Title
Predicting Shoreline Changes Along the California Coast Using Deep Learning Applied to Satellite Observations

Journal
JGR Machine Learning and Computation, 1(e2024JH000172)

Authors
Adusumilli, Susheel
Cirrito, Nicholas
Engeman, Laura
et al.

Publication Date
2024-06-27

DOI
10.1029/2024JH000172

Peer reviewed

JGR Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2024JH000172

Key Points:

- We apply deep learning to satellite-derived shorelines to hindcast changes in beach width using tide and wave model outputs as predictors
- We validate these hindcasts against a large volume of unique in situ observations collected over Southern California beaches
- The deep learning algorithm provides the first satellite-derived estimates of seasonal changes in beach slope, key for flood predictions

Correspondence to:
S. Adusumilli,
suadusum@ucsd.edu

Citation:
Adusumilli, S., Cirrito, N., Engeman, L., Fiedler, J. W., Guza, R. T., Lange, A. M. Z., et al. (2024). Predicting shoreline changes along the California coast using deep learning applied to satellite observations. *Journal of Geophysical Research: Machine Learning and Computation*, *1*, e2024JH000172. https://doi.org/10.1029/2024JH000172

Received 14 FEB 2024
Accepted 14 JUN 2024

Predicting Shoreline Changes Along the California Coast Using Deep Learning Applied to Satellite Observations

Susheel Adusumilli[1], Nicholas Cirrito[1], Laura Engeman[1,2], Julia W. Fiedler[1,3], R. T. Guza[1], Athina M. Z. Lange[1], Mark A. Merrifield[1], William O'Reilly[1], and Adam P. Young[1]

[1]Scripps Institution of Oceanography, UC San Diego, La Jolla, CA, USA, [2]California Sea Grant, UC San Diego, La Jolla, CA, USA, [3]University of Hawaii at Manoa, Honolulu, HI, USA

Abstract Understanding and predicting changes in shoreline location are critical for coastal planners. In situ monitoring is accurate but not widely available. Satellite observations of shorelines have global coverage, but their accuracy and predictive capacity have not been fully explored. Abundant beach surveys and extensive wave observations in Southern California provide a unique ground truth for the interpretation of satellite-derived recently wetted waterlines. We combine 23 years of waterline position estimates from satellite imagery with nearshore wave hindcasts and tides to train and test a deep neural network (DNN). The trained DNN uses only tides and waves as predictors at transects with satellite coverage and wave estimates to predict beach width and, for the first time, seasonal average beach slopes. Beach width changes hindcast using DNN have at least fair skill (>0.3) for 50% of transects, where the skill was calculated relative to the mean of extensive new in situ survey data from in San Diego County. The DNN also predicted shoreline changes to within 10 m (the nominal uncertainty in satellite-derived shorelines) for 64% of transects lacking ground truth. DNN and survey estimates of seasonal beach slope changes also agree qualitatively.

Plain Language Summary Monitoring changes in the width and slope of sandy beaches is crucial for effective coastal management and planning to mitigate risks of coastal erosion and flooding. Survey observations of beach widths using mobile laser scanners or drones are accurate but expensive to collect and unavailable for a large fraction of California's coastline. In this study, we combine satellite observations of changes in shoreline positions and a regional wave model to relate (using machine learning) waves and changes in beach width and slope. We use extensive survey observations in Southern California to validate the machine learning framework. We find that our machine learning framework allows us to make useful predictions of beach width change using only widely available tide and wave model outputs as input for 50% of the beaches considered. We also quantify seasonal changes in beach slope for the first time, important because steeper slopes in winter (combined with large waves) lead to increased flooding.

1. Introduction

Beaches worldwide are increasingly threatened by erosion and flooding associated with high shoreline total water levels (TWL). TWL drivers include sea level, climate and seasonal cycles, tides, ocean eddies, storm surge, and local waves (e.g., Idier et al., 2019). On the U.S. west coast, coastal flooding is often due to coinciding energetic winter swells and high tides (Serafin et al., 2017). Flooding duration, intensity, and frequency are projected to increase over the next century (Sweet et al., 2022; Thompson et al., 2021; Vitousek et al., 2017). The significant uncertainties and impacts of these events are key research areas for coastal management and adaptation (California Climate Change Assessment, Pierce et al., 2018).

The surfzone bathymetry and beach-face slope control the evolution of waves across the swash zone and the resulting shoreline setup and swash (e.g., Battjes, 1974). Forecasts of wave runup, an important component of TWL, are typically made using simplified models that include a specification of the beach-face slope to transform incident waves to runup (Stockdon et al., 2006). The beach-face slope is related to grain size (Bujan et al., 2019), with coarse sand beaches typically steeper than fine sand beaches. Beach-face and surfzone slopes are both key parameters in empirical models of shoreline runup and total swash excursion (Harley et al., 2015; Lange et al., 2022). Conventional survey techniques readily measure the beach-face slope at individual beaches. At large spatial scales (i.e., regional), a linear beach-face slope averaged over the full satellite record can be estimated from satellite imagery (Vos et al., 2020). Here, we develop an alternative approach that extends satellite-based slope estimates from annual to seasonal.

A significant challenge in a satellite-derived shoreline change estimation is transforming the observed recently wetted total water level waterline position, which varies in space and time with the tide, wave runup, and beach face composition to an estimate of the concurrent position of a constant elevation reference contour on the beach face (e.g., the MSL shoreline). Vos et al. (2020) correct for the tide using a mean beach slope but do not attempt to remove the wave signal when estimating a (typically unknown but greater than MSL) shoreline reference contour elevation. The reference contour elevation in this case includes the long-term mean wave runup at the coastal transect. By including skillful wave runup, shoreline change accuracy can be improved at short time scales (Castelle et al., 2021; Graffin et al., 2023; Konstantinou et al., 2023; Vitousek et al., 2023). However, empirical parameterizations for runup corrections can have large errors, and in some cases, corrections increase the error in satellite-estimated shoreline location.

Recent studies have demonstrated the usefulness of machine learning techniques to predict various nearshore processes (e.g., Ellenson et al., 2020; Iglesias et al., 2009; Lopez et al., 2017; Pape et al., 2007) and specifically to predict changes in beach width (Gomez-de la Pena et al., 2023; Montaño et al., 2021; Siegelman et al., 2023; Simmons & Splinter, 2022; Zeinali et al., 2021). While equilibrium (e.g., Yates et al., 2009) and numerical models (e.g., XBeach, CShore) provide insights into the processes driving beach change, machine learning models such as neural networks (Simmons & Splinter, 2022) and random forests (Siegelman et al., 2023) are comparatively assumption-free and have better predictive skills.

Here, we develop a deep neural network (DNN) framework to predict changes in MSL shoreline position and beach slope from wave and tide estimates. Agreement between DNN estimates and extensive (not previously reported) field observations of beach slope and width is promising. Observations of shoreline change and waves are described in Section 2. The DNN model is developed in Section 3 and tested in Section 4.

2. Observations of Shoreline Change and Waves in California

2.1. Changes in Shoreline Positions Estimated Using CoastSat

Publicly available changes in the width of California beaches were derived in Vos et al. (2023a, 2023b) by applying the CoastSat (Vos et al., 2019) algorithm to Landsat 5, 7, 8, and 9 satellites. CoastSat applies supervised machine learning for pixel classification and sub-pixel border segmentation to map the boundary between sand and water. The mapped sand-water interfaces are intersected with cross-shore beach transects spaced at 100 m intervals. Changes in shoreline positions on each transect are available as time series, $X_s(t)$. Shoreline change from these existing satellite data products, spanning from Jan. 2000 to Dec. 2021, was extended to Mar. 2023 by applying the CoastSat algorithm to Landsat 7, 8, and 9 (Figure 1a). The existing and extended time series were similar during 2021 when they overlap, but in the cases where they differed (e.g., due to differences in the quality flags used to remove bad data), they were combined by minimizing the median difference between the two records during the overlap period. Following (Vos et al., 2023a), we removed data points with over 40 m of change between time steps or a modified normalized difference water index (MNDWI, used to find the sand/water threshold) above 0.5.

The accuracy of CoastSat satellite-derived horizontal shoreline positions from comparison with in situ observations was around 10 m at sites with 2–3 m spring tide (Vos, 2023). We find similar accuracy for beaches in California (detailed in Section 4.1). On average, 496 shoreline positions were available for each of 5240 transects, totaling about 2,600,000 positions. Cloud cover led to missing data and differed with location (Figures 2a–2c). Additionally, the sun-synchronous Landsat orbits affect the distribution of sampled tide levels (Figures 2d–2f). The consequences of sparse data during some time periods and tide states on results are discussed below.

We excluded shoreline positions of island beaches and all CoastSat transects >200 m away from the nearest Coastal Data Information Program (CDIP) Monitoring and Prediction (MOP) system wave estimate (O'Reilly et al., 2016), leaving 6,951 transects (86%). The 100 m alongshore spacing of the retrieved CoastSat transects is the same as MOP wave model outputs in Southern California and smaller than the 200 m MOP spacing in Northern California.

