On the role of hydrodynamic and morphologic variables on neural network prediction of shoreline dynamics

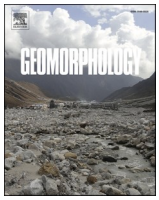
Nadia Senechal, Giovanni Coco

Geomorphology 451 (2024) 109084

Contents lists available at ScienceDirect

Geomorphology

journal homepage: www.journals.elsevier.com/geomorphology



On the role of hydrodynamic and morphologic variables on neural network prediction of shoreline dynamics

Nadia Senechal [a,*], Giovanni Coco [b]

[a] Univ. Bordeaux, CNRS, Bordeaux INP, EPOC, UMR 5805, F-33600 Pessac, France
[b] School of Environment, University of Auckland, 23 Symonds Street, Auckland 1142, New Zealand

ARTICLE INFO

Keywords:
Neural network
Shoreline
Bar
Morphology

ABSTRACT

Predicting shoreline change is a key issue in coastal research. Predictors, process-based or data-driven, tend to be developed and tested on high-frequency and high-quality data sets. Combining hydrodynamic and morphological variables extracted from video images and artificial neural network allows us to evaluate if sparse data could still provide physically-sound shoreline change predictions. The data set covered a 3-year period with shoreline position data (with an accuracy of ±5 m) available 73 % of the time and 66 % for the morphological parameters (beach state or bar location). The best configuration of the trained shallow (one hidden layer) Feedforward Artificial Neural Network (ANN), includes 10 input variables and 10 nodes allowing to capture the shoreline dynamic at different time scales, from the storm-event to the seasonal scale, and to predict the shoreline position on a 1-year period with a RMSE of about 6.7 m. Increasing the complexity of the architecture of the ANN by increasing the number of hidden layers did not improve the predictions. By modifying the number of input variables in the algorithm, the ANN also allows us to highlight the mitigation effect of the bar during the storm event and its role as sediment buffer during seasonal accretion.

1. Introduction

More than one-third (2.75 billion) of the world's population lives both within 100 km distance of the coast and <100 m above sea level (Reimann et al., 2023). Sixty percent of the world's 39 largest cities with a population of over 5 million are located within 100 km of the coast, including 12 of the world's 16 cities with populations >10 million (Nicholls et al., 2007). As pointed out by Newton et al. (2012), the condition for success in integrated coastal zone management is the active involvement of all actors: scientists, stakeholders, governance actors as well as an informed and educated public.

Sandy beaches comprise over 30 % of the world's ice-free coasts (Splinter and Coco, 2021) and, for those beaches situated in wave-dominated environments, shoreline dynamics can vary over a wide range of different temporal and spatial scales (Splinter and Coco, 2021). As a result, building comprehensive datasets of shoreline change remains a challenge. The use of satellite imagery (Vos et al., 2019, 2023) and the emergence of community-based programs (Harley et al., 2019; Harley and Kinsela, 2022) is now making the development of shoreline datasets a reality for every beach worldwide. With more data becoming available from these systems, it becomes imperative to develop models that can take advantage of such measurements, make predictions and help management.

Over the past decades, predicting shoreline change has been a key issue worldwide and many approaches have been developed from models based on conservation equations to equilibrium-based models (e.g. Splinter et al., 2018; D'Anna et al., 2021; Schepper et al., 2021), from data decompositions techniques (e.g. Montaño et al., 2021) to machine learning algorithms (e.g. Montaño et al., 2020; Zeinali et al., 2021). Despite these advances, there is still the need for improved quantitative shoreline prediction in order to mitigate present and future impacts of climatic changes. One type of models that has received attention in recent years is equilibrium-based models, which rely on the disequilibrium concept by evaluating the difference between instantaneous and equilibrium wave energy. These models generally perform quite well on beaches with a clear seasonal cycle and are able to capture both the seasonal and short-term scales despite the limited set of parameters included (e.g. Splinter et al., 2018; D'Anna et al., 2021; Schepper et al., 2021; Ibaceta et al., 2022). However they require long-term high quality data sets and they also require the adjustment of some of the coefficients to take into consideration local conditions (e.g. memory decay, time up-scaling, time down-scaling). In contrast, machine learning algorithms,

* Corresponding author.
E-mail addresses: nadia.senechal@u-bordeaux.fr (N. Senechal), g.coco@auckland.ac.nz (G. Coco).

https://doi.org/10.1016/j.geomorph.2024.109084
Received 18 September 2023; Received in revised form 25 January 2024; Accepted 27 January 2024
Available online 1 February 2024

Table 1
Input variables used in the model. Previous means previous to the shoreline extraction.(blue) variables related to wave forcing, (green) variables related to the tide (yellow) variables related to the beach morphology (orange) variables related to the surf zone width.

Variable description	Number
Offshore Significant wave height averaged the previous tide	1
Offshore Significant wave height averaged the previous day	2
Offshore Significant wave height averaged the previous week	3
Offshore Significant wave height averaged the previous month	4
Cumulated Longshore wave energy flux over the previous day	5
Cumulated Longshore wave energy flux over the previous week	6
Cumulated Longshore wave energy flux over the previous month	7
Tidal range the previous day	8
Maximum Tidal range the previous week	9
Alongshore Averaged inner Bar position the previous day	10
Standard deviation associated with alongshore averaged inner Bar position	11
Mean alongshore averaged inner Bar position estimated the previous week	12
Intertidal Beach State the previous day	13
High Tide offshore breakpoint position	14

among them Artificial Neural Networks (ANN), allow considering a much broader set of parameters. However, it is usually claimed that their performance depends highly on the quantity and the quality of the data available (e.g. Montaño et al., 2020).

As underlined by Splinter and Coco (2021), data is key, especially real-time high frequency and high-quality in-situ coastal monitoring programs with the focus to capture the shoreline and nearshore bathymetry. To meet this demand, the design of low-coast community beach program, such as CoastSnap (Harley and Kinsela, 2022), which empowers local communities to collect quantitative measurements of coastline change using their smartphones, improves data coverage but it almost inevitably produce a sparse dataset (i.e., the time between images available is not regular). Here we propose to evaluate if machine learning, specifically Artificial Neural Networks (ANN), can help providing physical insight in complex environment dynamic using a sparse dataset of oceanographic variables. The purpose of this work is thus to use the simplest ANN with simple morphological parameters, which could be easily extracted from this community beach program data set and with simple hydrodynamic parameters which are widely collected (e.g. wave buoy) or modeled, and to evaluate if it provides reasonable predictions of shoreline change.

2. Study area and field data

2.1. Study area

The field data used here were first published and discussed in Sénéchal et al. (2015). Data were collected at Biscarrosse beach (Bbeach) situated on the southern French Atlantic coast (Fig. 1). Bbeach is an open coast sandy beach with a median grain size of about 0.35 mm facing the Atlantic swells in a meso- to macrotidal environment (e.g. Biausque and Sénéchal, 2019). Bbeach generally exhibits a gentle slope in the lower intertidal area (typically tanβ~0.01–0.02) and a steeper upper beach (typically tanβ~0.03–0.05). The beach generally exhibits a double bar profile with a single intertidal bar and a single subtidal bar (Almar et al., 2009). The inner bar often exhibits 3D patterns. However, observations based on three years of daily low tide images (Péron and Sénéchal, 2011) indicate that the most typical states observed fall in the category of Low Tide Terrace, LTT, and Transverse Bar and Rip, TBR (following the classification by Wright and Short, 1984). Péron and Sénéchal (2011) also indicate that both up-state and down-state transitions were dependent on the previous beach state and that no 'direct jump' from the lowest 'reflective' state to the highest 'dissipative' beach

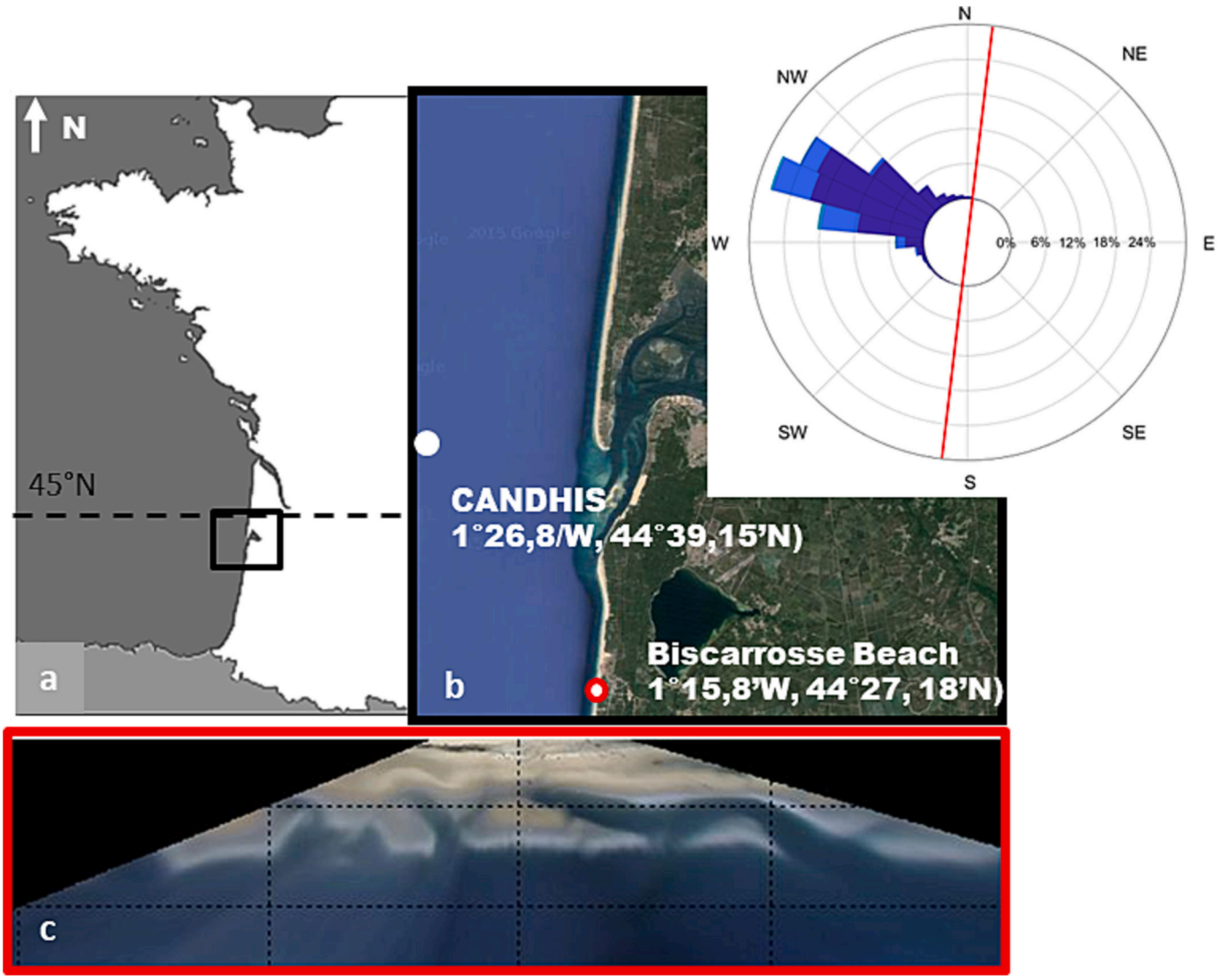


Fig. 1. Field site: (a) Location on the south Atlantic coastline of France; (b) location (1°26.8′W, 44°39.15′N) where data were extracted from the nearest CANDHIS wave buoy or obtained using hindcast. Wave rose with the mean direction of the coast indicated by the red solid line. (c) example of a rectified video image of Biscarrosse Beach at low tide. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

