Blind testing of shoreline evolution models

Jennifer Montaño[1,*], Giovanni Coco[1], Jose A.A. Antolínez[2], Tomas Beuzen[3], Karin R. Bryan[4], Laura Cagigal [1,2], Bruno Castelle[5], Mark A. Davidson[6], Evan B. Goldstein[7], Raimundo Ibaceta[3], Déborah Idier[8], Bonnie C. Ludka[9], Sina Masoud-Ansari[1], Fernando J. Méndez[2], A. Brad Murray[10], Nathaniel G. Plant[11], Katherine M. Ratliff[10], Arthur Robinet[5,8], Ana Rueda[2], Nadia Sénéchal[5], Joshua A. Simmons[3], Kristen D. Splinter[3], Scott Stephens[12], Ian Townend[13], Sean Vitousek[14, 15,] Kilian Vos[3]

[1] School of Environment, Faculty of Science, University of Auckland, Auckland, 1142, New Zealand

[2] Departamento de Ciencias y Tecnicas del Agua y del Medio Ambiente, Universidad de Cantabria, Santander, Spain,

[3] Water Research Laboratory, School of Civil and Environmental Engineering, UNSW Sydney, NSW, 2052, Australia

[4] School of Science, University of Waikato, Private Bag 3105, Hamilton, New Zealand

[5] UMR EPOC, University of Bordeaux/CNRS, Bordeaux, France

[6] Coastal Processes Research Group, School of Biological and Marine Sciences, Plymouth University, Drake Circus, PL4 8AA Plymouth, UK

[7] Department of Geography, Environment, and Sustainability, University of North Carolina at Greensboro, NC 27412, USA

[8] BRGM, 3 avenue Claude Guillemin, 45060 Orléans cédex, France

[9] Scripps Institution of Oceanography, University of California, San Diego, United States

[10] Division of Earth and Ocean Sciences, Nicholas School of the Environment, Center for Nonlinear and Complex Systems, Duke University Durham, NC, USA

[11] U.S. Geological Survey St. Petersburg Coastal and Marine Science Center, 600 4th Street South, St. Petersburg, FL, USA

[12] National Institute of Water and Atmospheric Research, Hamilton, New Zealand

[13] University of Southampton, Southampton SO17 1BJ, UK

[14] Pacific Coastal and Marine Science Center, U.S. Geological Survey Santa Cruz, CA, USA

[15] Department of Civil and Materials Engineering, University of Illinois at Chicago, IL, USA

* Corresponding author: Jennifer Montaño (jmon177@aucklanduni.ac.nz)

Beaches around the world continuously adjust to daily and seasonal changes in wave and tide conditions, which are themselves changing over longer time-scales. Different approaches to predict multi-year shoreline evolution have been implemented; however, robust and reliable predictions of shoreline evolution are still problematic even in short-term scenarios (shorter than decadal). Here we show results of a modelling competition, where 19 numerical models (a mix of established shoreline models and machine learning techniques) were tested using data collected for Tairua beach, New Zealand with 18 years of daily averaged alongshore shoreline position and beach rotation (orientation) data obtained from a camera system. In general, traditional shoreline models and machine learning techniques were able to reproduce shoreline changes during the calibration period (1999-2014) for normal conditions but some of the model struggled to predict extreme and fast oscillations. During the forecast period (unseen data, 2014-2017), both approaches showed a decrease in models' capability to predict the shoreline position. This was more evident for some of the machine learning algorithms. A model ensemble performed better than individual models and enables assessment of uncertainties in model architecture. Research-coordinated approaches (e.g., modelling competitions) can fuel advances in predictive capabilities and provide a forum for the discussion about the advantages/disadvantages of available models.

Introduction

Quantitative prediction of beach erosion and recovery is essential to planning resilient coastal communities with robust strategies to adapt to erosion hazards. Over the last decades, research efforts to understand and predict shoreline evolution have intensified as coastal erosion is likely to be exacerbated by climatic changes [1–5]. The social and economic burden of changes in shoreline position are vast, which has inspired development of a growing variety of models based on different approaches and techniques; yet current models can fail (e.g. predicting erosion in accreting conditions). The challenge for shoreline models is, therefore, to provide reliable, robust and realistic predictions of change, with a reasonable computational cost, applicability to a broad variety of systems, and some quantifiable assessment of the uncertainties.