10.1029/2024JH000172

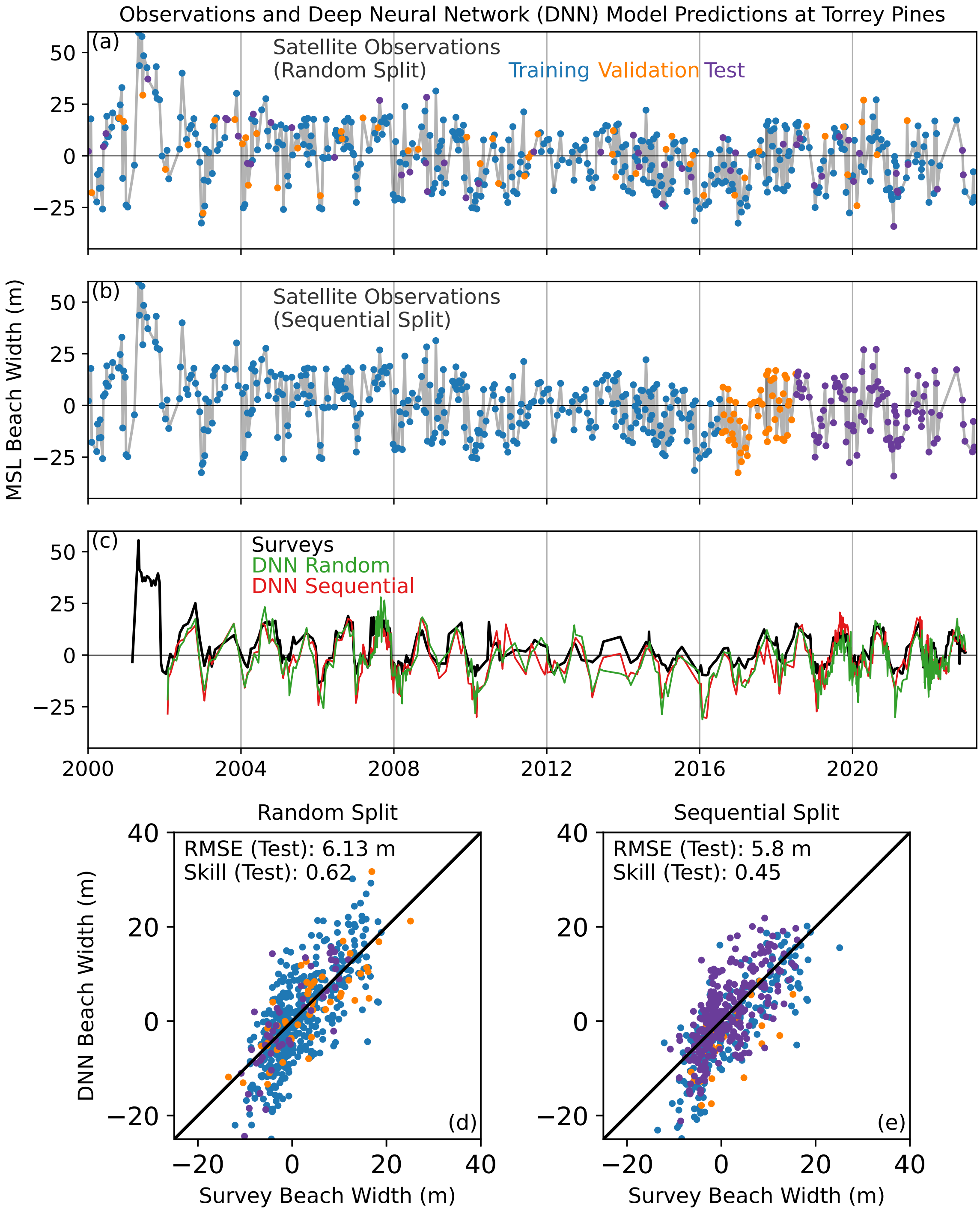

Figure 1. Observations and Deep Neural Network (DNN) model predictions of shoreline change at Torrey Pines State Beach (MOP 582). Satellite observations of mean sea level (MSL) shoreline location (essentially beach width fluctuation) divided into training, validation, and test data. (a) Divided randomly and (b) divided sequentially. The gray curves connect the dots. (c) MSL shoreline location versus time field surveys (black). DNN with random (green) and sequential (red) splits. For (d) random and (e) sequential splits, DNN versus surveyed beach width: training (blue), validation (orange), and test (purple). RMSE and skill are relatively insensitive to sequential or random splitting (compare (d) with (e), skill defined in Equation 5).

2.2. Wave and Tide Model Outputs

Hindcasts of wave spectra (frequencies: 0.04–0.25 Hz) were obtained from the MOP model (O'Reilly et al., 2016). MOP, a non-stationary, linear, spectral refraction wave model, is initialized using offshore deep-water directional wave buoys. Wave properties are estimated at 10-m depth in Southern California and 15-m

10.1029/2024JH000172

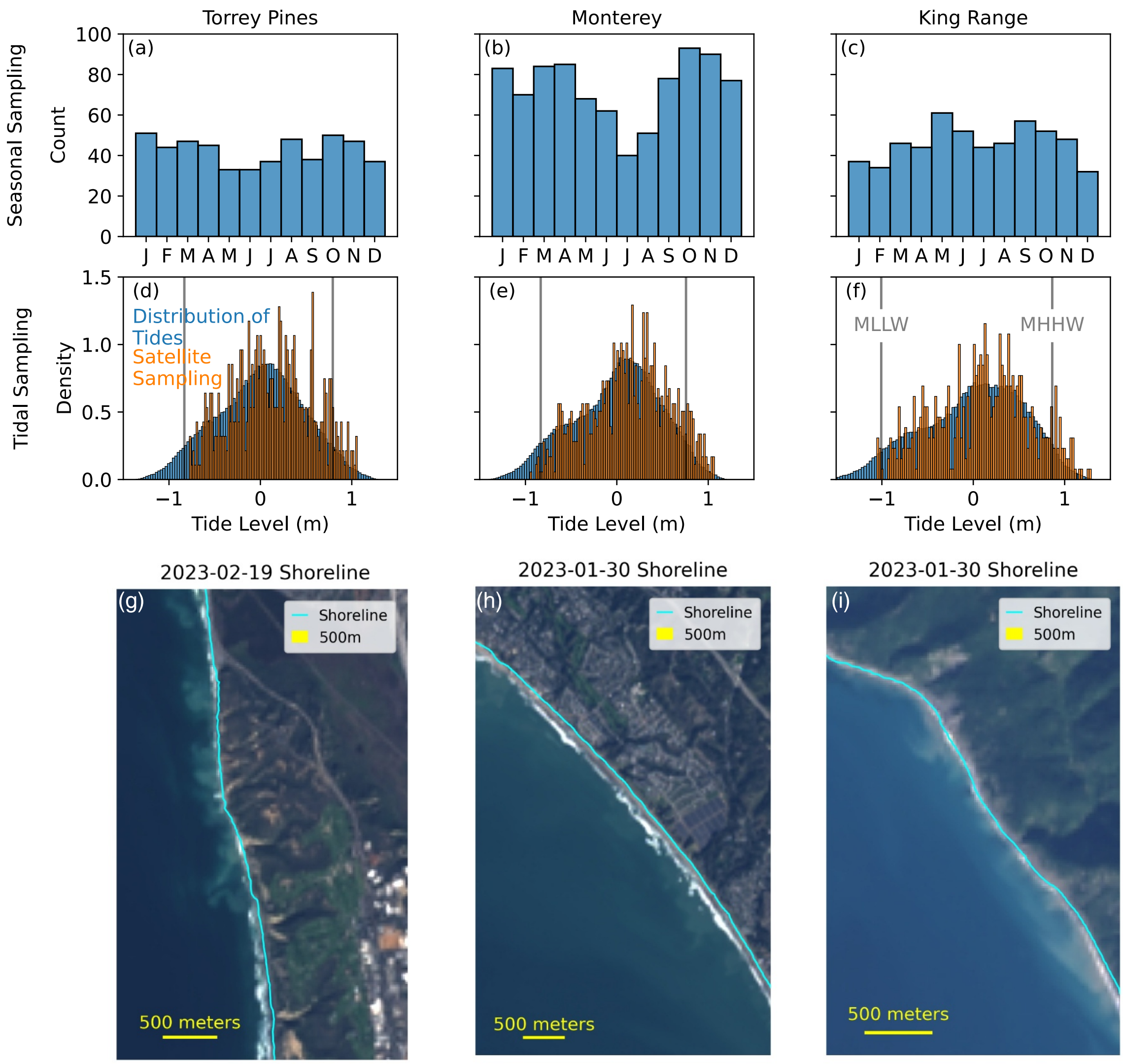

Figure 2. Sampling of satellite data in the tidal cycle for three beaches; Torrey Pines (left, Southern California), Monterey (central California), and King Range (Northern California). (a–c) Monthly frequency of satellite-derived shorelines. Clouds result in fewer shoreline estimates in some months. (d–f) Histograms of tide level for the full tidal distribution (blue) and the satellite sampling (orange). Vertical lines indicate MLLW and MHHW, and 0 is MSL. The lowest tides (<1 m) are almost never sampled due to the sun-synchronous satellite orbit. (g–i) Imagery over the three beaches, with corresponding CoastSat shoreline.

depth in Northern California. MOP outputs are best suited for quantifying mean wave climate conditions as opposed to extreme or individual wave events. The model does not consider wave breaking but includes island shadowing, wave shoaling and refraction, changing shoreline orientation, and offshore shoals. MOP wave model outputs are unavailable pre-2000, and 1 Jan 2000–30 March 2023 is considered here.

Tidal levels are from the FES2014 model (Lyard et al., 2021), following Vos et al. (2019). Monthly sea level anomalies relative to a 1993–2012 reference period derived from using satellite altimetry (Sánchez-Román et al., 2023) are used to include other contributions to still water level. Bilinear interpolation is used to map anomalies to the nearest CoastSat transect. Anomalies are usually <0.2 m and do not substantially change the results.