state was observed. Previous studies have shown that the dynamic of the upper beach face presents a dominant seasonal pattern essentially driven by the seasonal wave conditions (e.g. Sénéchal et al., 2015). However the short term 'storm-post-storm' component is also important in this area (Sénéchal et al., 2015; Angnuureng et al., 2017; Biausque and Sénéchal, 2019). In particular Biausque and Sénéchal (2019), using intensive topographic surveys spanning two consecutive winter seasons, showed that short term (within a few days) inter-storm recovery could mitigate the seasonal erosion. Data also showed that shoreline dynamics can be driven by both hydrodynamic (tide, wave) and morphological parameters (bar location, e.g. Sénéchal et al., 2009, 2015; Biausque and Sénéchal, 2020). Diverse management strategies are deployed along the beach-dune system. The back dune is covered by grass in order to be accessible and more attractive to tourists, the southern section is fixed by seawalls (100 m long) and the northern dune is protected by sand fences (Fig. 1, C). However, except for the very southern end of the beach, the effect of these anthropogenic works on the alongshore averaged shoreline dynamic can be considered negligible (Biausque and Sénéchal, 2018).

2.2. Hydrodynamic parameters

In the present study, we used the wave buoy measurements provided by the nearest CANDHIS directional wave buoy location (1°26.8′W, 44°39.15′N) moored in about 54 m water depth nearly 25 km north-west of the field area (Fig. 1). The wave data consist in hourly bulk parameters for the sea state conditions including the significant wave period (H_s), the peak period (T_p) and the direction of the peak period (D_p). In case of failure of the wave buoy, the bulk parameters were extracted from the nearest WW3 model (HOMERE) output using a corrected significant wave height. Correction was obtained from the technical validation proposed by Castelle et al. (2020). In this paper, after correction, H_s hindcast show strong agreement with measured wave data with a coefficient of determination $R^2 = 0.94$ and a root mean square error RMSE = 0.25 m. Accuracy decrease for T_p ($R^2 = 0.25$ and RMSE = 1.7 s).

Consistent with Biausque and Sénéchal (2019), wave energy flux (P_{tot}) has been estimated in 54 m water depth following the linear approach as (1):

$$P_{tot} = \frac{1}{8}\rho g H_s^2 C_g \quad (1)$$

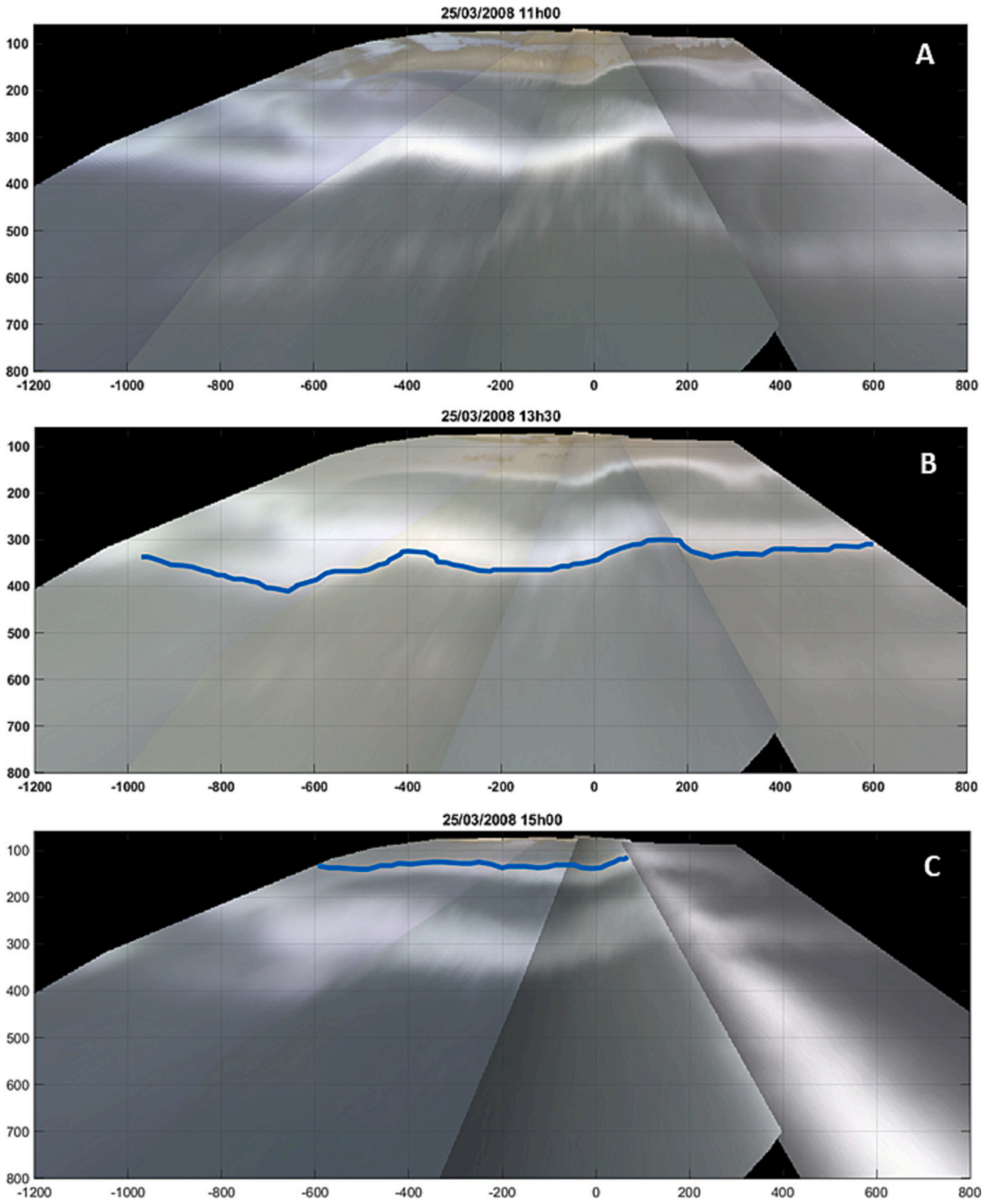


Fig. 2. Merged rectified images of Bbeach on March 25th 2008. Current wave conditions are: $H_s = 2.3$ m and $T_p = 9$ s (A) at low tide to determine the beach state; (B) at water level 1.7 m used for bar extraction (blue line) and (C) at water level 2.6 m used for shoreline extraction (blue line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where C_g is the wave group velocity (2), ρ the density of the ocean's water (1025 kg/m^3) and g the gravitation constant (9.81 m/s^2).

$$Cg = \frac{1}{2}\, c\left(1 + \frac{2kd}{sinh2kd}\right) \tag{2}$$

where c is the phase velocity (3),

$$c = \sqrt{\frac{g}{k} tanhkd} \tag{3}$$

and k the wavenumber and d the depth. The wavenumber k associated with the peak period has been calculated by resolving the linear dispersion relation using a Newton approach.

The longshore component of the energy flux has then been calculated as:

$$P_{ly} = P_{toi} cos\theta sin\theta \tag{4}$$

where θ is the incidence of the wave.

Tide data were provided by both the tidal model developed by the French naval hydrographic and oceanographic service ("*SHOM*") and tide data collected in the Arcachon lagoon (REFMAR).

2.3. Shoreline detection

Datum-based shorelines generally consist of the cross-shore position of a specified elevation contour. Given we used the data set published by Sénéchal et al. (2015), the shoreline was defined in the same way. Thus, the shoreline has been identified as the elevation associated with the isocontour 2.6 m, which corresponds to the elevation where the seasonal berm generally forms and which is consistent with previous works (e.g.

