P. L., Erikson, L. H., O'Neill, A., Foxgrover, A. C., & Herdman, L. (2016). *CoSMoS v3.0 phase 2 flood-hazard projections: San Diego County (Technical Report)*. U.S. Geological Survey Data Release. <https://doi.org/10.5066/F7T151Q4>
- Barnard, P. L., Hoover, D., Hubbard, D. M., Snyder, A., Ludka, B. C., Allan, J., & Serafin, K. A. (2017). Extreme oceanographic forcing and coastal response due to the 2015–2016 El Niño. *Nature Communications*, 8, 1–8. <https://doi.org/10.1038/ncomms14365>
- Barnard, P. L., Short, A. D., Harley, M. D., Splinter, K. D., Vitousek, S., Turner, I. L., & Heathfield, D. K. (2015). Coastal vulnerability across the Pacific dominated by El Niño/southern oscillation. *Nature Geoscience*, 8, 801–808. <https://doi.org/10.1038/NGEO2539>
- Barnard, P. L., van Ormondt, M., Erikson, L. H., Eshleman, J., Hopke, C., Ruggiero, P., & Foxgrover, A. C. (2014). Development of the coastal storm modeling system (CoSMoS) for predicting the impact of storms on high-energy, active-margin coasts. *Natural Hazards*, 74(2), 1095–1125. <https://doi.org/10.1007/s11069-014-1236-y>
- Bass, B., & Bedient, P. (2018). Surrogate modeling of joint flood risk across coastal watersheds. *Journal of Hydrology*, 558, 159–173. <https://doi.org/10.1016/j.jhydrol.2018.01.014>
- Bilskie, M. V., Hagen, S. C., Medeiros, S. C., & Passeri, D. L. (2014). Dynamics of sea level rise and coastal flooding. *Geophysical Research Letters*, 41(3), 927–934. <https://doi.org/10.1002/2013GL058759>
- Blanning, R. W. (1975). Construction and implementation of metamodels. *Simulation*, 26(6), 177–184. <https://doi.org/10.1177/003754977502400606>
- Breiman, L. E. O. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/a:1010933404324>
- Cagigal, L., Rudea, A., Anderson, D., Ruggiero, P., Merrifield, M., Montano, J., & Mendez, F. J. (2020). A multivariate, stochastic, climate-based wave emulator for shoreline change modelling. *Ocean Modelling*, 154. <https://doi.org/10.1016/j.ocemod.2020.101695>
- Callaghan, D. P., Nielsen, P., Short, A., & Ranasinghe, R. (2008). Statistical simulation of wave climate and extreme beach erosion. *Coastal Engineering*, 55(5), 375–390. <https://doi.org/10.1016/j.coastaleng.2007.12.003>
- Camus, P., Méndez, F. J., Losada, I. J., Menéndez, M., Espejo, A., Pérez, J., & Guanche, Y. (2014). A method for finding the optimal predictor indices for local wave climate conditions. *Ocean Dynamics*, 64(7), 1025–1038. <https://doi.org/10.1007/s10236-014-0737-2>
- Camus, P., Mendez, F. J., Medina, R., & Cofiño, A. S. (2011). Analysis of clustering and selection algorithms for the study of multivariate wave climate. *Coastal Engineering*, 58(6), 453–462. <https://doi.org/10.1016/j.coastaleng.2011.02.003>
- Crosby, S. C., O'Reilly, W. C., & Guza, R. T. (2016). Modeling long-period swell in Southern California: Practical boundary conditions from buoy observations and global wave model predictions. *Journal of Atmospheric and Oceanic Technology*, 33(8), 1673–1690. <https://doi.org/10.1175/JTECH-D-16-0038.1>
- Dietrich, J. C., Zijlema, M., Westerink, J. J., Holthuijsen, L. H., Dawson, C., Luettich, R. A., & Stone, G. W. (2011). Modeling hurricane waves and storm surge using integrally-coupled, scalable computations. *Coastal Engineering*, 58(1), 45–65. <https://doi.org/10.1016/j.coastaleng.2010.08.001>
- Erikson, L., Barnard, P., O'Neill, A., Wood, N., Jones, J., Finzi-Hart, J., & Foxgrover, A. (2018). Projected 21st century coastal flooding in the Southern California Bight. Part 2: Tools for assessing climate change driven coastal hazards and socio-economic impacts. *Journal of Marine Science and Engineering*, 6, 19. <https://doi.org/10.3390/jmse6030076>
- Erikson, L. H., Espejo, A., Barnard, P. L., Serafin, K. A., Hegermiller, C. A., O'Neill, A., & Mendez, F. J. (2018). Identification of storm events and contiguous coastal sections for deterministic modeling of extreme coastal flood events in response to climate change. *Coastal Engineering*, 140, 316–330. <https://doi.org/10.1016/j.coastaleng.2018.08.003>

- Erikson, L. H., Hegermiller, C. A., Barnard, P. L., Ruggiero, P., & van Ormondt, M. (2015). Projected wave conditions in the Eastern North Pacific under the influence of two CMIP5 climate scenarios. *Ocean Modelling*, 96, 171–185. <https://doi.org/10.1016/j.ocemod.2015.07.004>
- Fox-Kemper, B., Hewitt, H., Xiao, C., Aalgeirsdóttir, G., Drijfhout, S., & Edwards, T. (2021). Ocean, cryosphere and sea level change. *Climate Change*.