2.3. In Situ Surveys of Beach Change

Satellite observations of beach change are compared with in situ elevation surveys collected quarterly (or more frequently) over several San Diego County beaches since 2001 using GPS or GNSS equipped platforms along the MOP transects (jet-ski, hand-pushed carts, and all-terrain vehicles, Ludka et al., 2019). Higher-frequency topography observations (monthly, weekly) made with mobile LiDAR systems or drone-based structure-from-motion are available from 2017 onwards (Young et al., 2021, 2023). Surveys from 2017 onwards at sites other than Torrey Pines have not been used in previous satellite validation studies. All data were quality controlled and interpolated into cross-shore MOP system transects at 1 m resolution. A total of 1,073 transects were extracted from in situ observations.

3. Methods

3.1. Transforming Wave Model Outputs to Runup Using the Integrated Power Law

MOP spectral outputs are transformed into estimated runup time series useful for comparisons with the satellite-derived shoreline data with the integrated power-law approximation (IPA) (Fiedler et al., 2020; Lange et al., 2022). IPA uses frequency-weighted integrals of the deepwater (in this case, 10–15 m) sea-swell spectra denoted $E(f)$, where f is the frequency to generate bulk runup statistics using

$$R(m, n, t) = c \int_{\text{deep}} E(f, t)^m f^n df, \tag{1}$$

where the exponents m and n can be varied to derive an ensemble of runup estimates $R(m, n, t)$ and c is a dimensionless constant. Typically, c includes the influence of the beach slope; however, before input into the DNN, the $R(m, n, t)$ are normalized with the constant c. Equation 1 differs from Fiedler et al. (2020) because separate terms for setup, infragravity, and sea-swell components are not estimated, and beach slope is not included explicitly. Instead, DNN parameterizes the role of each $R(m, n, t)$ in the predictions. IPA is an alternative to parameterizations that only incorporate the significant wave height and a wave period (Hunt, 1959; Stockdon et al., 2006) and is intended to improve runup predictions under the bimodal and broad sea-swell distribution common in California (O'Reilly et al., 2016).

3.2. Combining CoastSat and MOP Wave Data With a Deep Neural Network

Tide and wave-driven changes in shoreline positions are large contributors to changes in satellite-driven shoreline positions $X_s(t)$ that should be accounted for to estimate shoreline elevation. A previous study (Vos, 2023) describes techniques for correcting $X_s(t)$ for tides using FES2014 and waves using the Stockdon et al. (2006) parameterization; however, a correction for variability in wave climate using sea-swell spectra has not been explored for beaches in California, where MOP provides near-shore wave climate. In this study, we describe a DNN that uses $X_s(t)$ and MOP waves to estimate seasonal mean slopes averaged over the full 23-year record, while disentangling the influences of tides and waves on shoreline change estimates (Figure 1); a similar separation of these contributors was described in Vitousek et al. (2023) (their Equation 4).

For each CoastSat transect, we developed a unique DNN with the satellite-derived shoreline changes relative to the time-average value from CoastSat as the target variable. The predictors considered for the DNN include tide elevation, monthly changes in still water level estimated from satellite altimetry, and a time series of hourly averaged runup predictors $R(m, n, t)$ (normalized using the constant c in Equation 1). We estimated $R(m, n, t)$ using the IPA, with wave spectra from the nearest MOP transect. We used 5 values of m ranging from 0.2 to 8 and 5 values of n ranging from −4 to 4, resulting in 25 total estimates of $R(t)$ values averaged for the hour in which the satellite data were collected; we refer to this as the "instantaneous" wave forcing. Past wave conditions are included as predictors by averaging $R(m, n, t)$ values over lags of 14 days, 1 month, 3 months, 6 months, 1 year, and 2 years (for a total of 6 lag periods). Siegelman et al. (2023) used similar lagged wave predictors in their machine learning framework. With 25 estimates of $R(t)$ and 6 lag periods for each estimate, a total of 150 lagged wave predictors were included in the DNN.

We use TensorFlow's Keras, an open-source deep learning library, to implement the DNN using the framework shown in Figure 3. The DNN was not pre-trained because this study is the first time these predictors and observations are used for this application. The data, consisting of normalized shoreline position changes for each transect, were divided into training, validation, and test subsets, taking up 70%, 10%, and 20% of the total volume, respectively (Figure 3). These subsets were assigned either randomly (where the data were divided without a specific order) or sequentially (the first 70% for training, the following 10% for validation, and the last 20% for

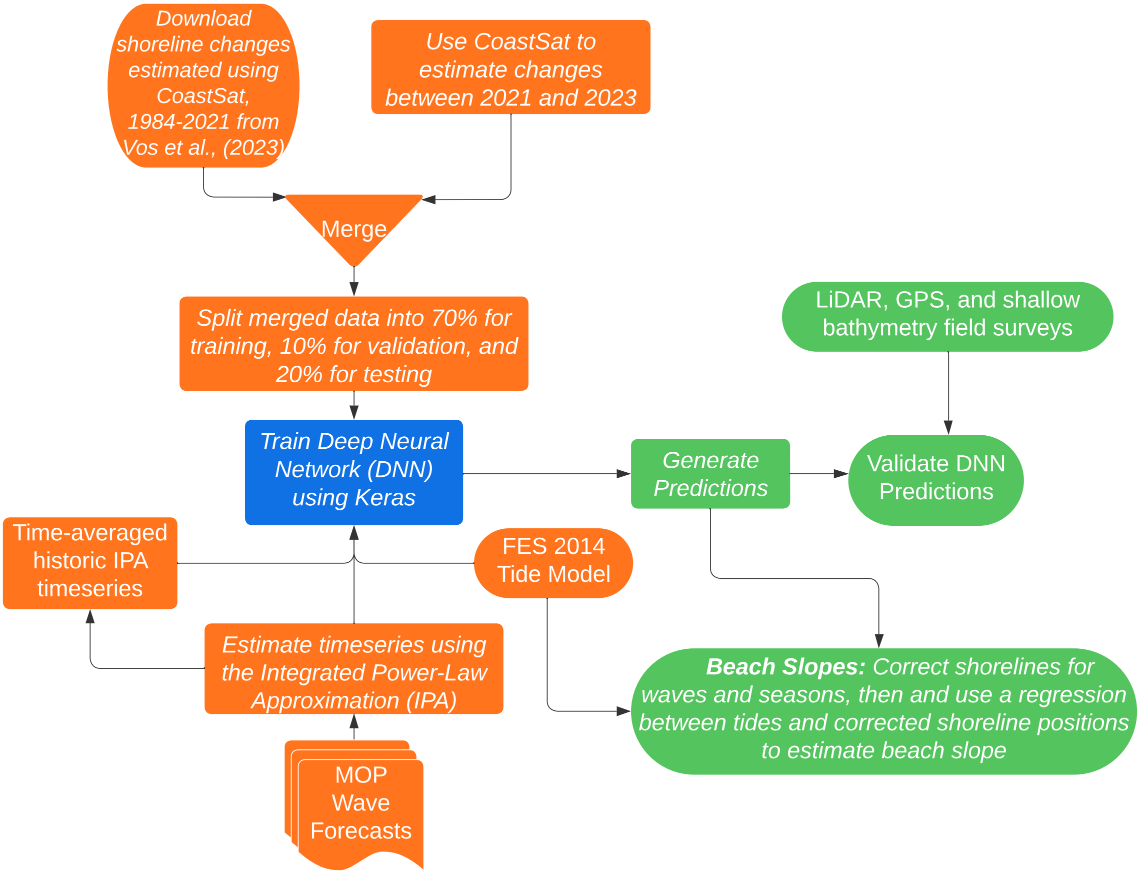

Figure 3. Flowchart of DNN analysis of satellite-derived shoreline change. Orange fields indicate data inputs into the DNN (in blue) either from CoastSat-derived shorelines or MOP Wave forecasts. Green fields indicate post-processing steps, where DNN outputs are compared against survey observations or used to generate beach slopes.

testing) (Figure 1a). Random and sequential sampling often yielded similar results (Figure 1) but there are examples (not shown) where sequential results are biased, for example, by only using training data from a subset of Landsat missions. Random sampling is used for the remainder of this study.

To demonstrate the utility of machine learning for predicting shoreline change, we opted for a simple DNN architecture designed to prevent overfitting. The architecture for this DNN included three layers: the first layer had 128 neurons with a rectified linear unit (ReLU) activation function (Nair & Hinton, 2010) using "He uniform" initialization (He et al., 2015); the second layer contained 256 neurons with the same activation and initialization but included a 50% dropout rate (Srivastava et al., 2014) to mitigate overfitting; the final layer consisted of a single neuron without an activation function. We tested DNN architectures with 3–4 layers with the number of neurons in each layer varying between 32 and 256 and tested dropout rates between 0% and 60%. We found that the 4-layer models were often prone to overfitting (especially in the case with >128 neurons in each layer). Changing the number of neurons between 64 and 256 in the preferred 3-layer model did not show significant differences. We test DNN over all beaches where satellite and survey data are available (86% of the shoreline).

The Adam (Adaptive Moment Estimation) optimization algorithm (Kingma & Ba, 2014), used to fit to the DNN, was slightly less sensitive to changes in the DNN architecture than other algorithms (although the differences were small and within the range of observation uncertainty). We used a mean square error (MSE) as a simple loss function. The training process was run for 1,000 epochs, which refers to one complete pass, with a batch size of 50 data points for each pass. This duration was a tradeoff between adequate learning time, computational efficiency, and model overfitting.