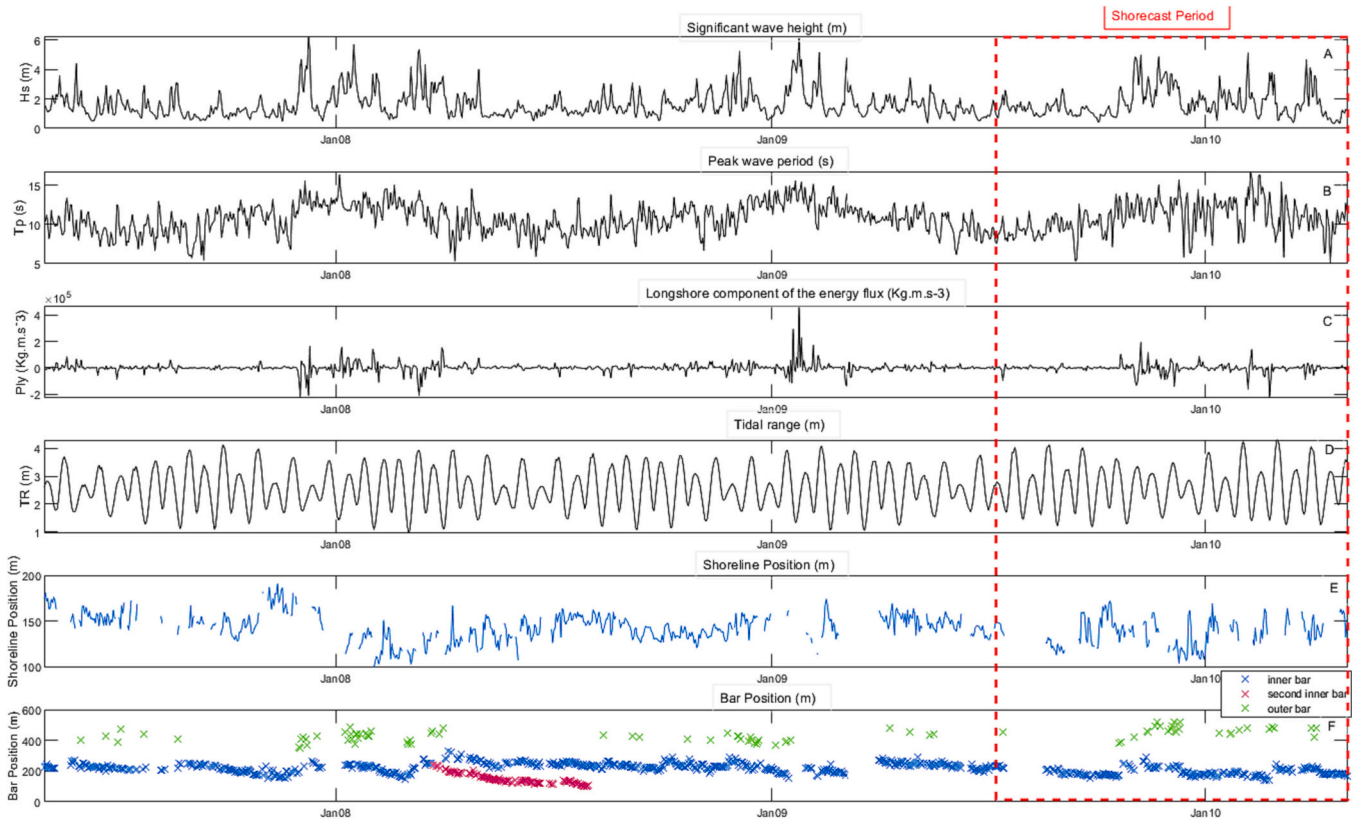


Fig. 3. Environmental conditions, shoreline and bar position at Bbeach. (A) Offshore significant wave height (m); (B) Peak period (s); (C) Longshore component of the energy flux estimated with Eq. (4) (kg·m·s^{-3}), (D) Tidal Range (m), (E) Daily alongshore-averaged position of the shoreline, (F) Daily alongshore-averaged position of the bars: (green cross) outer subtidal sandbar (m) (blue) the inner first intertidal bar and (pink) the second intertidal bar when observed. The red box represents the period used for blind forecast of the FNN model and will be referred here as shorecast period. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Sénéchal et al., 2009, 2015). In the initial work of Sénéchal et al. (2015), shoreline locations were captured from daily oblique photo imagery from May 2007 to April 2010. The video station is composed of five color cameras located on the top of the foredune (height above mean water level is about 26 m). The total alongshore distance covered is approximately 2 km (Fig. 1C) at low tide. The images are recorded during daylight hours, at a frame rate of 2 images/s during 10-minute bursts, resulting in four 10-minute time-exposure images per hour (Almar et al., 2009). A semi-manually technique based on the Minimum Shoreline Variability method (MSV, Almar et al., 2011) is used to extract daily shoreline proxy from 10-minute time-exposure rectified and merged video images (Fig. 2). Daily images were selected using a computed mean water level including wave set-up and the data set was then limited to days when the overall accuracy in the shoreline cross-shore position was ±5 m. At this water level the alongshore covered a distance ranging between 500 m and 700 m depending on the number of camera available (Fig. 2C). Each shoreline position corresponds thus to the alongshore averaged cross-shore position of the mean waterline. The resulting dataset consists of 797 shorelines over a 3-year period (about 73 % of the possible data) and covers the different seasons.

2.4. Morphological parameters

Sand bar position was monitored using the time-exposure video images (see Sénéchal et al., 2015 for a complete description of the methodology). Sand bar position was manually digitized tracking the cross-shore location of the image intensity peaks in the alongshore direction (Fig. 2B). To reduce wave- and tide-artificial shift, bar extraction was performed at the same water level each day and in the presence of offshore significant wave height <2.5 m for the inner bar. The resulting dataset consists of 720 inner bar positions (66 % of the covered period) with an accuracy of ±8 m and 110 days of outer bar position (10 % of the covered period) with an accuracy of ±10 m.

A qualitative description has also been used to characterize the beach intertidal morphology (see Péron and Sénéchal, 2011). Using rectified images extracted at low tide (Fig. 2A), the intertidal beach state was thus visually classified into six states: Low Tide Terrace (LTT); Transitional Low Tide Terrace Transverse Bar and Rip (LTT-TBR); Transverse Bar and Rip (TBR); Rhythmic Bar and Beach (RBB); Dissipative (D) and MBS (Multiple Bar System). Observations showed that despite high energetic conditions (mean annual H_s was 1.66 m), the inner bar exhibited mostly complex 3D patterns, TBR and LTT states being the most frequently observed states and multiple intertidal bar systems being also observed. The same numbers of down-state and up-state transitions were observed and both transitions were observed to be dependent on the previous beach state.

2.5. Environmental conditions

Wave climate clearly exhibits a seasonal trend with longer periods (12–14 s) and more energetic wave conditions during winter and shorter periods (8–10 s) and lower energetic wave conditions during summer (Fig. 3A & B). The winter season is also characterized by the presence of storms that can be either isolated or in cluster (Sénéchal et al., 2015). Waves are essentially propagating from the W-N-W (280°N-300°N) corresponding to a north incidence with the coastline orientation at Bbeach (Fig. 1). Tidal range varies between 1 m and 4.2 m (Fig. 3C).

The shoreline exhibits seasonal cycles with an overall accretion during spring and summer and significant erosion during the winter period. The dynamic of the shoreline is also highly sensitive to the short-

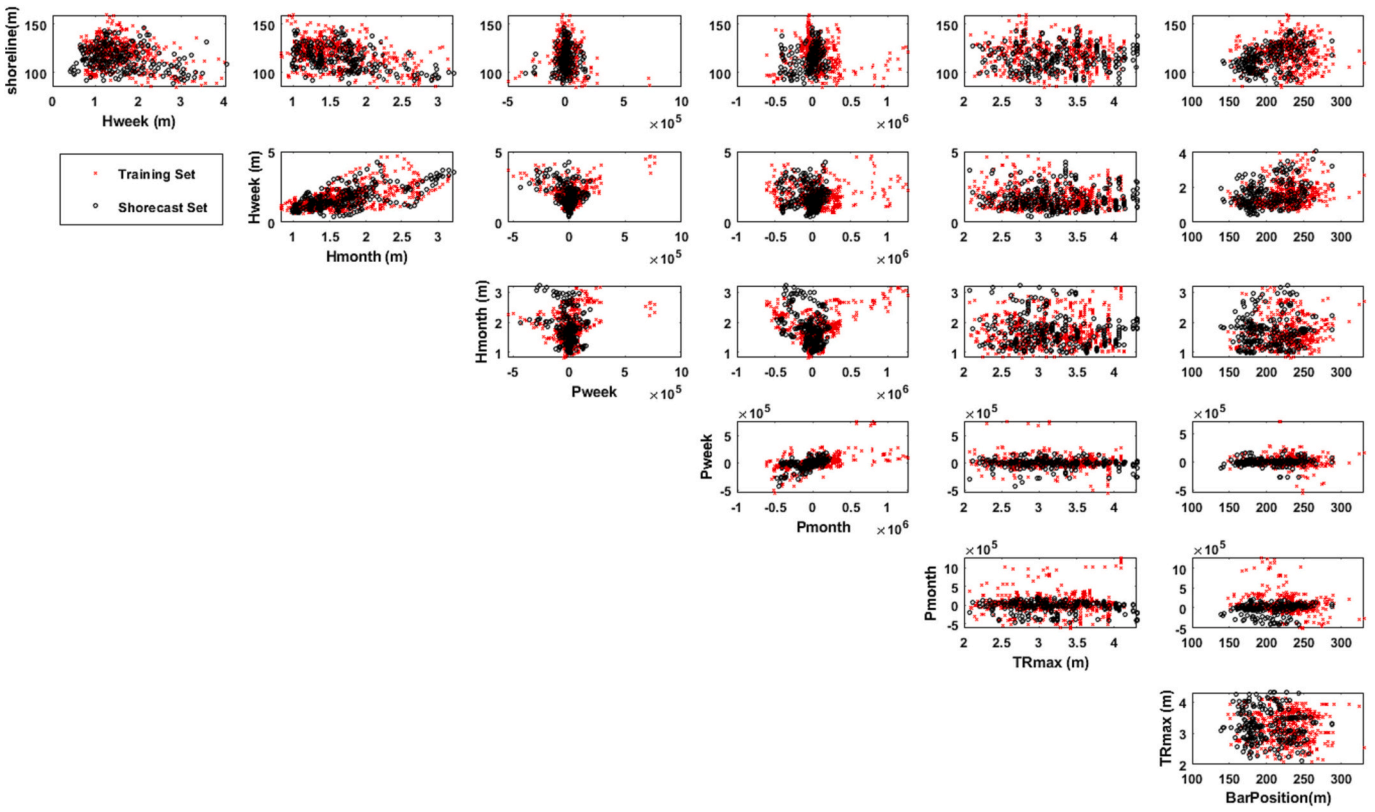


Fig. 4. Results of data selection and correlation between the input variables. In red is the training data set (~67 % of the original dataset) and in black is the Shorecast test (~33 %). Input variables considered: shoreline position, offshore significant wave height averaged the week (H_{week}, variable 3) and the month (H_{month}, variable 4) before, cumulated longshore wave fluxes over the previous week (P_{week}, variable 6) or month (P_{month}, variable 7), maximum tidal range the week before (TR_{max}, variable 9) and alongshore averaged inner bar position the day before (BarPosition, variable 10). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

term storm/post-storm forcing (Biausque and Sénéchal, 2019). The alongshore-averaged shoreline position varies between 100 m and 170 m from the dune toe where the video system is deployed (Fig. 3E). We observe that at the end of the winter 2008, the beach exhibited a double bar system in the intertidal domain (Fig. 3F). The intertidal beach can generally be classified as TBR (68 % of observations, Péron and Sénéchal, 2011) and its cross-shore dynamic is strongly driven by the mean water level (Almar et al., 2010). The outer bar system, in contrast, is sensitive to offshore incoming waves and wave breaking occurring at high tide is essentially observed only during energetic conditions.

3. Methodology

3.1. Artificial Neural Network

Calkoen et al. (2021) recently compared Machine-Learning methods for forecasting sandy shoreline evolution using satellite-derived shorelines. Their results indicated that the performance of the simple Feed-forward network (FFN) was similar to the one of more complex Neural Networks (NN) like Recurrent Neural Networks (e.g. deep Autoregression) or Long Short Term Memory (LSTM) NN. Montaño et al. (2020) using video derived shoreline and diverse Machine-Learning models in term of approach or architecture (e.g. k-Nearest Neighbor, Autoregressive NN with exogenous inputs, LSTM, Bayesian Network) also concluded that all models displayed similar performance in shoreline dynamics at event and seasonal scales. Since the main goal of the present study is to explore how morphological and hydrodynamic parameters affect shoreline change prediction, we used the simplest basic shallow neural network, the FFN. Basically in this kind of network, the information moves in only one direction-forward-from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network in contrast to more complex NN.

The architecture of the FFN consists of three types of layers: the input layer, the hidden layer, and the output layer. Each layer is made up of units known as neurons: the neurons of the input layers will be called input variables (see Section 3.2 for their description), the neurons of the hidden layer will be called nodes and in the present work and there is only one neuron in the output layer which corresponds to the predicted shoreline position. The layers are interconnected by weights. Each hidden layer's node takes the weighted sum of the outputs from the previous layer (the input layer or the previous hidden layer if several hidden layers are present), apply an activation function, and pass the result to the next layer. The final layer (output layer) finally produces the output (shoreline position) for the given inputs. Each neuron in one layer is thus connected to every neuron in the next layer. There are also no memory cells or gates that allow FFN to retain or forget information over time selectively in contrast to more complex NN (e.g. LSTM).

There are two phases in FFN process: 1- the feedforward phase: the input variables are fed into the network and propagate forward through the network until the output layer is reached, and a predication is made. 2- the backpropagation phase: once a prediction is made, the error (a measure of the difference between the predicted and the actual output) is calculated. This error is then propagated back through the network, and the weights are adjusted to minimize this error (using for example a gradient descent optimization algorithm).