- Gallien, T. (2016). Validated coastal flood modeling at Imperial Beach, California: Comparing total water level, empirical and numerical overtopping methodologies. *Coastal Engineering*, 111, 95–104. <https://doi.org/10.1016/j.coastaleng.2016.01.014>
- Garner, G. G., Hermans, T., Kopp, R. E., Slangen, A. B. A., Edwards, T. L., Levermann, A., & Pearson, B. (2021). Framework for assessing changes to sea-level (facts)IPCC ar6 sea-level rise projections, version 20210809. Retrieved from <https://podaac.jpl.nasa.gov/announcements/2021-08-09-Sea-level-projections-from-the-IPCC-6th-Assessment-Report>
- Ghanbari, M., Arabi, M., Obeysekera, J., & Sweet, W. (2019). A coherent statistical model for coastal flood frequency analysis under nonstationary sea level condition. *Earth's Future*, 162–177. <https://doi.org/10.1029/2018EF001089>
- Goldstein, E. B., Coco, G., & Plant, N. G. (2019). A review of machine learning applications to coastal sediment transport and morphodynamics. *Earth-Science Reviews*. <https://doi.org/10.1016/j.earscirev.2019.04.022>
- Gouldby, B., Méndez, F. J., Guanche, Y., Rueda, A., & Mínguez, R. (2014). A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis. *Coastal Engineering*, 88, 15–26. <https://doi.org/10.1016/j.coastaleng.2014.01.012>
- Guanche, Y., Mínguez, R., & Méndez, F. J. (2014). Autoregressive logistic regression applied to atmospheric circulation patterns. *Climate Dynamics*, 42(1–2), 537–552. <https://doi.org/10.1007/s00382-013-1690-3>
- Hegermiller, C. A., Antolinez, J. A. A., Rueda, A., Camus, P., Perez, J., Erikson, L., & Mendez, F. J. (2017). A multimodal wave spectrum – Based approach for statistical downscaling of local wave climate. *American Meteorological Society*, 47, 375–386. <https://doi.org/10.1175/JPO-D-16-0191.1>
- Hsu, C. H., Olivera, F., & Irish, J. L. (2018). A hurricane surge risk assessment framework using the joint probability method and surge response functions. *Natural Hazards*, 91, 7–28. <https://doi.org/10.1007/s11069-017-3108-8>
- Irish, J. L., Song, Y. K., & Chang, K. A. (2011). Probabilistic hurricane surge forecasting using parameterized surge response functions. *Geophysical Research Letters*, 38(3), 1–5. <https://doi.org/10.1029/2010GL046347>
- James, S. C., Zhang, Y., & O'Donncha, F. (2018). A machine learning framework to forecast wave conditions. *Coastal Engineering*, 137, 1–10. <https://doi.org/10.1016/j.coastaleng.2018.03.004>
- 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>
- Jia, G., Taflanidis, A. A., Nadal-Caraballo, N. C., Melby, J. A., Kennedy, A. B., & Smith, J. M. (2016). Surrogate modeling for peak or time-dependent storm surge prediction over an extended coastal region using an existing database of synthetic storms. *Natural Hazards*, 81(2), 909–938. <https://doi.org/10.1007/s11069-015-2111-1>
- Kim, S. W., Melby, J. A., Nadal-Caraballo, N. C., & Ratcliff, J. (2015). A time-dependent surrogate model for storm surge prediction based on an artificial neural network using high-fidelity synthetic hurricane modeling. *Natural Hazards*, 76(1), 565–585. <https://doi.org/10.1007/s11069-014-1508-6>
- Kopp, R. E., Gilmore, E. A., Little, C. M., Lorenzo-Trueba, J., Ramenzoni, V. C., & Sweet, W. V. (2019). Usable science for managing the risks of sea-level rise. *Earth's Future*, 7, 1235–1269. <https://doi.org/10.1029/2018EF001145>