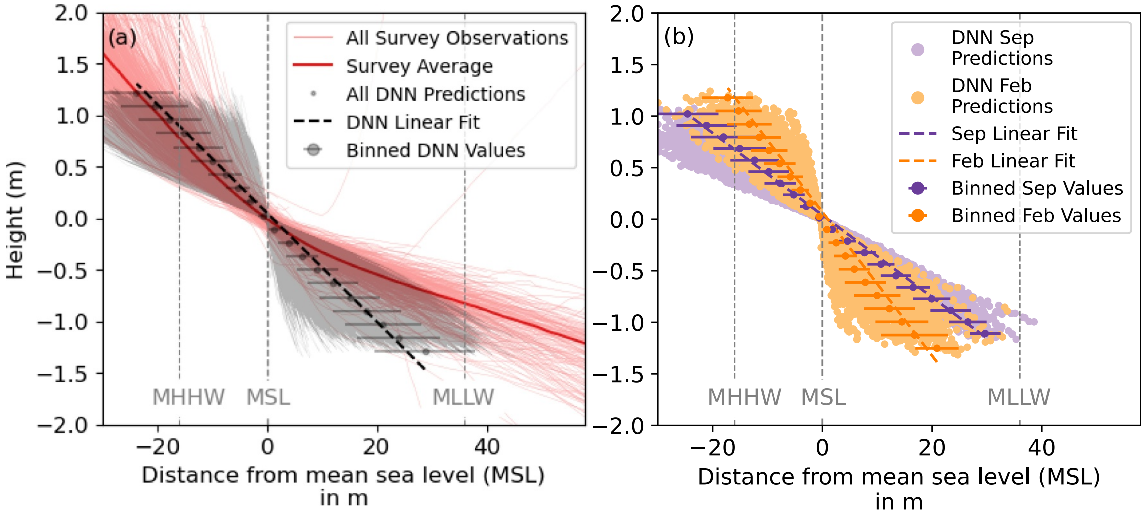

Figure 4. Torrey Pines State Beach (2000–2023, MOP 582) elevation (relative to MSL) versus cross-shore distance from MSL Beach shape from 654 field surveys compared with DNN predictions of beach slope at Torrey Pines State Beach (2000–2023). (a) Beach profiles centered around mean sea level (MSL) from 654 field surveys (red curves average is bold). Hourly DNN predictions (gray) of the tide component of shoreline change, $X_{\text{tide}}(t)$ (Equation 2). The DNN-slope is the slope of the dashed black line, a linear fit to binned averages of DNN-estimated $X_{\text{tide}}(t)$. Panel (b) similar to (a) except DNN beach slope predictions for February (orange, slope 0.08) and September (purple, slope 0.05) are shown separately. Survey shapes are not shown.

3.3. Estimating Beach Slopes Using Deep Neural Networks

We use the tide component of shoreline change, $X_{\text{tide}}(t)$, to estimate a linear beach slope. Hourly DNN predictions of shoreline change are denoted as $X_{\text{DNN}}(t)$. A second set of shoreline change predictions, denoted $X_{\text{msl}}(t)$, were generated using a constant mean sea level (MSL) value for the tide (but the wave-driven contribution was allowed to vary as before). $X_{\text{tide}}(t)$ was estimated as

$$X_{\text{tide}}(t) = X_{\text{DNN}}(t) - X_{\text{msl}}(t) \tag{2}$$

$X_{\text{tide}}(t)$ was binned in 20 discrete tide height intervals and errors in the binned values estimated as the standard deviation of heights within the bins (Figure 4). The beach slope is estimated as the regression slope with the binned $X_{\text{tide}}(t)$ values as the x-variable and tide heights as the y-variable. Instead of simple linear regression, orthogonal distance regression was used to account for errors in the x-variable (Figure 4a). The analysis is repeated separately for each month (Figure 4b) and a 3-month moving average filter was applied to derive seasonal changes. The relatively few (<3%) beaches with estimated slopes >0.2 were excluded consistent with Vos et al. (2020) and the upper limit of sandy beach-face slopes (Bujan et al., 2019).

4. Results

4.1. Validating Estimated Tide and Wave Contributions to Shoreline Change

The ability of the DNN framework to predict changes in beach widths is illustrated at Torrey Pines State Beach (TPSB; MOP Transect 582, Figure 1). Results are similar with CoastSat data divided into training, validation, and test subsets both randomly (Figure 1a) and sequentially (Figure 1b).

Shorelines corrected using the DNN for tide and wave contributions are compared to corrections derived by multiplying beach slopes from Vos et al. (2020) with tides and several published parameterizations for wave runup (such as Stockdon et al., 2006). Corrected shorelines are denoted $X_{s,\text{corr}}(t)$. $X_{s,\text{corr}}(t)$ and survey observations are compared at 933 transects using a MSE metric normalized by the standard deviation (STD) of the surveyed MSL beach width $X_{\text{survey}}(t)$,

$$S_{\text{survey}} = \frac{\text{MSE}(X_{\text{survey}}(t) - X_{s,\text{corr}}(t))}{\text{MSE}(X_{\text{survey}}(t) - X_s(t))}. \tag{3}$$

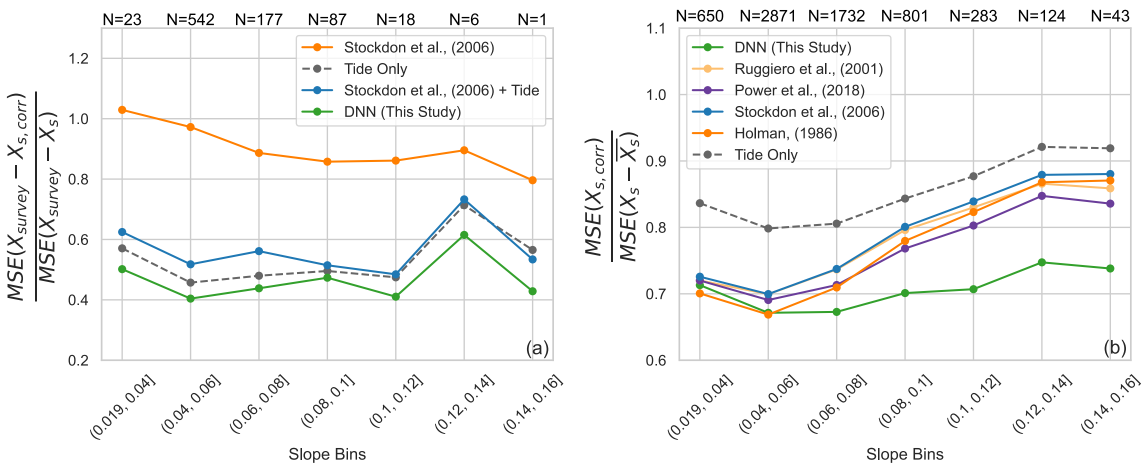

Figure 5. Comparison between beach width modeled using deep neural networks (DNN) and estimates using other published parameterizations in a standard (e.g., Vos, 2023) approach. Transects are binned by the average beach slope value for 2020–2023. Total number of beaches transects in each slope bin (N) is shown at the top of (a) selected Southern California transects and (b) all California transects. Note different vertical scales. (a) Normalized MSE differences S_{survey} (Equation 3) between beach width at mean sea level from surveys (X_{survey}) and satellites with various corrections ($X_{s,\text{corr}}$) for all Southern California transects with ground truth ($N_{\text{total}} = 854$). Shown are corrections for waves only (Stockdon et al., 2006; orange line), tides only (black dashed line), tides and waves using Stockdon et al. (2006) (blue line), and tides and waves using the DNN framework (green line). (b) Normalized MSE differences between satellite-derived shoreline changes uncorrected and corrected for various wave forcing parameterizations ($X_{s,\text{corr}}$) and DNN for all California transects ($N_{\text{total}} = 6{,}504$) (see Equation 4). Corrections use standard tide and wave effects (R2% from Beuzen et al., 2019; Holman, 1986; Power et al., 2019; Ruggiero et al., 2001) (see legend).

S_{survey} normalizes the error in $X_{s,\text{corr}}$ by the uncorrected error.

The DNN-based correction performs better at all beach slopes than corrections that only account for tides and corrections that account for both tides and wave climate through the Stockdon et al. (2006) parameterization (Figure 5a). These corrections are applied by adding tide and wave contributions to total water level with beach slopes derived from Vos et al. (2020); these results do not change when we use beach slopes derived from this study. Figure 5a shows about 5%–10% error reduction for the DNN-based correction compared with tide only. The correction based on the tides and the Stockdon et al. (2006) parameterization counter-intuitively increases the error (relative to tide only) by around 10% for low slope beaches. This suggests that the Stockdon et al. (2006) parameterization by itself is inaccurate at low beach slopes.

Corrections using various parameterizations (Figure 5b), including one machine learning approach by Power et al. (2019), are compared using fractional MSE differences between corrected and uncorrected satellite-derived shorelines

$$M_s = \frac{\text{MSE}(X_{s,\text{corr}}(t))}{\text{MSE}\left(X_s(t) - \overline{X_s(t)}\right)}, \tag{4}$$

where $\overline{X_s(t)}$ is the average. These parameterizations are not designed to account for changes in satellite-derived shorelines and are outperformed by DNN, especially at high beach slopes where M_s for the DNN was lower by over 10% (Figure 5b).