Training a FFN network involves using a dataset to adjust the weights of the connections between neurons. Thus the shoreline data set is first divided into two different sub data sets: a calibration data set and a blind forecast data set. To ensure that the calibration data set covered the different seasons, we used the initial 66 % of shoreline data of the time

Table 2
Model performance when trained with 10 input variables (1–7, 9–11) and different number of nodes. The values in bold correspond to the calibration period and the other values correspond to the shorecast period.

	7 nodes	10 nodes	12 nodes	15 nodes
RMSE	7.76/**9.81**	6.7/**8.71**	7.67/**8.92**	6.8/**8.3**
CE	0.59/**0.41**	0.72/**0.61**	0.60/**0.59**	0.68/**0.65**
R	0.67/**0.72**	0.76/**0.82**	0.68/**0.77**	0.65/**0.85**
R^2	0.45/**0.51**	0.58/**0.67**	0.46/**0.59**	0.42/**0.72**

series, the remaining 34 % were used for the blind shorecast data set (red box in Fig. 2). The calibration data set has been further divided into a training data set and a test data set. During the training process, the training data set is passed through the network multiple times, and each time, the weights are updated to reduce the error in prediction made on the test data set. The network was trained using the Levenberg-Marquardt algorithm and the goal is to reproduce the shoreline position as the test data, as accurately as possible. The training is carried into 120 epochs, which means that each data in the training data set has been used 120 times to update the internal model parameters. The activation function used in the hidden layer is a sigmoid function while the transfer function used in the output layer is a linear transfer function. The initial learning rate used in the gradient descent optimization algorithm was set to the default value 0.001. We used the Mathworks, Deep Learning Toolbox of Matlab support to set up the FFN.

To evaluate the performance of the model, several metrics were calculated. Indeed previous works have highlighted the difficulty in choosing a metric for assessing model performance as it might favor one model to another one (e.g. Montaño et al., 2020; Calkoen et al., 2021). It is thus important to consider multiple metrics and different approaches to evaluate the model. Model performance was evaluated using: the root mean squared error (RMSE), the coefficient of efficiency (CE) and the correlation coefficient R. The coefficient efficiency (CE) was calculated as:

$$CE = 1 - \left(\frac{\sum_{i}^{n} (Xi - Xe)^2}{\sum_{i}^{n} (Xi - \overline{Xi})^2} \right) \quad (5)$$

where Xi is the observed input, Xe is the estimated value and $\overline{Xi}$ is the averaged of observed inputs. A CE closer to 1 means higher accuracy of the model. Model performance was also assessed in terms of quantile-quantile (prediction vs measurements) which provides information about extreme events, the direction of shoreline change and mean behavior.

3.2. Input variables

In total, 14 input variables have been evaluated describing the offshore wave forcing (variables 1 to 7), the tide (variables 8 and 9), the morphology of the intertidal area (variables 10 to 13) and a proxy of the surf zone width at high tide (variable 14). The accuracy of the method relies on the training data set and particularly on its capacity to represent the largest and most diverse environment's conditions (Fig. 4).

3.3. Model calibration

The first step of the calibration was ensuring that the Artificial Neural Network provided reasonable results. The FFN is a shallow network with only one hidden layer so calibration was only done on the number of nodes in such layer. Table 2 resumes the model performance: on the training data set (bold, referred to as calibration) and on the blind forecast (referred to as shorecast).

We observe that the FFN displays good performance on the calibration data set for the different number of nodes. Given the simplicity of

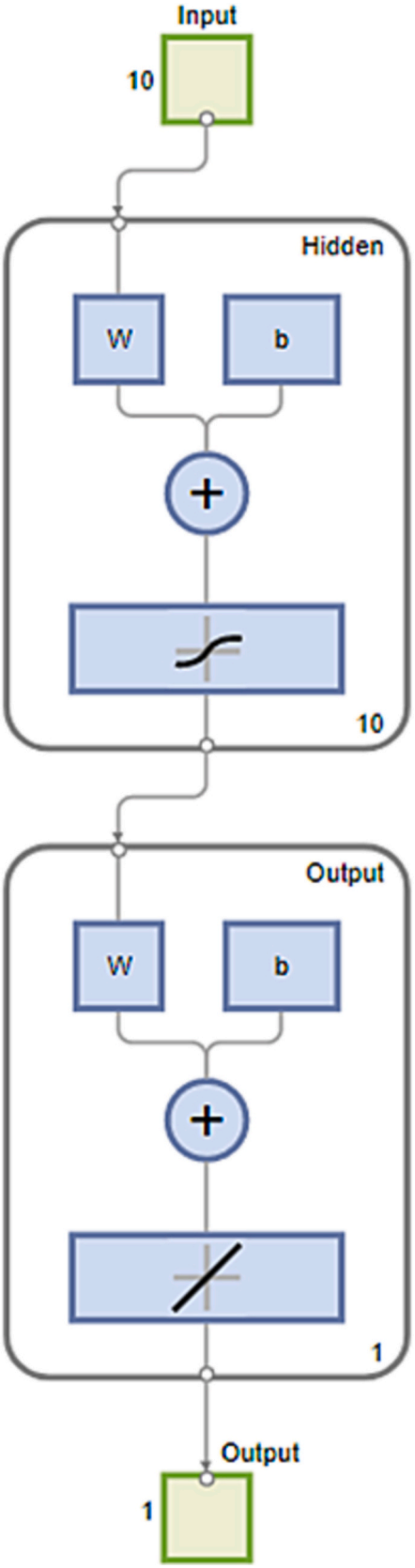


Fig. 5. Architecture of the FFN with 10 input variables and 10 nodes.

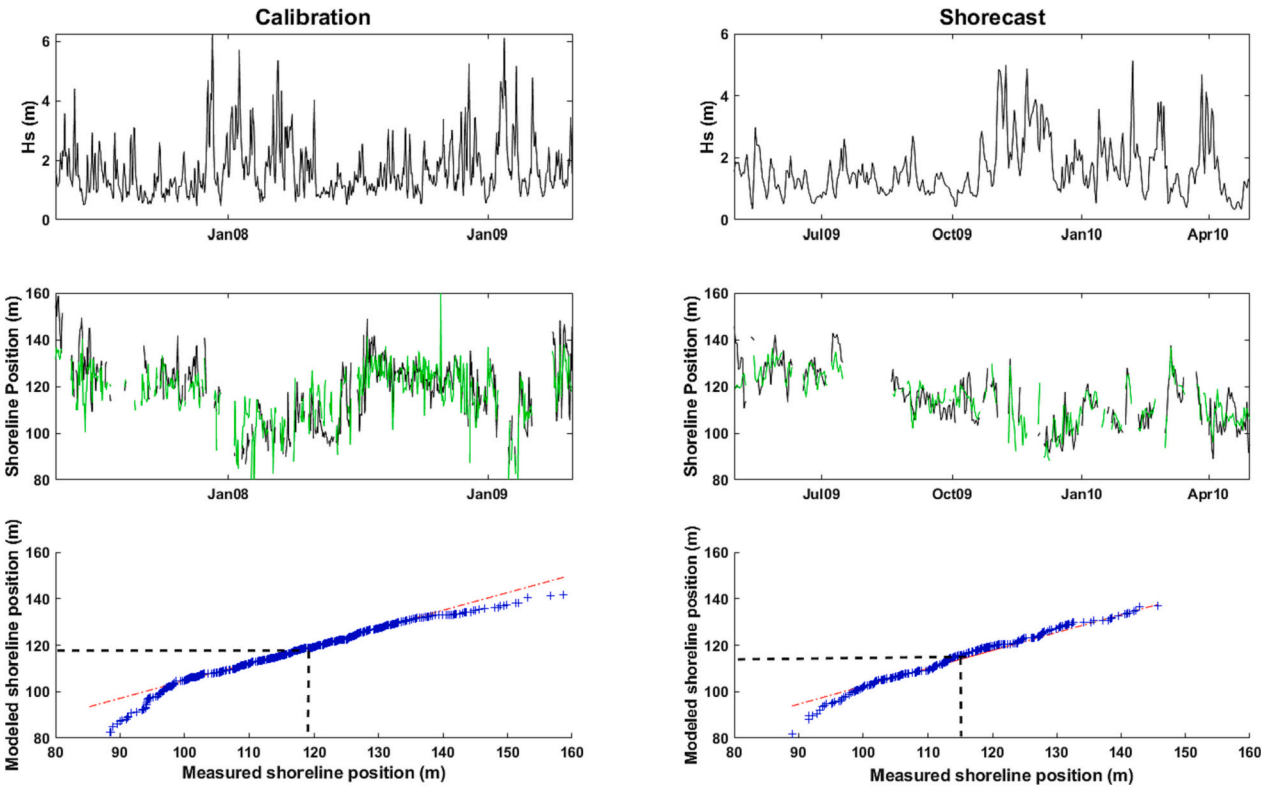


Fig. 6. Models outputs and performance obtained with 10 nodes and 10 exogenous parameters. The middle figures represent (left) the (black) measured and (green) modeled shoreline position during the calibration period and (right) during the shorecast period. The bottom figures represent the Quantile-Quantile plots of model behavior during (left) calibration and (right) shorecast. Dashed lines indicate the average shoreline position. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the input variables and the resolution of the shoreline measurements (±5 m), the RMSE is reasonable. Using the different metrics, it appears that the configuration with 15 nodes displays the best performance. However, if one takes into consideration the blind forecast only, it appears that the best configuration is obtained with 10 nodes. Fig. 5 illustrates the architecture of the FFN with the sigmoid activation function used in the hidden layer and the linear transfer function used in the output layer for this specific configuration. The initial learning rate was set to 0.001 and when the training process was stopped, the learning rate was 0.01.

Fig. 6 illustrates that during the calibration period, the FFN is able to reproduce seasonal cross-shore (alongshore averaged) shoreline behavior as well as short-term behavior both accretive and erosive. For example the FFN is able to reproduce the seasonal accretive period at the beginning of the summer '08 (right-hand of the black rectangle) and faster oscillations during fall '07. The quantile-quantile plot displays a left-skewed shape. During the shorecast period, the FFN is also found to be able to reproduce the seasonal trend and the high frequency erosive-accretive event. The quantile-quantile plot also displays a left-skewed shape.

4. Results

The FNN has then been trained with different numbers of input variables. As a comparison, we evaluated how each input variable fits the data using a linear regression method. To do this, the best fit using linear regression has been calculated on the training period and the parameters have then been applied to the shorecast period. Table 3 resumes all the different cases. Cases 1 to 14 represent the result of the linear regression method for each input variable. Cases 15 to 38 were obtained by using the FFN with different input variables. Cases 15 to 27 only include input variables associated with the wave forcing and the tide (cases 17 and 26). Cases 28 to 38 also include parameters related to the intertidal beach morphology. Bulk metrics of model performances (RMSE, CE and R) for each tested case are presented in Fig. 7 for (blue) the calibration period and (green) the shorecast period.