Shoreline evolution occurs over temporal scales ranging from seconds (e.g., individual waves) to hours (e.g., storms), months (e.g., seasonal wave energy modulation) and decades (e.g., wave climate). Shoreline changes occurring over much larger timescales (decadal to centennial) can be the result of other factors like longshore sediment transport gradients, changes in sediment supply, tectonic processes, anthropogenic interventions, and sea level rise (SLR)[6–8]. Cross-shore sediment transport is generally considered to be the main control of shoreline evolution at seasonal and inter-annual time-scales[9,10] whilst longshore processes (specifically on open coastlines) become more relevant over much longer timescales (decades-centuries) [7,11].

To test and improve the ability of models to predict shoreline changes, we carried out a workshop/competition on shoreline evolution modelling, "*Shoreshop"*, with participants from 15 institutions worldwide. Modelers were asked to simulate shoreline evolution obtained using a camera system at Tairua beach (New Zealand, Figure 1a,b) and submit the

results of the simulations without prior knowledge of how the shoreline actually evolved for the last 3 years (2014– 2017) of the total study period (1999 – 2017). Data in the grey shading (Figure 1,c-e) were not shown to the modelers to ensure that they were not tempted to adjust parameters of their model framework after exposure to the results.

This article summarizes the main outputs from a study where well-known models that have been broadly used in diverse study sites worldwide[12-18] combined with new approaches were compared and evaluated objectively and with no possibility of parameter tuning during the last three years of the study. As all the models are tested with the same input dataset, there is no bias associated with data sources (regardless of any inherent uncertainties in the data that was used) allowing us to objectively assess the predictive capability of the models. As shown in other disciplines[19], modelling competitions are a powerful tool to promote advances since they favour research-coordinated approaches, and, importantly, encourage the community to share datasets to assess and compare models while ensuring reproducibility which allows for objective assessment.

Data and models

In this study we concentrate on models that make use a set of cross-shore profiles or shorelines captured from aerial or oblique photography to relate changes in the shoreline position to the prevailing forcing conditions. Although 3D digital terrain models are now becoming available, there are few long-term (multi-year) data sets of this type. Figure 1 shows the entire dataset used in the study (1999-2017). Shoreline position as a function of alongshore location and time is shown in Figure 1c. Changes in the shoreline orientation, here referred to as shoreline rotation, evaluated as the slope of the trend-line fitted to the shoreline position before alongshore-averaging, are often observed (Figure 1c,e). The average alongshore position (Figure 1d) shows seasonality with progradation and retreat events generally occurring in summer and winter. The grey shading in Figure 1c-e highlights the period hidden from modelers (2014 – 2017). Three hourly wave characteristics (wave height, peak period, and direction), obtained from a wave hindcast using the SWAN model forced with Wavewatch III model, are shown in Figure 1f-h. More information about the study site characteristics and input data used during the study can be found in the Methods section.