- Kopp, R. W., Horton, R. M., Little, C. M., Mitrovica, J. X., Oppenheimer, M., Rasmussen, D. J., et al. (2014). Probabilistic 21st and 22nd century sea-level projections at a global network of tide gauge sites. *Earth's Future*, 2, 383–406. <https://doi.org/10.1002/2014ef000239>
- Loepky, J. L., Sacks, J., & Welch, W. J. (2009). Choosing the sample size of a computer experiment: A practical guide. *Technometrics*, 51(4), 366–376. <https://doi.org/10.1198/TECH.2009.08040>
- Ludka, B. C., Guza, R. T., O'Reilly, W. C., & Yates, M. L. (2015). Field evidence of beach profile evolution toward equilibrium. *Journal of Geophysical Research: Oceans*, 120, 7574–7597. <https://doi.org/10.1002/2015JC010893>
- Marcos, M., Rohmer, J., Voudoukas, M., Mentaschi, L., Le Cozannet, G., & Amores, A. (2019). Increased extreme coastal water levels due to the combined action of storm surges and wind-waves. *Geophysical Research Letters*, 1, 2019GL082599. <https://doi.org/10.1029/2019GL082599>
- Marjani, S., Alizadeh-Choobari, O., & Irannejad, P. (2019). Frequency of extreme El Niño and La Niña events under global warming. *Climate Dynamics*, 53, 5799–5813. <https://doi.org/10.1007/s00382-019-04902-1>
- Merrifield, M. A., Thompson, P. R., & Lander, M. (2012). Multidecadal sea level anomalies and trends in the western tropical Pacific. *Geophysical Research Letters*, 39(13), 2–6. <https://doi.org/10.1029/2012GL052032>
- Moftakhar, H. R., AghaKouchak, A., Sanders, B. F., & Matthew, R. A. (2017). Cumulative hazard: The case of nuisance flooding. *Earth's Future*, 5, 214–223. <https://doi.org/10.1002/2016EF000494>
- Moftakhar, H. R., Salvadori, G., AghaKouchak, A., Sanders, B. F., & Matthew, R. A. (2017). Compounding effects of sea level rise and fluvial flooding. *Proceedings of the National Academy of Sciences*, 114(37), 9785–9790. <https://doi.org/10.1073/pnas.1620325114>
- Moftakhar, H. R., Schubert, J. E., AghaKouchak, A., Matthew, R., & Sanders, B. F. (2019). Linking statistical and hydrodynamic modeling for compound flood hazard assessment in tidal channels and estuaries. *Advances in Water Resources*, 128, 28–38. <https://doi.org/10.1016/j.advwatres.2019.04.009>
- Muis, S., Apecechea, M. I., Dullaart, J., de Lima Rego, J., Madsen, K. S., Su, J., & Verlaan, M. (2020). A high-resolution global dataset of extreme sea levels, tides, and storm surges, including future projections. *Frontiers in Marine Science*, 7, 263. <https://doi.org/10.3389/fmars.2020.00263>
- Nerem, R. S., Beckley, B. D., Fasullo, J. T., Hamlington, B. D., Masters, D., & Mitchum, G. T. (2018). Climate-change-driven accelerated sea-level rise detected in the altimeter era. *Proceedings of the National Academy of Sciences of the United States of America*, 115(9), 2022–2025. <https://doi.org/10.1073/pnas.1717312115>
- Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding – A global assessment. *PLoS One*, 10(3). <https://doi.org/10.1371/journal.pone.0118571>
- Nicholls, R., Hanson, S., Lowe, J., Slangen, A., Wahl, T., Hinkel, J., & Long, A. (2021). Integrating new sea-level scenarios into coastal risk and adaptation assessments: An ongoing process. *WIREs Climate Change*. <https://doi.org/10.1002/wcc.706>