The differences observed in the correlation coefficient between the calibration and the shorecast are not statistically significant (Fig. 7). The RMSE is generally higher for the calibration period but this is not surprising, given that modelling extremes is difficult and the shoreline reached its most eroded location and its most accreted location during this period (see Fig. 4). The coefficient efficiency (CE) is larger for the shorecast period than for the calibration period when parameters of different nature (hydrodynamic and morphological) are considered (cases 28 to 38). Overall, results indicated that cases including both hydrodynamic conditions (waves with or without tide) and morphological parameters perform better than the ones relying only on one kind of parameters; this improvement is consistently observed for the three metrics used to assess model performance.

4.1. Hydrodynamic variables

Focusing first on wave conditions (Fig. 7, cases 1 to 7), we observe that the simple linear regression approach does not fit the data when considering only the longshore fluxes (Fig. 7, cases 5, 6 7), consistent with previous observations indicating that Bbeach is primarily driven by cross-shore processes (e.g. Angnuureng et al., 2017; Biausque and Sénéchal, 2018, 2019, 2020). When analyzing each of the significant wave height descriptors in isolation, taking into consideration wave conditions on longer time scales (e.g. previous month, case 4) improves significantly the results compared to the cases where only wave on short time scales are considered (cases 1 and 2). This is consistent with the observations reported in Sénéchal et al. (2015). They observed a negative relationship between the wave conditions and the shoreline

Table 3
The 14 cases (1 to 14) evaluating data fitting with the linear regression for each individual input variable and the 24 trained configurations (cases 15 to 38) with the various input variables (see Table 1 for the detailed definition of each variable).

	Hs – Tide	Hs – Day	Hs – Mean	Hs - month	Flux – Day	Flux - week	Flux - Month	Tide – Day	Max Tide - Week	Inner Bar Day	Standard Deviation	Inner Bar week	Beach State	Offshore Breakpoint
case 1	■													
case 2		■												
case 3			■											
case 4				■										
case 5					■									
case 6						■								
case 7							■							
case 8								■						
case 9									■					
case 10										■	■			
case 11										■				
case 12												■		
case 13													■	
case 14														■
case 15	■	■												
case 16	■		■											
case 17	■			■					■					
case 18	■	■	■											
case 19	■	■		■										
case 20	■	■	■	■										
case 21				■			■							
case 22		■	■	■			■							
case 23			■	■			■							
case 24	■	■	■	■			■							
case 25	■	■	■	■	■	■								
case 26	■	■	■	■	■		■		■					
case 27	■	■	■	■	■		■							
case 28	■	■	■	■	■		■			■	■			
case 29				■			■			■	■			
case 30				■			■		■	■	■			
case 31	■	■	■	■	■		■		■	■	■			
case 32	■	■	■	■	■		■		■			■		
case 33	■	■	■	■	■		■		■	■	■	■		
case 34	■	■	■	■	■		■		■				■	
case 35	■	■	■	■	■		■		■	■				
case 36	■	■	■	■	■		■		■	■				■
case 37	■	■	■	■	■	■	■	■	■	■	■			■
case 38	■	■	■	■	■	■	■	■	■	■	■			

position, which is consistent with a shoreline migrating landward under energetic conditions (eroding conditions) and offshore under calm conditions (accreting conditions) but they also reported that the alongshore averaged shoreline position was only poorly related to the daily variation in offshore conditions suggesting that short-duration events (typically <1 day) may only have a limited effect on the shoreline position.

Fig. 8 illustrates how the FNN is able to reproduce the seasonal cross-shore (alongshore averaged) shoreline behavior as well as the short-term behavior both accretive and erosive although there is a shift in the recovery rate at the seasonal scale (case 24) (see left panels in Fig. 8). Including the tide allows improving the recovery rate at the seasonal scale (case 26) (see right panels in Fig. 8) but at the same time it also increases the erosion rate at short scales leading to an overestimation of the erosion at this time scale. The Quantile-Quantile plot displays a light tailed shape. The bulk parameter of performance also improved: the RMSE decreases from 10.6 m in case 24 to 9.8 m in case 26, the coefficient efficiency increases from 0.24 in case 24 to 0.35 in case 26 and the correlation coefficient increases from 0.35 to 0.48.

4.2. *Morphological variables*

Fig. 9 illustrates the impact of the morphological variables. Because of the lack of observations related to some variables, particularly the ones on the inner bar position, the shorecast period in case 20 (left) is sparser. However if we focus on the shorecasts obtained simultaneously by the two cases, we observe that the shorecasts obtained in case 29 (right) are better than the shorecasts obtained in case 20. The improvement is observed during recovery periods and also during the erosive events. In particular, including the inner bar limits the erosion during consecutive energetic events, consistent with the physical processes. Indeed several studies reported that the presence of the bar will protect the upper part of the beach from severe erosion especially during storm events (e.g. Almar et al., 2009; Angnuureng et al., 2017). The seasonal recovery is also consistent with previous observations indicating that the berm dynamic is linked to the presence and the position of the inner bar (e.g. Sénéchal et al., 2009). In term of prediction performance on the shorecast period (blind prediction of the models), the RMSE drops from 10.5 m in case 20 to 7.7 m in case 29, the coefficient efficiency rises from 0.25 to 0.60 and the correlation coefficient

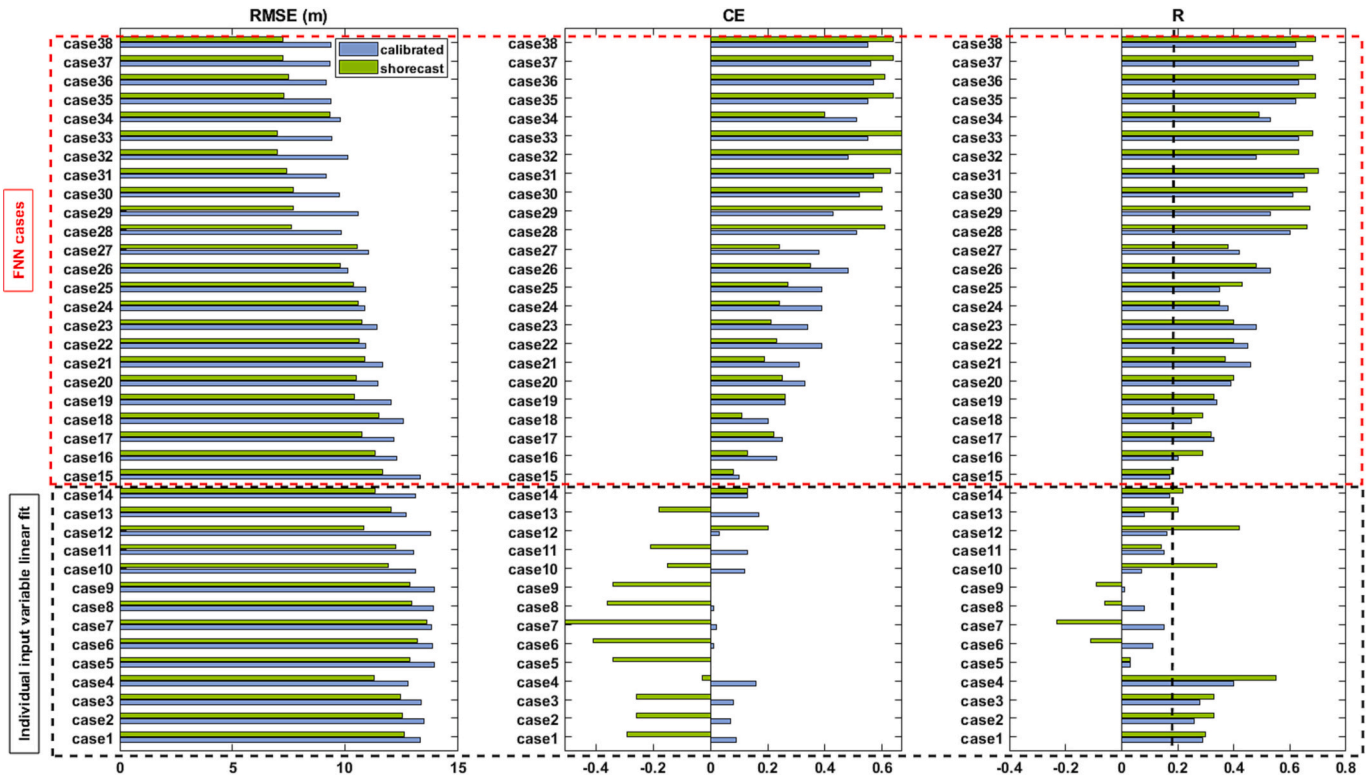


Fig. 7. Model performance: RMSE, coefficient efficiency (CE), correlation coefficient (R) for each individual input variable linear fit (cases 1 to 14, black box) and the FNN (cases 15 to 38, red box) and both the (blue) calibration period and (green) the shorecast period. The vertical dashed line indicates the value of R significant at 99 %. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

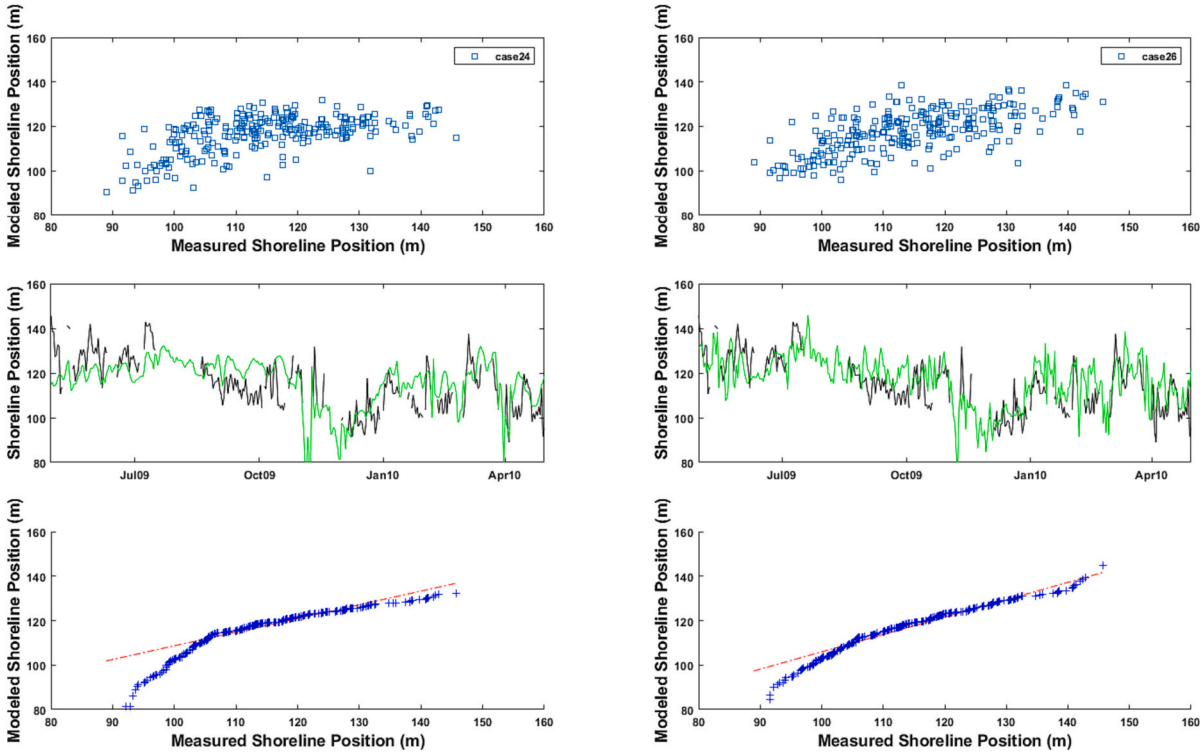


Fig. 8. Model performance with wave inputs at different temporal scales and tide: (left) wave parameter at different time scales (case 24) and (right) when adding tide (case 26). The upper plots represent the modeled versus the measured shoreline; the middle panels represent the time series of (black) measured shoreline and (green) modeled shoreline; the bottom figures represent the Quantile-Quantile plots of model behavior. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