4.2. Validating Predictions of Shoreline Change

The DNN predicts shoreline change by incorporating historical wave observations and is a reliable predictor of shoreline changes for MOP 582 at Torrey Pines (Figure 1c). Beach width changes relative to the time-average over the record for each transect were estimated by removing the instantaneous wave, tide, and sea level contributions from the DNN predictions of shoreline change. DNN predictions are compared with independent survey data (Figures 1d and 1e) using a skill-like metric (alternative normalizations are possible)

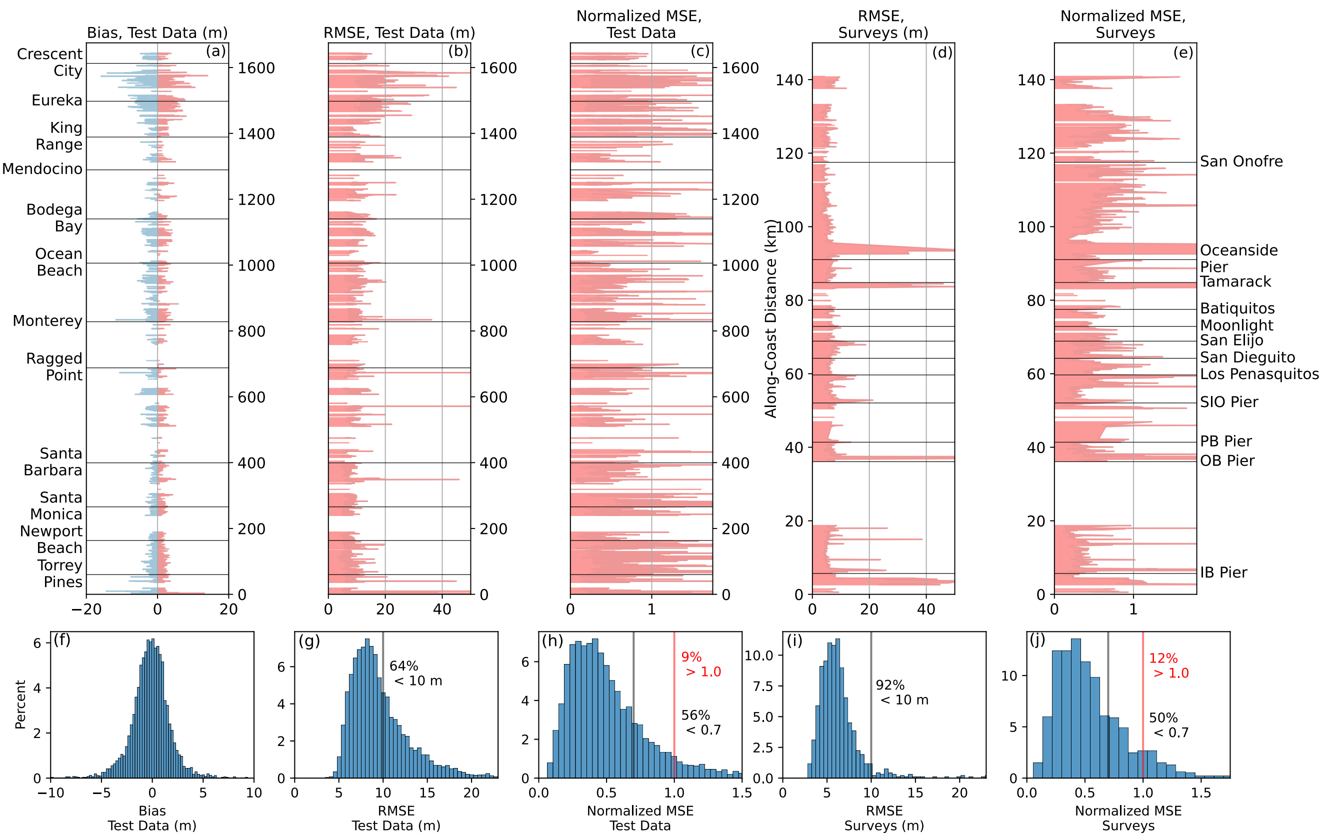

Figure 6. Testing the DNN model. (a, f) Bias (mean of the residuals) between test data from satellite observations and DNN predictions. The small bias (usually less than 3 cm) is noisy and changes sign at small (100 m) alongshore scales. (b, g) Root mean square error (RMSE) between test data from satellite observations and DNN model predictions. (c, h) Normalized MSE (see Equation 4). 56% of transects have normalized MSE < 0.7 (skill > 0.3) on test data. (d, i) RMSE of residuals between data from DNN model predictions and surveys (available between 0 and 145 km). (e, j) Normalized MSE version of (d, i). 49.7% of transects have MSE < 0.7 (skill > 0.3) with observations. (f–j) Histograms of changes shown in (a–e).

$$\text{skill} = 1 - \frac{\text{MSE}(X_{\text{DNN}}(t) - X_{\text{survey}}(t))}{\text{MSE}\left(X_{\text{DNN}}(t) - \overline{X_{\text{DNN}}(t)}\right)}, \tag{5}$$

where $\overline{X_{\text{DNN}}(t)}$ is the average.

DNN predictions are compared with California-wide satellite (Figures 6a–6c) and Southern California survey data (Figures 6d and 6e). The bias observed between the test data, which constituted satellite-derived shoreline change estimates not employed in the DNN's training, and the DNN predictions consistently fell within ±5 m (Figure 6a). Root mean square error (RMSE) values between the test data and DNN predictions typically remained below 20 m (Figure 6b). The median RMSE value between DNN predictions and beach widths from independent survey data in Southern California was 6 m (Figure 6d). These RMSE values between DNN predictions and survey beach widths were similar to the values between DNN predictors and test data, because the variability in beach width in Southern California is typically smaller than the variability in central and Northern California. For 8.7% of transects, normalized MSE > 1, skill < 0 (Figures 6c and 6h) indicating DNN is inaccurate. These DNNs fail internal tests and were excluded from comparison with in situ surveys and future analyses. For 44% of test data, normalized MSE values were >0.7 (skill < 0.3), which suggested poor skill for these DNNs; however, we included these DNNs in future analyses so that they could be independently verified against in situ data.

4.3. Beach Slope Predictions

The portion of the beach sampled by the satellite-derived shorelines is strongly dependent on the tidal level during the satellite acquisitions and influences the beach slopes determined using the DNN. In California, the portion of

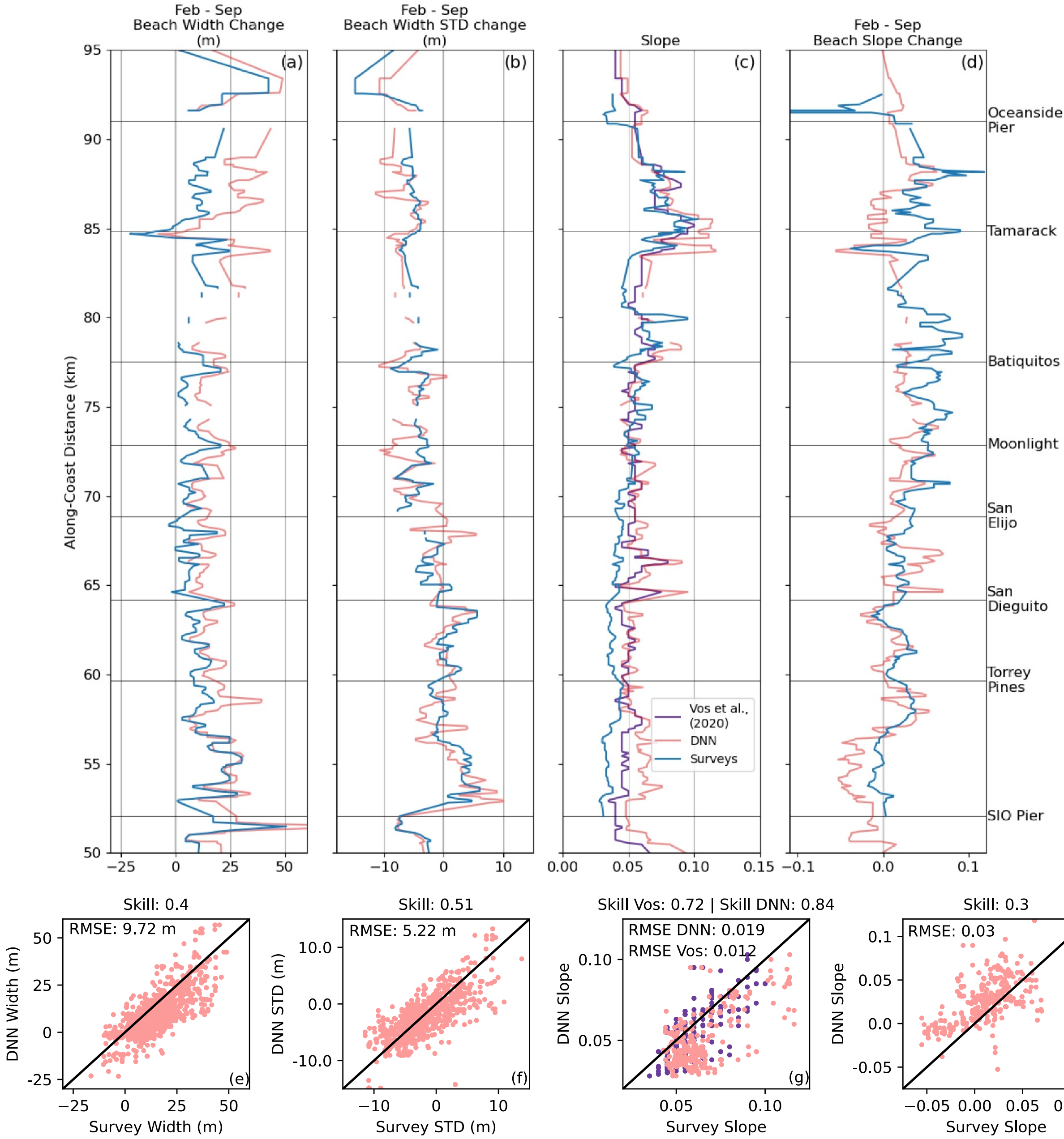

Figure 7. Validation of changes in beach width and slope for the relatively well-surveyed 50 km span between SIO pier and Oceanside Pier. DNN (pink), in situ surveys (blue), (a) Beach width change between Feb and Sep (b) Feb–Sep beach width change standard deviation, (c) average (2000–2023) beach slope, (d) slope change Feb–Sep. in (c) Vos et al. (2020) (blue). (e–h) Scatter plots of panels (a–d). RMSE and skill are given in each panel. DNNs with normalized MSE > 1 excluded (Figure 6c), then data were smoothed alongshore with a 300 m moving median filter.

the beach between MSL and MHHW is captured significantly better by the Landsat satellites than the portion below MSL (Figures 2d–2f). Anomalies of $X_{\mathrm{tide}}(t)$ estimated using the DNN are therefore more accurate above MSL than below it. This is evident in comparisons of $X_{\mathrm{tide}}(t)$ against survey observations of beach shape centered at their MSL locations (Figure 4a) at TPSB, which shows that $X_{\mathrm{tide}}(t)$ is more accurate at high tides than at low tides. Therefore, in this analysis, we restrict the calculation of beach slopes to the region between MSL and MHHW.