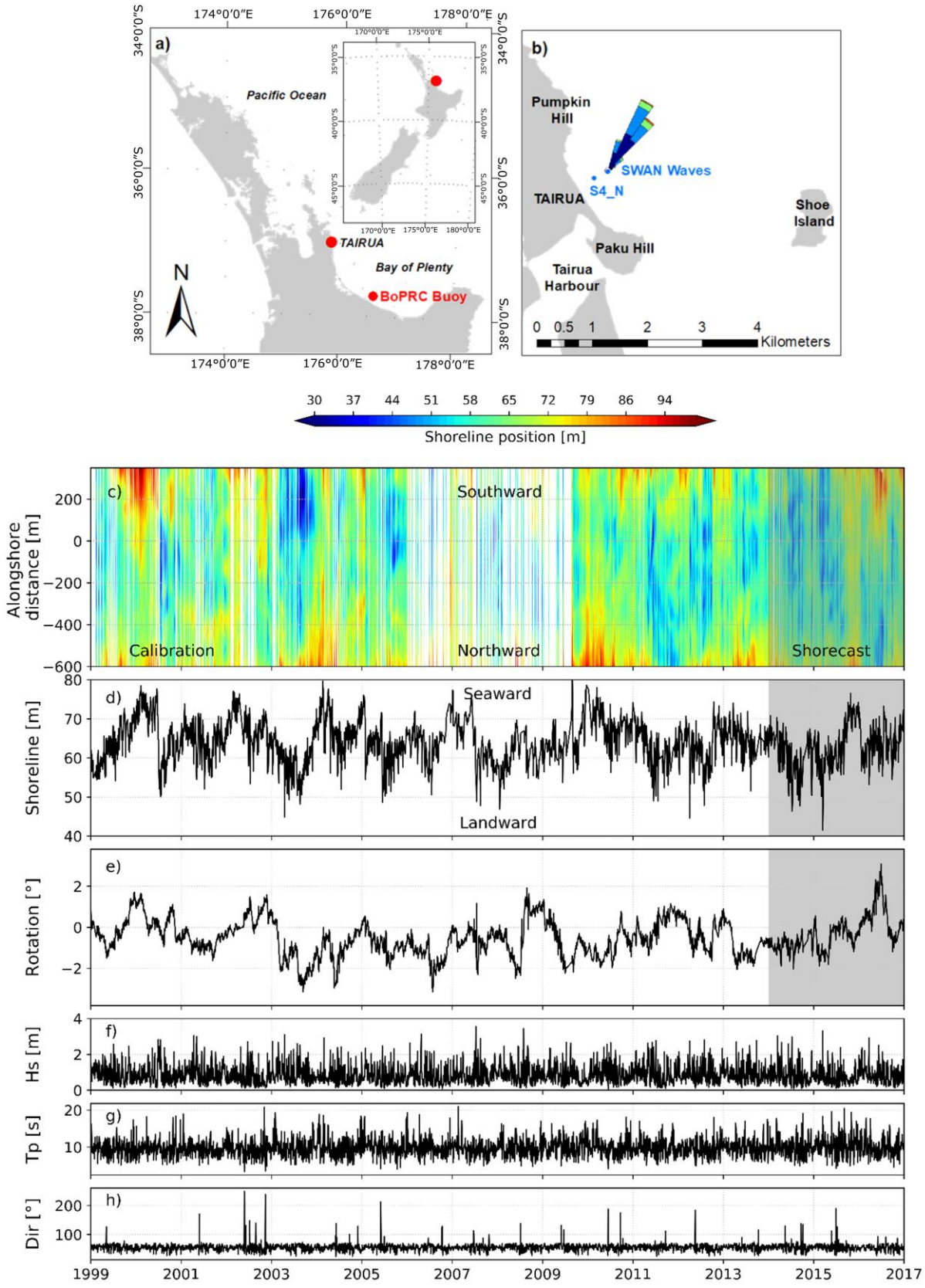

Figure 1. Study site and input conditions. a) Location of Tairua, New Zealand North Island b) Detail of Tairua Beach . Pressure sensor (S4_N) location used for SWAN model validation c) Alongshore shoreline position at Tairua beach. Red represents shoreline advance

and blue shoreline retreat over time. d) Daily alongshore-averaged position e) Shoreline rotation (orientation) with positive values representing southward accretion (anti-clockwise rotation) and negative values representing northward accretion (clockwise rotation). f) Significant wave height g) Peak period h) Wave direction. Grey shading show the data that was hidden from modelers (*Shorecast* period, 2014-2017).

We focus on daily shoreline predictions using a variety of modelling approaches (ranging from established shoreline models to Machine Learning algorithms), which, in the context of the 3 years of testing data (2014-2017), we hereafter referred to as the "*Shorecast*". A total of 19 models were used. Twelve models (indicated with HM) were built following various formulations of the well-established equilibrium concept, where the beach rate of change is governed by the difference between present and equilibrium conditions [12-13]. Some of these models were also used to predict shoreline rotation (indicated with R). Seven models were built using Machine Learning techniques (indicated with ML). Table 1 summarises the models used during the study. More information about the models can be found in the Methods section and the supporting information.