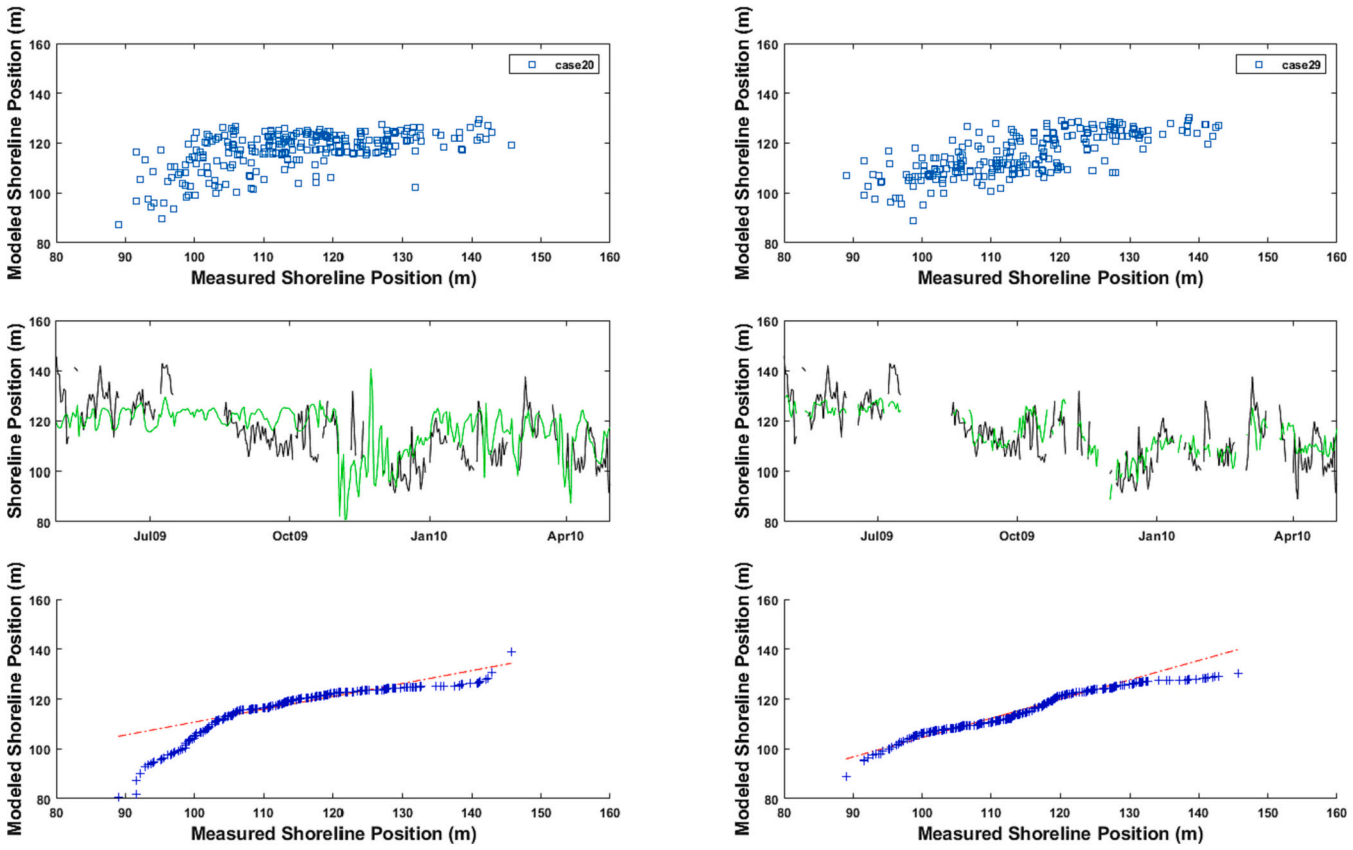


Fig. 9. Model performance with the same number but different kind of input variables: (left) only wave parameters (case 20); and (right) wave and morphological parameters (case 29). The upper plots represent the modeled versus the measured shoreline; the middle panels represent the time series of (black) measured shoreline and (green) modeled shoreline; the bottom figures represent the Quantile-Quantile plots of model behavior. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

increases from 0.40 to 0.67.

Because recording the inner bar position is not always possible, we evaluated if providing qualitative information of the beach state morphology will improve model outputs. Fig. 10 displays the results obtained in case 26 and in case 34 (including beach state). We observe that including the beach state does not improve the prediction of the impact of storms on shoreline position. This result is not surprising given the limited states exhibited by Bbeach (only 6 different states) and the TBR state being observed 68 % of the observed period.

5. Discussion

5.1. Uncertainty and model robustness

Since the model developed is entirely based on data, it is critical to quantitatively assess the type of uncertainty and where such uncertainly is found. While some parameters are relatively easy and widely available (e.g. waves, tide), other parameters are not always easily accessible (e.g. bar location, beach state, memory decay). In the present work, there are two main types of uncertainties. The first ones arise from the data set collected and are specifically related to the estimation of the mean sea level. This uncertainty was estimated to be ±5 m for the averaged shoreline position. The uncertainties associated with the inner bar location were estimated to be ±8 m. Since at times we used wave conditions obtained through a numerical model, a potential different source of error should also be considered, especially with respect to the period and the direction. The second source of uncertainty arises from the model itself: machine learning methods do not take into consideration the physical processes but instead try to find hidden patterns and relations among drivers and response. The number, the choice and the definition of the input variables used in the machine learning may considerably modify the results. In the present study we decided to use simple variables that can be easily computed without knowledge of the typical timescales of the field study as required for some more complicated parameters (e.g. memory decay, time-upscaling, time downscaling). The calibration of the model indicated that the model was able to reproduce the cross-shore shoreline behavior at both seasonal (several months) and short-time (several days) scales with reasonable good performance parameters consistent with previous works (e.g. Montaño et al., 2020). During the calibration period, the model failed at capturing correctly an erosive event in early April 2008 although results were consistent with the physical processes. This might be due to the difficulty for the algorithm to capture the most extreme and/or rare events. In this particular case, this erosive event was associated to a rare combination of morphological and hydrodynamic parameters. Indeed two inner bars in the lower intertidal domain were observed and this event also resulted from the combination of spring high tide, dissipative beach configuration and weak oblique wave conditions (see Coco et al., 2014).

The robustness of the model was then evaluated during the shorecast period (one year of blind prediction). The model showed comparable or even better skills on this period, probably because the shorecast period did not include as many extreme events as the calibration dataset. One could argue that the model performance and model robustness could be improved by using a more complex architecture for the neural networks. Table 4 resumes the model performance by using a more complex architecture. The model has been run using the same input variables used for the calibration but with different number of hidden layers with the sigmoid activation function used in the hidden layer and the linear transfer function used in the output layer. We observe no significant

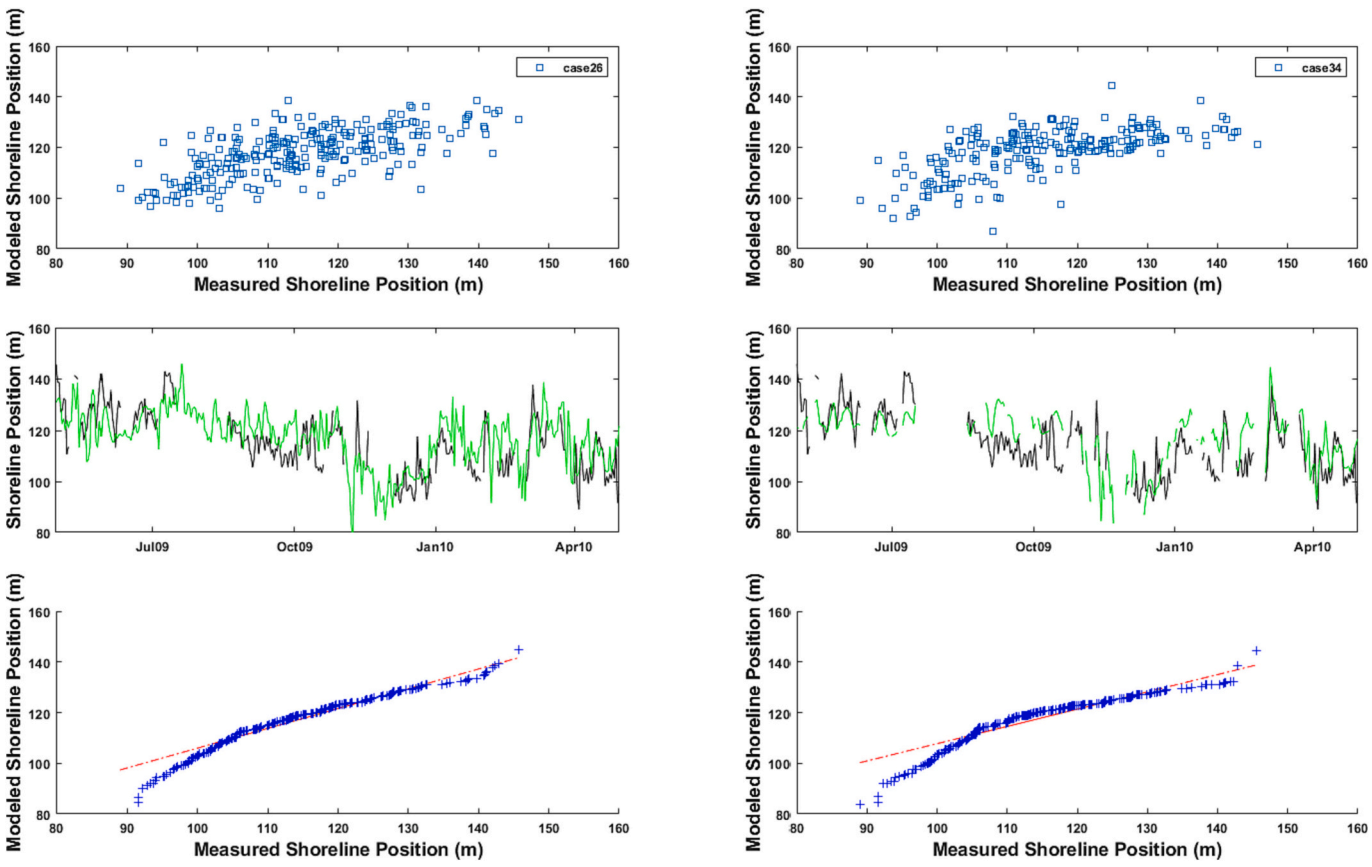


Fig. 10. Model performance including beach state classification: (left) only hydrodynamic parameters (case 26); and (right) including beach state (case 34). The upper plots represent the modeled versus the measured shoreline; the middle panels represent the time series of (black) measured shoreline and (green) modeled shoreline; the bottom figures represent the Quantile-Quantile plots of model behavior. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Model performance when trained with 10 input variables (1–7, 9–11), 10 nodes and different number of hidden layers. The values in bold correspond to the calibration period and the other value correspond to the shorecast period.