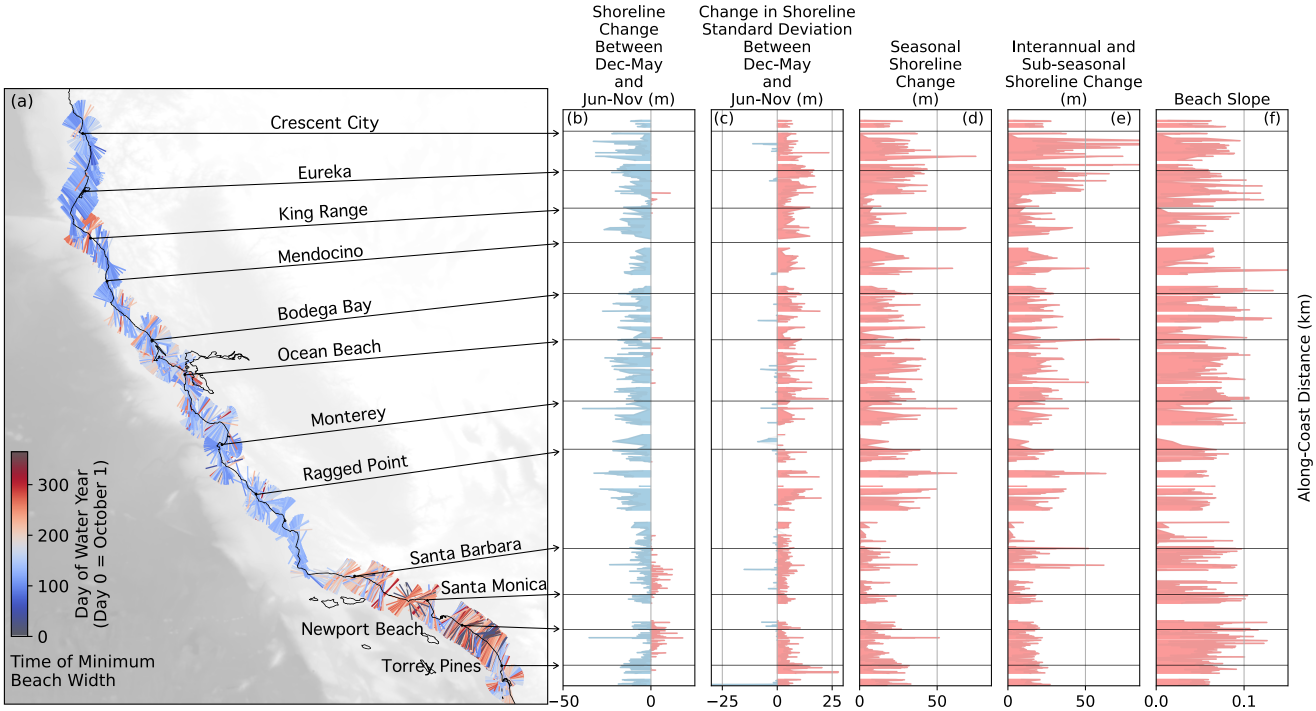

Figure 8. Contributors to shoreline change in California. (a) Time (as day of water year, starting October 1) of minimum beach width. Line directions show beach orientation. (b) Change in shoreline position between winter average (December–May) and summer average (June–November). Negative values indicate wider summer than winter beaches. The alongshore variation of (a) time of minimum and (b) winter minus summer are similar, with strong variation south of Santa Barbara. (c) Change in standard deviation of shoreline positions between winter and summer. Positive values indicate a shoreline is more variable in winter than summer. (d) Seasonal amplitude of shoreline change. (e) Combined amplitude of sub-seasonal and interannual components of shoreline change. (f) Average beach slope (2000–2023) from DNN.

Most (90%) California beach slopes derived using the DNN-based technique were between 0.03 and 0.1 (Figure 8f). Slope estimates are similar to Vos et al. (2020), with a mean and standard deviation difference of 0.01 ± 0.02. For Southern California, RMSE slope errors compared to surveys are larger for DNN than Vos et al. (2020) (Figures 7c and 7g) (RMSE = 0.012 for Vos et al. (2020) and 0.019 for DNN); however, unlike the technique described in Vos et al. (2020), the DNN can also estimate the seasonal change in beach slope (Figure 7d). At TPSB, $X_{\mathrm{tide}}(t)$ values during February are steeper and show less horizontal variability than during a fall month of September (Figure 2b). Similarly, we find that beach slopes in September are typically lower than those in February for other beaches in Southern California (Figure 7d). Slope change data show moderate skill (0.3) and similar patterns of variability compared with survey data.

4.4. Characteristics of Shoreline Changes in California

Shoreline change data from CoastSat, with corrections for instantaneous tide and wave contributors estimated using the DNN, provide large-scale insights of beach change in California at a variety of timescales. At seasonal timescales, average shoreline positions are more landward during winters (Dec–May) and more seaward during summers (Jun–Nov) (Figures 8a and 8b), and the standard deviation of shoreline positions (indicative of variability in beach width) is correspondingly higher during winters compared to summers (Figure 8c). The amplitude of the seasonal variations is higher for beaches further north (Figure 8d). There are some seasonal biases in the availability of satellite data across different regions in California due to changes in cloud cover (Figures 2a–2c), but these biases do not affect these results of seasonal variability averaged over two decades. Sub-seasonal and interannual variations in shoreline positions are also higher in the northern beaches compared to the south (Figure 8e).

5. Discussion

The DNN framework extracts useful physically relevant information from satellite-derived shorelines. Previous studies (e.g., Castelle et al., 2021; Konstantinou et al., 2023) highlight the need for corrections of satellite-derived shorelines for wave contributions. Konstantinou et al. (2023) stress the need for site-specific corrections. Our DNN model achieves exactly that by establishing unique transformations of tide heights and wave spectra for each beach to obtain optimal corrections. We use DNNs because they are commonly used, easy to implement, and flexible with the capability to capture the nonlinear interactions between waves, tides, and beach morphology. The simplicity allows testing DNN predictive ability over the large spatial reach with high computational efficiency. We note that other similar machine learning techniques show promise, such as random forests (Siegelman et al., 2023), applied to two transects in California, and convolutional neural networks (Gomez-de la Peña et al., 2023), applied to one beach in New Zealand. Simple long-short-term memory, feedforward neural networks, and convolutional neural networks were applied to satellite-derived shorelines for a large number of beaches distributed across the globe in Calkoen et al. (2021). The results from the machine learning techniques used in all three of these studies did not differ significantly from the DNN but were an improvement over "traditional" time series prediction techniques, such as ordinary least squares and autoregressive models.

Estimates of shoreline change corrected for tide and wave contributions using DNN showed better agreement (by around 10%) with survey observations than corrections applied using previous published techniques (Figure 5a), implying that DNN-corrected shorelines represent improved estimators of changes in physical beach width. We note, however, that these techniques were originally designed to provide accurate runup estimates rather than to correct satellite-derived shorelines. Correcting the tide and wave contributions to shoreline change also helps identify physical drivers of beach change; for example, some south-facing beaches in Southern California (Figure 8a) show reversed seasonal behavior because southern swell is maximum in summer whereas northerly waves are maximum in winter.

The low RMSE values between DNN predictions of beach width and survey observations (<15 m in Southern California) suggest that machine learning is a valuable tool for estimating the response of beach width to wave forcing. There is close correspondence between survey observations at TPSB and DNN predictions from 2019 onwards (Figure 1c), a period during which the satellite-derived shoreline change data were not used to train the model in the sequential split case. Areas with low DNN skill (where normalized MSE > 0.7, Figures 6c and 6e) highlight regions where wave, tides, and sea level anomalies alone were not reliable predictors of beach width. These areas may be influenced by episodic alongshore transport or beach nourishment (not included in this DNN) or are subjected to higher satellite shoreline retrieval errors (e.g., cliff shadowing and seepage faces).

Future DNNs could include the role of alongshore transport (e.g., Vitousek et al., 2023) and nourishments. S_{xy}, the shear component of the radiation stress can be used as a predictor to account for alongshore sediment transport. Graph neural networks (e.g., Corso et al., 2024) can be used to model changes in connected transects simultaneously. As expected, DNN performance is severely degraded by beach nourishments. Figure 9 shows poor (negative) skills for an Imperial Beach transect (in Southern California) nourished in September 2012 (Ludka et al., 2018). DNN model results are also degraded by nourishments at Cardiff State Beach and Newport Beach (not shown). Future work could simulate nourishments as step functions in space and time.