- O'Neill, A., Erikson, L., Barnard, P., Limber, P., Vitousek, S., Warrick, J., & Lovering, J. (2018). Projected 21st century coastal flooding in the Southern California Bight. Part 1: Development of the third generation CoSMoS model. *Journal of Marine Science and Engineering*, 6(2), 59. <https://doi.org/10.3390/jmse6020059>
- O'Reilly, W., Olfe, C., Thomas, J., Seymour, R., & Guza, R. (2016). The California coastal wave monitoring and prediction system. *Coastal Engineering*, 116, 118–132. <https://doi.org/10.1016/j.coastaleng.2016.06.005>
- Palmsten, M., & Splinter, K. (2016). Observations and simulations of wave runup during a laboratory dune erosion experiment. *Coastal Engineering*, 115, 58–66. <https://doi.org/10.1016/j.coastaleng.2016.01.007>

- Parker, K., Ruggiero, P., Serafin, K. A., & Hill, D. F. (2019). Emulation as an approach for rapid estuarine modeling. *Coastal Engineering*, 150, 79–93. <https://doi.org/10.1016/j.coastaleng.2019.03.004>
- Pawka, S. (1983). Island shadows in wave directional spectra. *Journal of Geophysical Research: Oceans*, 88(C4), 2579–2591. <https://doi.org/10.1029/JC088iC04p02579>
- Pawlowicz, R., Pawlowicz, R., Beardsley, R. C., Beardsley, R., Lentz, S., & Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T-TIDE. *Computers & Geosciences*, 28(8), 929–937. [https://doi.org/10.1016/s0098-3004\(02\)00013-4](https://doi.org/10.1016/s0098-3004(02)00013-4)
- Perez, J., Menendez, M., & Losada, I. J. (2017). GOW2: A global wave hindcast for coastal applications. *Coastal Engineering*, 124, 1–11. <https://doi.org/10.1016/j.coastaleng.2017.03.005>
- Pullen, T., Liu, Y., Otinari Morillas, P., Wyncoll, D., Malde, S., & Gouldby, B. (2018). Bayonet GPE: A generic and practical overtopping model that includes uncertainty. *Maritime Engineering*, 171(3). <https://doi.org/10.1680/jmaen.2017.31>
- Rasmussen, C. E., & Williams, C. K. (2006). *Gaussian processes for machine learning* (No. 39).
- Ray, R. D., & Foster, G. (2016). Future nuisance flooding at Boston caused by astronomical tides alone. *Earth's Future*, 4(12), 578–587. <https://doi.org/10.1002/2016EF000423>
- Razavi, S., Tolson, B. A., & Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, 48(7). <https://doi.org/10.1029/2011WR011527>
- Resio, D. T., Irish, J., & Cialone, M. (2009). A surge response function approach to coastal hazard assessment – Part 1: Basic concepts. *Natural Hazards*, 51(1), 163–182. <https://doi.org/10.1007/s11069-009-9379-y>
- Roelvink, D., Reniers, A., Dongeren, A. V., Thiel, J. V., Vries, D., McCall, R., & Lescinski, J. (2009). Modelling storm impacts on beaches, dunes and barrier islands. *Coastal Engineering*, 56(11–12), 1133–1152. <https://doi.org/10.1016/j.coastaleng.2009.08.006>
- Rueda, A., Camus, P., Tomás, A., Vitousek, S., & Méndez, F. J. (2016). A multivariate extreme wave and storm surge climate emulator based on weather patterns. *Ocean Modelling*, 104, 242–251. <https://doi.org/10.1016/j.ocemod.2016.06.008>
- Rueda, A., Hegermiller, C. A., Antolinez, J. A. A., Camus, P., Vitousek, S., Ruggiero, P., & Mendez, F. J. (2017). Multiscale climate emulator of multimodal wave spectra: MUSCLE-spectra. *Journal of Geophysical Research: Oceans*(122), 1400–1415. <https://doi.org/10.1002/2016JC011957>
- Sallenger, A. H., Jr. (2000). Storm impact scale for barrier islands. *Journal of Coastal Research*, 16(3), 890–895.