	1 hidden layer	2 hidden layers	3 hidden layers	4 hidden layers	5 hidden layers
RMSE	6.7/**8.71**	7.31/**9.81**	6.93/**8.89**	7.10/**8.38**	7.51/**8.72**
CE	0.72/**0.61**	0.63/**0.51**	0.67/**0.60**	0.66/**0.63**	0.62/**0.61**
R	0.76/**0.82**	0.56/**0.75**	0.70/**0.79**	0.60/**0.82**	0.61/**0.80**
R^2	0.58/**0.67**	0.31/**0.56**	0.49/**0.62**	0.48/**0.67**	0.37/**0.64**

improvement either on the calibration or on the shorecast period. One could also argue that the model could be improved by using other neural networks: recurrent network, radial basis function network, time delay neural network or generalized regression neural network. However, there is no general consensus about this point. Zeinali et al. (2021) applying different network approaches on a 34 years of monthly topographic surveys did not obtain more accurate results. Similarly, Calkoen et al. (2021) applying various Machine Learning time-series forecast algorithms (e.g. Simple Long-Short-Term Memory, Deep Autoregression, Multi-Quantile Convolutional Neural Network, Simple Feed-forward) to satellite-derived shorelines concluded differences in accuracy scores within this group of algorithms were notably small and obtained lower accuracy scores than in the present study. More recently, deep learning techniques have resulted in the development of a model that more accurately predicts shoreline position at Tairua beach (Gomez-de la Peña et al., under review).

5.2. How many input variables?

It is often claimed that models fail at providing good prediction of cross-shore shoreline dynamic because we lack all relevant variables to describe the different timescales involved in shoreline change. Schepper et al. (2021) recently showed that the addition of a time-upscaling and a down-upscaling terms significantly improved predictions of cross-shore shoreline changes at different timescales and their interactions. Fig. 11 illustrates that increasing the number of variables will not necessarily improve the model performance. In this figure, both the training and the shorecast periods are represented. Both models are able to reproduce the seasonal pattern of the shoreline dynamic and also give similar results at the short-term event scale. The differences in model performance parameters between the two cases are not statistically significant: the RMS is 8.0 m and 8.2 m (for case 30 and 37, respectively) the coefficient efficiency are 0.56 and 0.60, and the coefficient correlation is 0.65 in both cases. The Quantile-Quantile plots are very similar. Several reasons can explain why the increase in the number of variables will not necessarily improve the model performance. The first reason is probably that the added variable is cross-correlated to another variable and thus does not provide a complementary information. However this seems not the main reason in the present study as the different input variables show weak correlation (Fig. 4). The second reason is that the variable might not be relevant physically. Another reason can be associated to the uncertainties associated with the parameters, which might increase and reduce the overall performance of the model.

Overall, also when dealing with machine learning, the issue is not only how many input variables to use but rather which input variables should be included in the model.

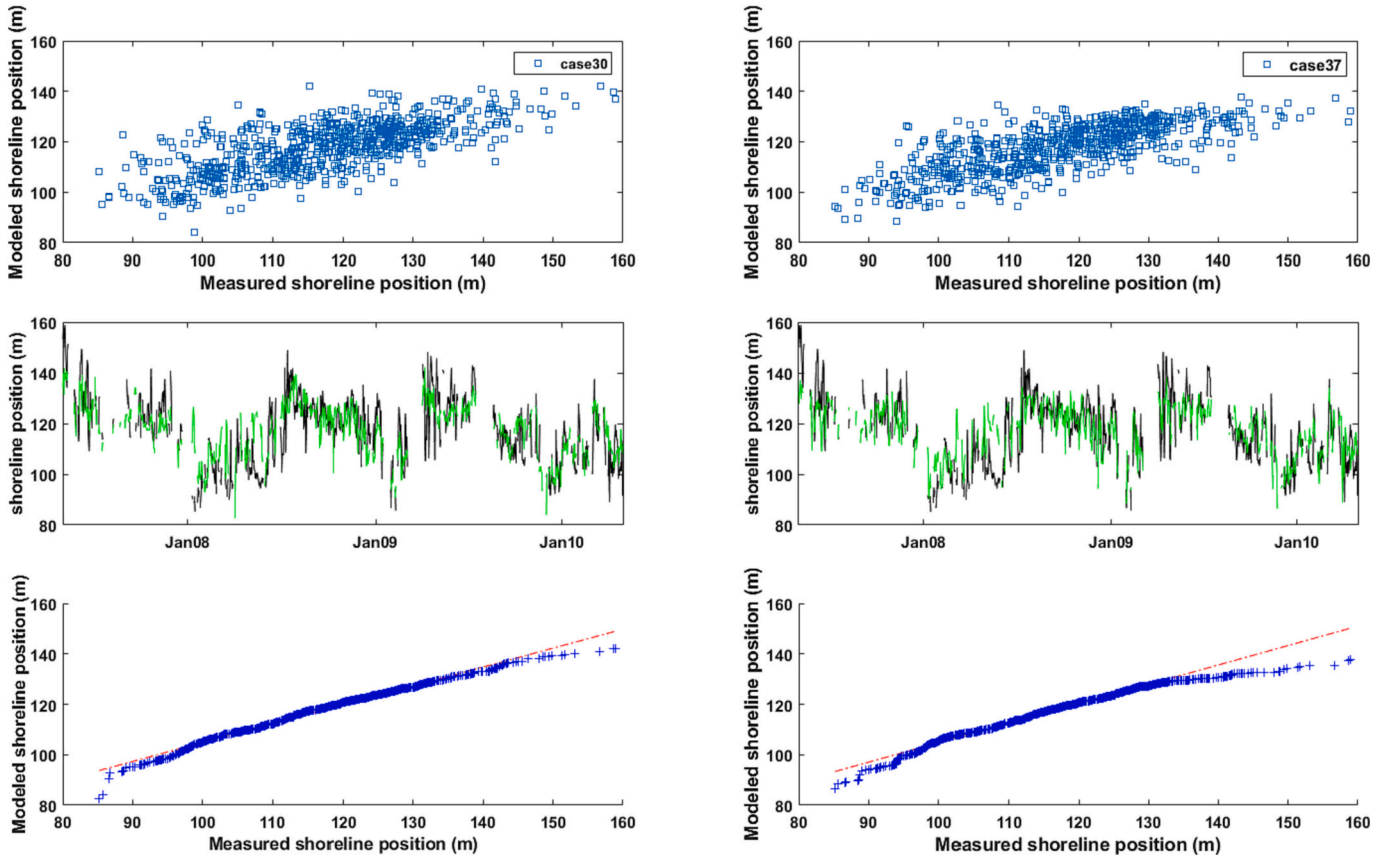


Fig. 11. Model performance with different number of exogenous parameters: (left) only 5 parameters (case 30); and (right) 12 parameters (case 37). The upper plots represent the modeled versus the measured shoreline; the middle panels represent the time series of (black) measured shoreline and (green) modeled shoreline; the bottom figures represent the Quantile-Quantile plots of model behavior. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.3. *Which input variables?*

The model provided results consistent with the physical processes reported previously for this area (Almar et al., 2009; Sénéchal et al., 2015; Angnuureng et al., 2017; Biausque and Sénéchal, 2018) despite the use of a sparse data set, as well as the complexity of the field site characterized by the presence of multiple bars, a *meso* to macrotidal environment and a strong seasonality in the wave conditions associated with a significant storm activity. In particular, the model was able to provide further insight on the role of the inner bar cross-shore position in the dynamic of the shoreline: its mitigating effect during storm conditions but also its role as sediment buffer during seasonal recovery, consistent with previous observations on wave-dominated sandy coastlines (e.g. Phillips et al., 2017). The model was also able to provide further insight on the role of tides. In particular the inclusion of the tide improved the model prediction of the erosive event observed in early April 2008 (see black rectangle in Fig. 12). In term of prediction performance on the training period, the RMSE drops from 10.9 m in case 24 to 10.1 m in case 26, the coefficient efficiency rises from 0.39 to 0.48 and the correlation coefficient increases from 0.38 to 0.53. This is consistent with the observations of Coco et al. (2014) on a nearby beach who reported a similar erosive event. Observations indicated that it resulted from the combination of low energetic oblique waves, a dissipative beach profile resulting from an up-state transition of the outer bar under consecutive severe conditions and spring tidal conditions allowing thus these low energetic waves to reach the upper beach face without breaking.

6. Conclusions

A sparse data set of hydrodynamic and morphological parameters that could be relatively easily extracted for example from low-coast community beach programs was used to develop a simple Feedforward Neural Network with only one hidden layer. The algorithm provided reasonable forecast of shoreline position, capturing both the seasonal scale and the short-term (storm/post-storm) scale. The model was also able to provide some insight into the physical processes driving the dynamic of the shoreline. With the recent world-wide growth of low-coast community beach programs and so the possibility to include many input variables into machine learning algorithms, there is the possibility to develop more reliable predictions of shoreline positions and gain further insight into the physical processes.