Beach slopes in California generally vary seasonally, with winter higher slopes often contributing to increased flooding (Figure 7d). Estimated beach slopes showed lower accuracy at lower elevations (Figure 1a) due to a lack of sampling during low tide (Figure 2a). Slopes estimated with DNN are complementary to a frequency-based technique (Vos et al. (2020) and extend slope estimates from long-term annual averages to long-term seasonal averages. Higher spatial and temporal resolution (than Landsat) data are becoming available and will help fill gaps and improve shoreline and beach slope estimates.

6. Summary

We combined 23 years of waterline position estimates from satellite imagery with nearshore wave hindcast and runup models in a flexible deep neural network framework to estimate seasonal changes in beach width and slope. The DNNs provide improved estimates of the tide and wave contributions to instantaneous shoreline change

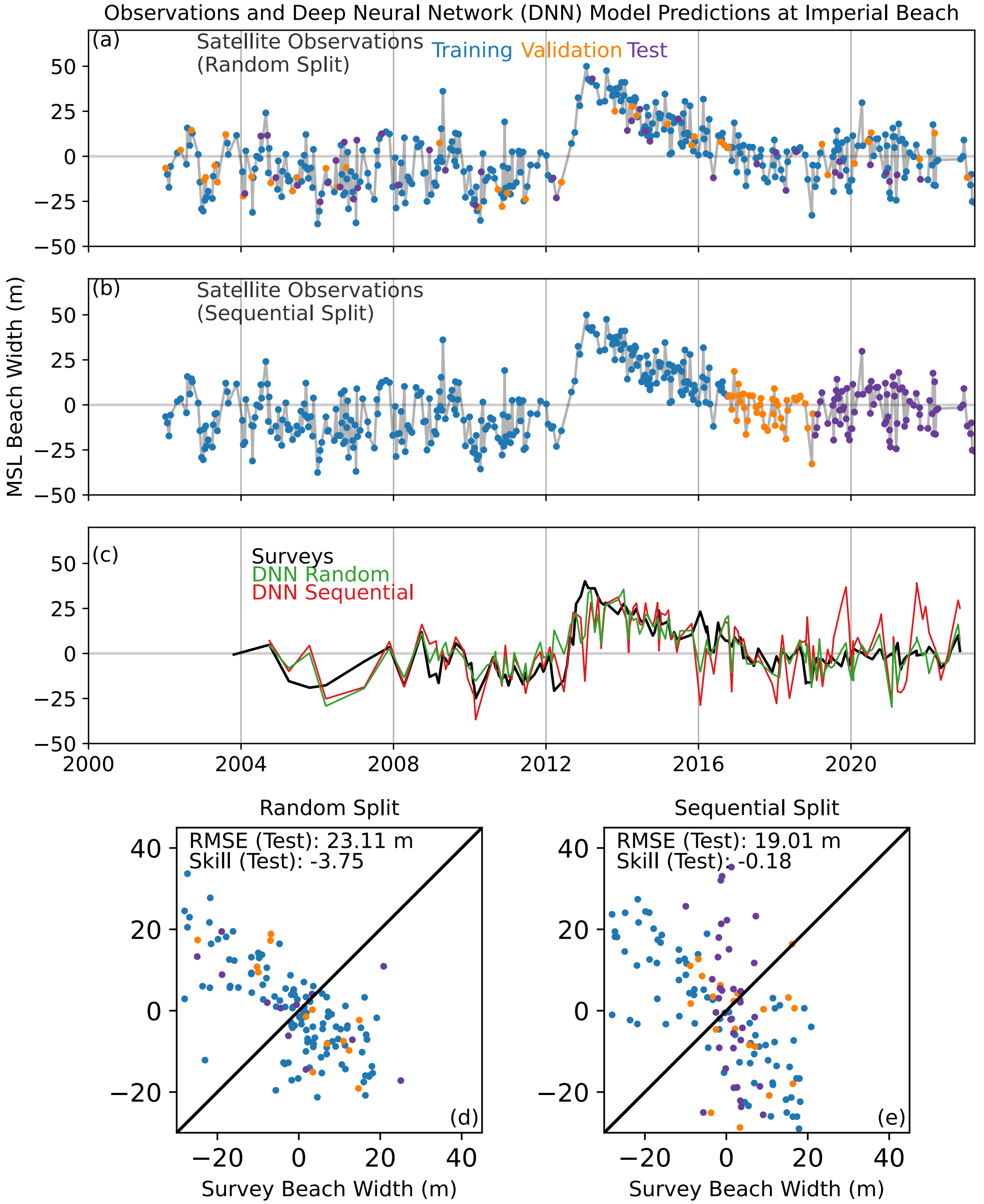

Figure 9. Observations and Deep Neural Network (DNN) model predictions of shoreline change at Imperial Beach (MOP 30), nourished in 2012. Satellite observations of mean sea level (MSL) shoreline location divided into training, validation, and test data. (a) Divided randomly and (b) divided sequentially. The gray curves connect the dots. (c) MSL shoreline location versus time field surveys (black). DNN with random (green) and sequential (red) splits. For (d) random and (e) sequential splits, DNN versus surveyed beach width: training (blue), validation (orange), and test (purple). Performance is poor (negative Test skills) compared with Torrey Pines (Figure 1) not nourished during the survey period.

compared to previous parameterizations, and DNN outputs agree with survey-derived beach widths with about 6 m accuracy (Figures 6d and 6i). Tide and wave model outputs can be used to predict seasonal beach widths and beach slopes (Figure 8) for over half of California's beaches with fair skill (normalized MSE < 0.7). These data could be used in TWL flooding estimates at these beaches.

Data Availability Statement

Data processed between 2000 and 2021 using CoastSat are available at Vos (2023). Data from Jan 2021 to Mar 2023 were processed using publicly available CoastSat code (Vos et al., 2019). Code to extract tide and wave model outputs and generate DNN models is available at Adusumilli (2024). The FES2014 and Sea Level Anomaly data products are at Aviso through Lyard et al. (2021) and Sánchez-Román et al. (2023), respectively.

Acknowledgments
This study was funded by the U.S. Army Corps of Engineers (W912HZ1920020) and the California Department of Parks and Recreation, Natural Resources Division Oceanography Program (C19E0026). In situ field surveys were collected and processed by the Coastal Processes Group field team.

References

Adusumilli, S. (2024). *iceisnice/coastsat-DNN: CoastSat DNN California (CoastSat-DNN-v0)*. Zenodo. https://doi.org/10.5281/zenodo.11245272

Battjes, J. A. (1974). Surf similarity. *Coastal Engineering*, *1974*, 466–480.

Beuzen, T., Goldstein, E. B., & Splinter, K. D. (2019). Ensemble models from machine learning: An example of wave runup and coastal dune erosion. *Natural Hazards and Earth System Sciences*, *19*(10), 2295–2309. https://doi.org/10.5194/nhess-19-2295-2019

Bujan, N., Cox, R., & Masselink, G. (2019). From fine sand to boulders: Examining the relationship between beach-face slope and sediment size. *Marine Geology*, *417*, 106012. https://doi.org/10.1016/j.margeo.2019.106012

Calkoen, F., Luijendijk, A., Rivero, C. R., Kras, E., & Baart, F. (2021). Traditional vs. machine-learning methods for forecasting sandy shoreline evolution using historic satellite-derived shorelines. *Remote Sensing*, *13*(5), 934. https://doi.org/10.3390/rs13050934

Castelle, B., Masselink, G., Scott, T., Stokes, C., Konstantinou, A., Marieu, V., & Bujan, S. (2021). Satellite-derived shoreline detection at a high-energy meso-macrotidal beach. *Geomorphology*, *383*, 107707. https://doi.org/10.1016/j.geomorph.2021.107707

Corso, G., Stark, H., Jegelka, S., Jaakkola, T., & Barzilay, R. (2024). Graph neural networks. *Nature Reviews Methods Primers*, *4*(1), 17. https://doi.org/10.1038/s43586-024-00294-7

Ellenson, A. N., Simmons, J. A., Wilson, G. W., Hesser, T. J., & Splinter, K. D. (2020). Beach state recognition using argus imagery and convolutional neural networks. *Remote Sensing*, *12*(23), 3953. https://doi.org/10.3390/rs12233953

Fiedler, J. W., Young, A. P., Ludka, B. C., O'Reilly, W. C., Henderson, C., Merrifield, M. A., & Guza, R. T. (2020). Predicting site-specific storm wave run-up. *Natural Hazards*, *104*(1), 493–517. https://doi.org/10.1007/s11069-020-04178-3

Gomez-de la Peña, E., Coco, G., Whittaker, C., & Montaño, J. (2023). On the use of Convolutional Deep Learning to predict shoreline change. *Earth Surface Dynamics*, *11*(6), 1145–1160. https://doi.org/10.5194/esurf-11-1145-2023

Graffin, M., Taherkhani, M., Leung, M., Vitousek, S., Kaminsky, G., & Ruggiero, P. (2023). Monitoring interdecadal coastal change along dissipative beaches via satellite imagery at regional scale. *Cambridge Prisms: Coastal Futures*, *1*, e42. https://doi.org/10.1017/cft.2023.30

Harley, M. D., Turner, I. L., & Short, A. D. (2015). New insights into embayed beach rotation: The importance of wave exposure and cross-shore processes. *Journal of Geophysical Research: Earth Surface*, *120*(8), 1470–1484. https://doi.org/10.1002/2014JF003390

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition (Vol. *7*).