- Santos, V. M., Wahl, T., Long, J. W., Passeri, D. L., & Plant, N. G. (2019). Combining numerical and statistical models to predict storm-induced dune erosion. *Journal of Geophysical Research: Earth Surface*, 124, 1817–1834. <https://doi.org/10.1029/2019JF005016>
- Serafin, K. A., & Ruggiero, P. (2014). Simulating extreme total water levels using time-dependent, extreme value approach. *Journal of Geophysical Research: Oceans*, 119, 6305–6329. <https://doi.org/10.1002/2014JC010093>
- Serafin, K. A., Ruggiero, P., Barnard, P. L., & Stockdon, H. F. (2019). The influence of shelf bathymetry and beach topography on extreme total water levels: Linking large-scale changes of the wave climate to local coastal hazards. *Coastal Engineering*, 150, 1–17. <https://doi.org/10.1016/j.coastaleng.2019.03.012>
- Serafin, K. A., Ruggerio, P., Parker, K., & Hill, D. (2019). What's streamflow got to do with it? A probabilistic simulation of the competing oceanographic and fluvial processes driving extreme along-river water levels. *Natural Hazards and Earth System Sciences*, 19, 1415–1431. <https://doi.org/10.5194/nhess-19-1415-2019>
- Stevenson, S., Wittenberg, A., Fasullo, J., Coats, S., & Otto-Bliesner, B. (2021). Understanding diverse model projections of future extreme El Niño. *Journal of Climate*, 34, 449–464. <https://doi.org/10.1175/JCLI-D-19-0969.1>
- Stockdon, H., Thompson, D., Plant, N., & Long, J. (2014). Evaluation of wave runup predictions from numerical and parametric models. *Coastal Engineering*, 92, 1–11. <https://doi.org/10.1016/j.coastaleng.2014.06.0045>
- Sweet, W. V., Kopp, R. E., Weave, C. P., Obeysekera, J., Horton, R. M., Thieler, E. R., & Zervas, C. (2017). *Global and regional sea level rise scenarios for the United States (Tech. Rep.)*. NOAA Technical Report NOS CO-OPS 083.
- Sweet, W. V., & Park, J. (2014). From the extreme to the mean: Acceleration and tipping points of coastal inundation from sea level rise. *Earth's Future*, 2(12), 579–600. <https://doi.org/10.1002/2014EF000272>
- Tadesse, M., Wahl, T., & Cid, A. (2020). Data-driven modeling of global storm surges. *Frontiers in Marine Science*, 7. <https://doi.org/10.3389/fmars.2020.00260>
- Vousdoukas, M., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L., & Feyen, L. (2018). Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. *Nature Communications*, 9, 2360. <https://doi.org/10.1038/s41467-018-04692-w>
- Wahl, T., & Chambers, D. P. (2015). Evidence for multidecadal variability in US extreme sea level records. *Journal of Geophysical Research: Oceans*, 120(3), 1527–1544. <https://doi.org/10.1002/2014JC010443>
- Wahl, T., & Chambers, D. P. (2016). Climate controls multidecadal variability in U. S. extreme sea level records. *Journal of Geophysical Research: Oceans*, 121, 1274–1290. <https://doi.org/10.1002/2015JC011075>
- Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., & Slanger, A. B. (2017). Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nature Communications*, 8, 1–12. <https://doi.org/10.1038/ncomms16075>
- Warner, J. C., Armstrong, B., He, R., & Zambon, J. B. (2010). Development of a coupled ocean-atmosphere-wave-sediment transport (COAWST) modeling system. *Ocean Modelling*, 35(3), 230–244. <https://doi.org/10.1016/j.ocemod.2010.07.010>
- Wolf, J. (2009). Coastal flooding: Impacts of coupled wave-surge-tide models. *Natural Hazards*, 49(2), 241–260. <https://doi.org/10.1007/s11069-008-9316-5>