CRediT authorship contribution statement

Nadia Senechal: Writing – review & editing, Writing – original draft, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Giovanni Coco:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: SENECHAL Nadia reports financial support was provided by University of Bordeaux. COCO Giovanni reports financial support was provided by School of Environment, University of Auckland. Senechal Nadia reports

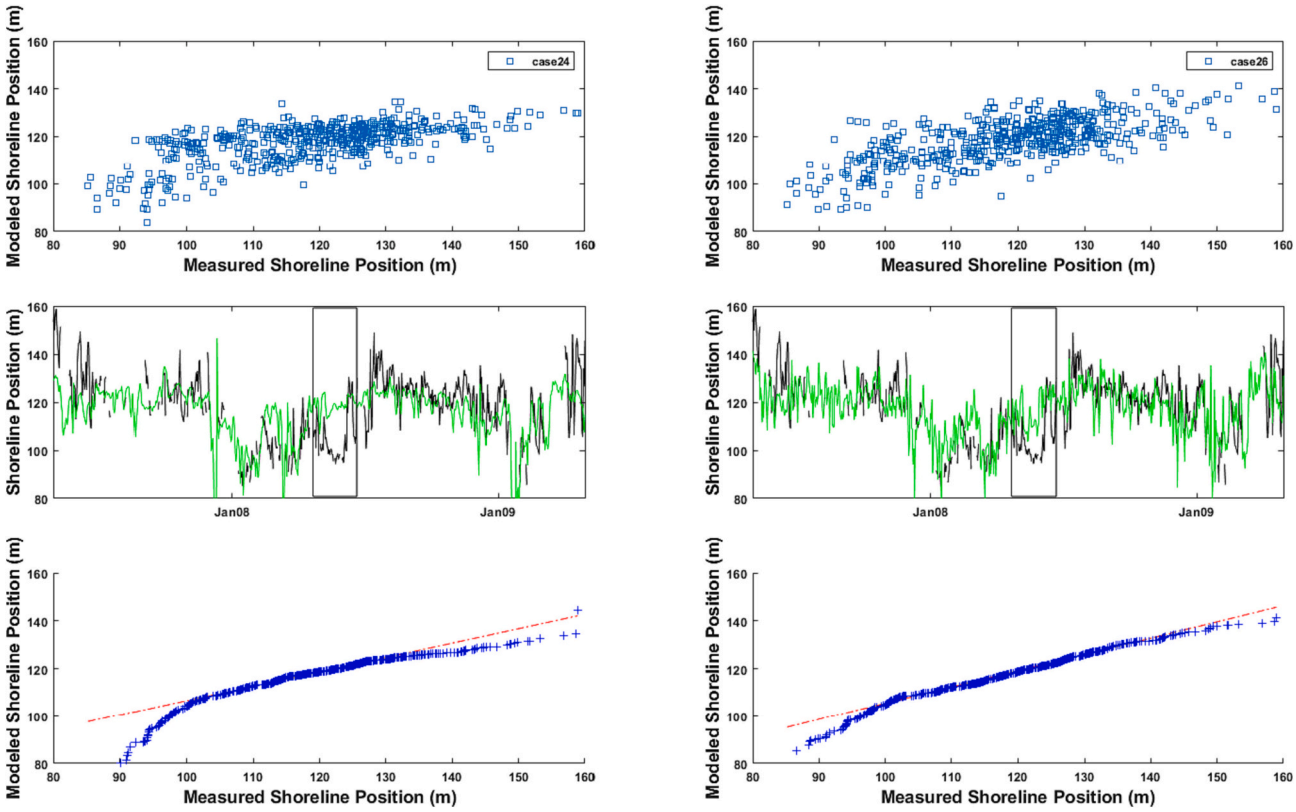


Fig. 12. Impact of the tide in Model performance on the training period: (left) wave parameter at different time scales (case 24) and (right) when adding tide (case 26). The upper plots represent the modeled versus the measured shoreline; the middle panels represent the time series of (black) measured shoreline and (green) modeled shoreline; the bottom figures represent the Quantile-Quantile plots of model behavior. The black box indicates the erosive event of April 2008 reported in Coco et al. (2014). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

financial support was provided by INSU SNO Dynalit. Senechal Nadia reports financial support was provided by OCNA. Coco Giovanni reports financial support was provided by project Our changing Coast.

Data availability

Data will be made available on request.

Acknowledgements

We would like to thank the SNO DYNALIT (Service National d'Observation du trait de Côte – INSU) and the OCNA (Observatoire de la Côte Nouvelle Aquitaine) for providing financial support. N.S. also thanks Dr. Jacques 2.8 for initiating and training her neuroplasticity. GC funded by the project "Our changing coast". We also thank the two reviewers for their useful comments.

References

Almar, R., Castelle, B., Ruessink, G., Sénéchal, N., Bonneton, P., Marieu, V., 2009. High-frequency video observation of a double sandbar system under high-energy wave forcing. J. Coast. Res. SI 56 (2), 1706–1710.

Almar, R., Castelle, B., Ruessink, B.G., Sénéchal, N., Bonneton, P., Marieu, V., 2010. Two- and three-dimensional double-sandbar system behaviour under intense wave forcing and a meso-macro tidal range. Cont. Shelf Res. 30, 781–792.

Almar, R., Ranasinghe, R., Sénéchal, N., Bonneton, P., Roelvink, D., Bryan, K., Marieu, V., Parisot, J.-P., 2011. Video based detection of shorelines at complex meso-macro tidal beaches. J. Coast. Res. https://doi.org/10.2112/JCOASTRES-D-10-00149.1.

Angnuureng, D.B., Almar, R., Sénéchal, N., Castelle, B., et al., 2017. Shoreline resilience to individual storms and storm clusters on a meso-macrotidal barred beach. Geomorphology 290, 265–276.

Biausque, M., Sénéchal, N., 2018. Storms impacts on a sandy beach including seasonal recovery: alongshore variability and management influences. Revue Paralia 11. https://doi.org/10.5150/revue-paralia.2018.n02.

Biausque, M., Sénéchal, N., 2019. Seasonal morphological response of an open sandy beach to winter wave conditions: the example of Biscarrosse beach, SW France. Geomorphology 332, 157–169.

Biausque, M., Sénéchal, N., 2020. Analysis of two contrasting seasonal recovery periods of an open sandy beach, using high frequency DGPS surveys. J. Coast. Res. 95 (sp1), 377–381. https://doi.org/10.2112/SI95-073.1.

Calkoen, F., Luijendijk, A., Rivero, R.C., Kras, E., Baart, F., 2021. Traditional vs. Machine-Learning methods for forecasting sandy shoreline evolution using historic satellite-derived shorelines. Remote Sens. (Basel) 13, 934. https://doi.org/10.3390/rs13050934.

Castelle, B., Bujan, S., Marieu, V., Ferreira, S., 2020. 16 years of topographic surveys of rip-channelled high-energy meso-macrotidal sandy beach. Sci. Data 7, 410. https://doi.org/10.1038/s41597-020-00750-5.

Coco, G., Sénéchal, N., Rejas, A., Bryan, K., Capo, S., Parisot, J.P., Brown, J.A., MacMahan, J.H.M., 2014. Beach response to a sequence of extreme storms. Geomorphology 204, 493–501. https://doi.org/10.1016/j.geomorph.2013.08.028.

D'Anna, M., Idier, D., Castelle, B., Vitousek, S., Le Coazannet, G., 2021. Reinterpreting the Bruun Rule in the context of equilibrium shoreline models. J. Mar. Sci. Eng. 9, 974. https://doi.org/10.3390/jmse9090974.

Gomez-de la Peña, E., Coco, G., Whittaker, C., Montaño, J., 2024. On the use of Convolutional Deep Learning to predict shoreline change. EGU sphere. https://doi.org/10.5194/egusphere-2023-958 (under review).

Harley, M.D., Kinsela, M.A., 2022. CoastSnap: a global citizen science program to monitor changing coastlines. Cont. Shelf Res. 245 https://doi.org/10.1016/j.csr.2022.104796.

Harley, M.D., Kinsela, M.A., Sanchez-Garcia, E., Vos, K., 2019. Shoreline change mapping using crowd-sourced smartphone images. Coast. Eng. 150, 175–189. https://doi.org/10.1016/j.coastaleng.2019.04.003.

Ibaceta, R., Splinter, K.D., Harley, M.D., Turner, I.L., 2022. Improving multi-decadal coastal shoreline change predictions by including model parameter non-stationarity. Front. Mar. Sci. 9, 1012041 https://doi.org/10.3389/fmars.2022.1012041.

Montaño, J., Coco, G., Antolinez, J., Beuzen, T., Bryan, K., et al., 2020. Blind testing of shoreline evolution models. Sci. Reports 10 (1). https://doi.org/10.1038/s41598-020-59018-y. Nature Publishing Group.

Montaño, J., Coco, G., Cagigal, L., Mendez, F., Rueda, A., Bryan, K., Harley, M.D., 2021. A multiscale approach to shoreline prediction. Geophys. Res. Lett. 48, e2020GL090587 https://doi.org/10.1029/2020GL090587.

Newton, A., Carruthers, T.J.B., Icely, J., 2012. The coastal syndromes and hotspots on the coast. Estuar. Coast. Shelf Sci. 96, 39–47.

Nicholls, R.J., Wong, P.P., Burkett, V.R., Codignotto, J.O., Hay, J.E., McLean, R.F., Ragoonaden, S., Woodroffe, C.D., 2007. Coastal systems and low-lying areas.

Climate Change 2007: impacts, adaptation and vulnerability. In: Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., Hanson, C.E. (Eds.), Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, pp. 315–356.

Péron, C., Sénéchal, N., 2011. Dynamic of a meso to macro-tidal doubled barred beach: inner bar response. J. Coast. Res. SI64, 120–124.

Phillips, M.S., Harley, M.D., Turner, I.L., Splinter, K.D., Cox, R.J., 2017. Shoreline recovery on wave-dominated sandy coastlines: the role of sandbar morphodynamics and nearshore wave parameters. Mar. Geol. 385, 146–159.

Reimann, L., Vafeidis, A.T., Honsel, L.E., 2023. Population development as a driver of coastal risk: current trends and future pathways. Cambridge Prisms Coast. Futures 1 (e14), 1–12. https://doi.org/10.1017/cft.2023.3.

Schepper, R., Almar, R., Bergsma, E., De Vries, S., Reniers, A., Davidson, M., Splinter, K., 2021. Modelling cross-shore shoreline change on multiple timescales and their interactions. J. Mar. Sci. Eng. 9, 582. https://doi.org/10.3390/jmse9060582.

Sénéchal, N., Gouriou, T., Castelle, B., Parisot, J.P., Capo, S., Bujan, S., Howa, H., 2009. Morphodynamic response of a meso- to macro-tidal intermediate beach based on a long-term data-set. Geomorphology 107, 263–274.

Sénéchal, N., Coco, G., Castelle, B., Marieu, V., 2015. Storm impact on the seasonal shoreline dynamics of a meso- to macrotidal open sandy beach (Biscarrosse, France). Geomorphology 228, 448–461.

Splinter, K.D., Coco, G., 2021. Challenges and opportunities in coastal shoreline prediction. Front. Mar. Sci. 8, 788657 https://doi.org/10.3389/fmars.2021.788657.

Splinter, K.D., Gonzalez, M.V.G., Oltman-shay, J., Rutten, J., Holman, R., 2018. Observations and modelling of shoreline and multiple sandbar behaviour on a high-energy meso-tidal beach. Cont. Shelf Res. 159, 33–45.

Vos, K., Harley, M.D., Splinter, K.D., Simmons, J.A., Turner, I.L., 2019. Sub-annual to multi-decadal shoreline variability from publicly available satellite imagery. Coast. Eng. 150, 160–174.

Vos, K., Splinter, K.D., Palomar-Vázquez, J., Pardo-Pascual, J.E., Almonacid-Caballer, J., Cabezas-Rabadán, C., Kras, E.C., Luijendijk, A.P., Calkoen, F., Almeida, L.P., Pais, D., 2023. Benchmarking satellite-derived shoreline mapping algorithms. Commun. Earth Environ. 4 (1), 345.

Wright, L.D., Short, A.D., 1984. Morphodynamic variability of surf zones and beaches: a synthesis. Mar. Geol. 56, 93–118.

Zeinali, S., Dehghani, M., Talebbeydokhti, N., 2021. Artificial neural network for the prediction of shoreline changes in Narrabeen, Australia. Appl. Ocean Res. 107, 102362 https://doi.org/10.1016/j.apor.2020.102362.