Holman, R. A. (1986). Extreme value statistics for wave run-up on a natural beach. *Coastal Engineering*, *9*(6), 527–544. https://doi.org/10.1016/0378-3839(86)90002-5

Hunt, I. A., Jr. (1959). Design of seawalls and breakwaters. *Journal of the Waterways and Harbors Division*, *85*(3), 123–152. https://doi.org/10.1061/jwheau.0000129

Idier, D., Bertin, X., Thompson, P., & Pickering, M. D. (2019). Interactions between mean sea level, tide, surge, waves and flooding: Mechanisms and contributions to sea level variations at the coast. *Surveys in Geophysics*, *40*(6), 1603–1630. https://doi.org/10.1007/s10712-019-09549-5

Iglesias, G., López, I., Castro, A., & Carballo, R. (2009). Neural network modelling of planform geometry of headland-bay beaches. *Geomorphology*, *103*(4), 577–587. https://doi.org/10.1016/j.geomorph.2008.08.002

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *CoRR*. https://www.semanticscholar.org/paper/Adam%3A-A-Method-for-Stochastic-Optimization-Kingma-Ba/a6cb366736791bcccc5c8639de5a8f9636bf87e8

Konstantinou, A., Scott, T., Masselink, G., Stokes, K., Conley, D., & Castelle, B. (2023). Satellite-based shoreline detection along high-energy macrotidal coasts and influence of beach state. *Marine Geology*, *462*, 107082. https://doi.org/10.1016/j.margeo.2023.107082

Lange, A. M. Z., Fiedler, J. W., Becker, J. M., Merrifield, M. A., & Guza, R. T. (2022). Estimating runup with limited bathymetry. *Coastal Engineering*, *172*, 104055. https://doi.org/10.1016/j.coastaleng.2021.104055

López, I., Aragonés, L., Villacampa, Y., & Serra, J. C. (2017). Neural network for determining the characteristic points of the bars. *Ocean Engineering*, *136*, 141–151. https://doi.org/10.1016/j.oceaneng.2017.03.033

Ludka, B. C., Guza, R. T., & O'Reilly, W. C. (2018). Nourishment evolution and impacts at four southern California beaches: A sand volume analysis. *Coastal Engineering*, *136*, 96–105. https://doi.org/10.1016/j.coastaleng.2018.02.003

Ludka, B. C., Guza, R. T., O'Reilly, W. C., Merrifield, M. A., Flick, R. E., Bak, A. S., et al. (2019). Sixteen years of bathymetry and waves at San Diego beaches. *Scientific Data*, *6*(1), 1. https://doi.org/10.1038/s41597-019-0167-6

Lyard, F. H., Allain, D. J., Cancet, M., Carrère, L., & Picot, N. (2021). FES2014 global ocean tide atlas: Design and performance. *Ocean Science*, *17*(3), 615–649. https://doi.org/10.5194/os-17-615-2021

Montaño, J., Coco, G., Cagigal, L., Mendez, F., Rueda, A., Bryan, K. R., & Harley, M. D. (2021). A multiscale approach to shoreline prediction. *Geophysical Research Letters*, *48*(1), e2020GL090587. https://doi.org/10.1029/2020gl090587

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th international conference on international conference on machine learning* (pp. 807–814).

O'Reilly, W. C., Olfe, C. B., Thomas, J., Seymour, R. J., & Guza, R. T. (2016). The California coastal wave monitoring and prediction system. *Coastal Engineering*, *116*, 118–132. https://doi.org/10.1016/j.coastaleng.2016.06.005

Pape, L., Ruessink, B. G., Wiering, M. A., & Turner, I. L. (2007). Recurrent neural network modeling of nearshore sandbar behavior. *Neural Networks*, *20*(4), 509–518. https://doi.org/10.1016/j.neunet.2007.04.007

Pierce, D. W., Kalansky, J. F., & Cayan, D. R. (2018). *Climate, drought, and sea level rise scenarios for California's fourth climate change assessment*. California Energy Commission and California Natural Resources Agency. Retrieved from https://cawaterlibrary.net/wp-content/uploads/2019/08/Pages-from-Projections_CCCA4-CEC-2018-006.pdf

Power, H. E., Gharabaghi, B., Bonakdari, H., Robertson, B., Atkinson, A. L., & Baldock, T. E. (2019). Prediction of wave runup on beaches using Gene-Expression Programming and empirical relationships. *Coastal Engineering*, *144*, 47–61. https://doi.org/10.1016/j.coastaleng.2018.10.006

Ruggiero, P., Komar, P. D., McDougal, W. G., Marra, J. J., & Beach, R. A. (2001). Wave runup, extreme water levels and the erosion of properties backing beaches. *Journal of Coastal Research*, *17*(2), 407–419.
Sánchez-Román, A., Pujol, M. I., Faugère, Y., & Pascual, A. (2023). DUACS DT2021 reprocessed altimetry improves sea level retrieval in the coastal band of the European seas. *Ocean Science*, *19*(3), 793–809. https://doi.org/10.5194/os-19-793-2023
Serafin, K. A., Ruggiero, P., & Stockdon, H. F. (2017). The relative contribution of waves, tides, and nontidal residuals to extreme total water levels on U.S. West Coast sandy beaches. *Geophysical Research Letters*, *44*(4), 1839–1847. https://doi.org/10.1002/2016GL071020
Siegelman, M. N., McCarthy, R. A., Young, A. P., O'Reilly, W., Matsumoto, H., Johnson, M. O., et al. (2023). Subaerial profiles at two beaches: Equilibrium and machine learning. *Authorea Preprints*. https://doi.org/10.22541/essoar.170000389.98250599/v1
Simmons, J. A., & Splinter, K. D. (2022). A multi-model ensemble approach to coastal storm erosion prediction. *Environmental Modelling & Software*, *150*, 105356. https://doi.org/10.1016/j.envsoft.2022.105356
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*(56), 1929–1958.
Stockdon, H. F., Holman, R. A., Howd, P. A., & Sallenger, A. H. (2006). Empirical parameterization of setup, swash, and runup. *Coastal Engineering*, *53*(7), 573–588. https://doi.org/10.1016/j.coastaleng.2005.12.005
Sweet, W. V., Hamlington, B. D., Kopp, R. E., Weaver, C. P., Barnard, P. L., Bekaert, D., et al. (2022). *Global and regional sea level rise scenarios for the United States: Updated mean projections and extreme water level probabilities along U.S. Coastlines. NOAA technical report NOS 01* (p. 111). National Oceanic and Atmospheric Administration, National Ocean Service.
Thompson, P. R., Widlansky, M. J., Hamlington, B. D., Merrifield, M. A., Marra, J. J., Mitchum, G. T., & Sweet, W. (2021). Rapid increases and extreme months in projections of United States high-tide flooding. *Nature Climate Change*, *11*(7), 7–590. https://doi.org/10.1038/s41558-021-01077-8
Vitousek, S., Barnard, P. L., Fletcher, C. H., Frazer, N., Erikson, L., & Storlazzi, C. D. (2017). Doubling of coastal flooding frequency within decades due to sea-level rise. *Scientific Reports*, *7*(1), 1399. https://doi.org/10.1038/s41598-017-01362-7
Vitousek, S., Vos, K., Splinter, K. D., Erikson, L., & Barnard, P. L. (2023). A model integrating satellite-derived shoreline observations for predicting fine-scale shoreline response to waves and sea-level rise across large coastal regions. *Journal of Geophysical Research: Earth Surface*, *128*(7), e2022JF006936. https://doi.org/10.1029/2022JF006936
Vos, K. (2023). Time-series of shoreline change along the Pacific Rim (v1.4) [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.7758183
Vos, K., Harley, M. D., Splinter, K. D., Walker, A., & Turner, I. L. (2020). Beach slopes from satellite-derived shorelines. *Geophysical Research Letters*, *47*(14), e2020GL088365. https://doi.org/10.1029/2020GL088365
Vos, K., Harley, M. D., Turner, I. L., & Splinter, K. D. (2023a). Pacific shoreline erosion and accretion patterns controlled by El Niño/Southern Oscillation. *Nature Geoscience*, *16*(2), 140–146. https://doi.org/10.1038/s41561-022-01117-8
Vos, K., Splinter, K. D., Harley, M. D., Simmons, J. A., & Turner, I. L. (2019). CoastSat: A Google Earth Engine-enabled Python toolkit to extract shorelines from publicly available satellite imagery. *Environmental Modelling & Software*, *122*, 104528. https://doi.org/10.1016/j.envsoft.2019.104528
Vos, K., Splinter, K. D., Palomar-Vázquez, J., Pardo-Pascual, J. E., Almonacid-Caballer, J., Cabezas-Rabadán, C., et al. (2023b). Benchmarking satellite-derived shoreline mapping algorithms. *Communications Earth & Environment*, *4*(1), 1. https://doi.org/10.1038/s43247-023-01001-2
Yates, M. L., Guza, R. T., & O'Reilly, W. C. (2009). Equilibrium shoreline response: Observations and modeling. *Journal of Geophysical Research*, *114*(C9), C09014. https://doi.org/10.1029/2009JC005359
Young, A. P., Guza, R. T., Matsumoto, H., Merrifield, M. A., O'Reilly, W. C., & Swirad, Z. M. (2021). Three years of weekly observations of coastal cliff erosion by waves and rainfall. *Geomorphology*, *375*, 107545. https://doi.org/10.1016/j.geomorph.2020.107545
Young, A. P., Matsumoto, H., Spydell, M. S., & Dickson, M. E. (2023). Cobble tracking observations at Torrey Pines State Beach, CA, USA. *Journal of Geophysical Research: Earth Surface*, *128*(9), e2023JF007199. https://doi.org/10.1029/2023JF007199
Zeinali, S., Dehghani, M., & Talebbeydokhti, N. (2021). Artificial neural network for the prediction of shoreline changes in Narrabeen, Australia. *Applied Ocean Research*, *107*, 102362. https://doi.org/10.1016/j.apor.2020.102362