Table 1. Models used during the "Shoreshop"

Hybrid Models (HM)		
	Name	**Modeller**
HM1	ShoreFor	Kristen Splinter
HM2-R1	ShoreFor-LX	Mark Davidson
HM3	Y09-HF	Jennifer Montaño
HM4	ShoreFor+uKF	Rai Ibaceta
HM5	Y09	Bonnie Ludka
HM6, HM7	[-]	Ian Townend
HM8, R2	LX-Shore	Arthur Robinet, Bruno Castelle, Deborah Idier
HM9, R3	CosMos-Coast	Sean Vitousek
HM10, R4	COCOONED	Jose A. A. Antolinez
R5, R6	[-]	Karin Bryan
Machine Learning (ML) Models		
	Technique	**Modeller**
kNN	k- Nearest Neighbor	Evan Goldstein
ANN-EI1, 2	Autoregressive NN with exogenous inputs	Giovanni Coco
NeuFor	Artificial NN	Josh Simmons
LSTM	Long-Short Term Memory	Sina Masoud Anasari
RF	Random Forest	Tom Beuzen
BNN	Bayesian N	Nathaniel Plant

Results

Calibration period

Fifteen years (1999-2014) were used for model calibration (Figure 2), while the last three years (2014-2017) were used for the blind prediction, *Shorecast* (Figure 3), in which modelers did not have access to the shoreline data. Both HM and ML approaches were able to reproduce seasonal cross-shore (alongshore averaged) shoreline behaviour during the calibration period. Figure 2 shows the models' performance for three years (2001-2004) of the calibration period (1999-2014). Regardless of the modelling approach, oscillations of shoreline position with periods larger than 3 months were well captured. Some of the HM and almost all ML models were able to reproduce accretion periods (e.g. beginning of 2002, Figure 2a,b). In general, HM predictions were smoother than ML predictions which reproduced faster oscillations (shorter than seasonal) and more extreme events in the shoreline position (Figure 2b and Figure 4b). For instance, almost all HM underestimate the erosion that occurred on August 2003 (Figure 2a) except for the models that use data assimilation (Kalman filters, HM4, HM9). There were no major differences among HM that sought to define an equilibrium condition using: wave history (HM1, HM2, HM4 and HM8) or shoreline position (HM3, HM5, HM6, HM7, HM9 and HM10). HM models improved when Kalman filters were used (HM4 and HM9). Despite being diverse in approach and architecture, all ML models displayed very high performance during the calibration period compared with the *Shorecast* period.

Some of the HM (HM2-R1, HM8-R2, HM9-R3 and HM10-R4) and others specifically developed for shoreline rotation (R5, R6) were also used to predict beach rotation (evaluated as the slope of the trend-line fitted to the shoreline before alongshore-averaging), but no ML model was tested. Some of the rotation models essentially simulate multiple 1D cross-shore models which then require multiple calibration of cross-shore profiles along the embayment (see the Supporting Information). Almost all the models followed the rotation pattern (clockwise or anticlockwise) even during extreme rotation events. An exception is R1, which displayed a smooth behaviour (Figure 2c and qq-plot Figure 4c). Conversely, some models consistently over-estimate rotation (e.g., R4).

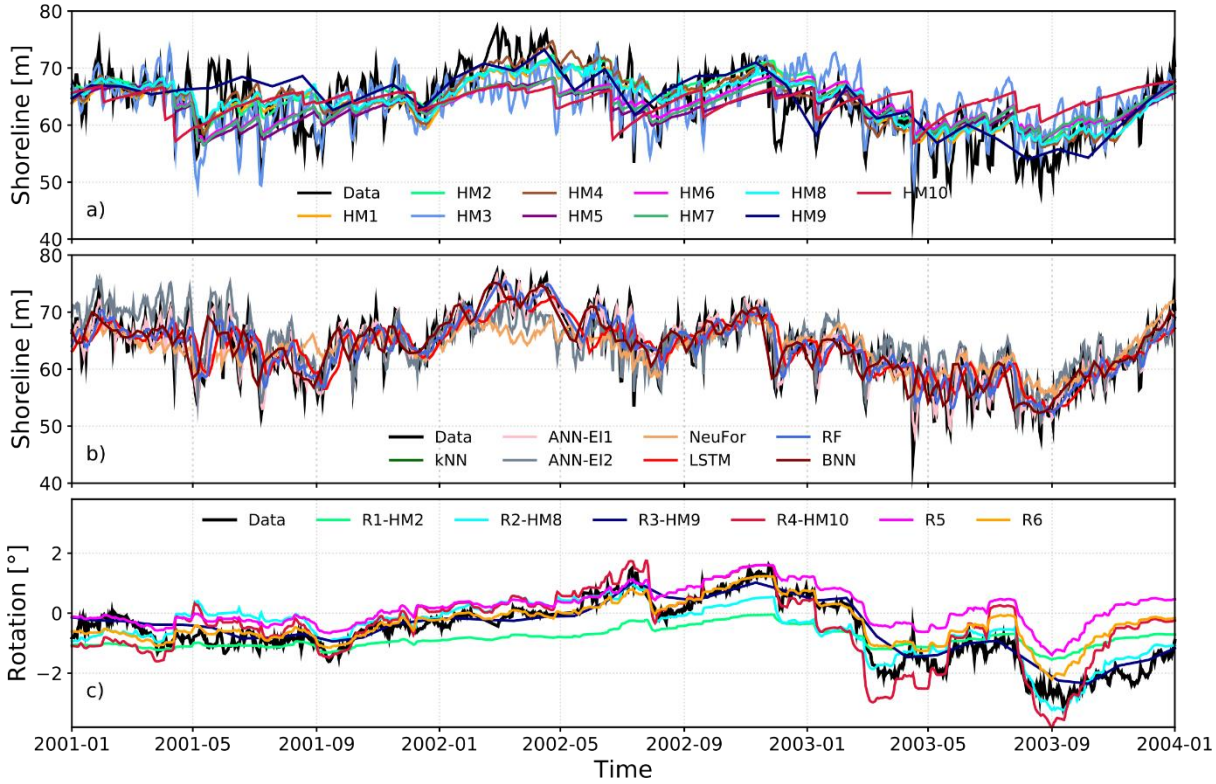

Figure 2. Three years of the entire calibration period (1999-2014). Examples of model outputs (see legend) compared to three years (2001-2004) of calibration data (black): a) Hybrid models; b) Machine Learning models; c) Shoreline rotation models. See Methods section and Supporting Information for model details.

Shorecast (Three years of blind prediction)

During the *Shorecast* (last three years of unseen data, grey shading Figure 1,c-e), HM collectively displayed similar behaviour (Figure 3a and Figure 4d) and this was distinct from the ML models (Figure 3b and Figure 4e). Both types of models were able to predict the seasonal changes. Underestimation of extreme erosion events by HM (models above dotted black line, Figure 4d) and inability of reproducing faster scale oscillations (order of 30 days) were more evident during the *Shorecast* (for example, see the beginning of 2015, Figure 3a). Some of the ML models captured extreme accretion-erosion events and faster scale oscillations not reproduced by HM. However, the models that better captured some localised large shoreline changes were also the ones that for other events produced the largest errors (Figure 3b and Figure 4e). Results from ML models changed more from the calibration to the testing phase compared to HM results (Figure 3b and Figure 4b,e), except for the HM that used Kalman filters which had a similar behaviour to the ML models. During the *Shorecast* period, the mean of the averaged alongshore shoreline position was slightly different to the mean for the calibration period (dashed grey lines in Figure 4a-f). Models tend to follow the calibration period mean shoreline position during the *Shorecast*, suggesting they are heavily dependent on the training dataset, which may indicate that they cannot predict/follow the long-term trend underlying the short term changes.

All models capture the general rotation patterns (shoreline orientation clockwise/anticlockwise) during the *Shorecast,* showing a better performance in terms of

metrics than during the calibration period, except for the models that used data assimilation (Figure 3d and Figure 4). This may suggest that models have less skill for extreme rotation events since the *Shorecast* period showed fewer and smaller beach rotation events compared with the calibration period. At times, models were able to predict shoreline rotation but underestimated/overestimated the shoreline position. Conversely, the shoreline rotation was poorly predicted at times when the erosion and accretion events were reasonably predicted. In fact shoreline advance or retreat is computed using the average alongshore shoreline position, while beach rotation (change in the shoreline orientation) considers all the alongshore transects in the trend-line fit.

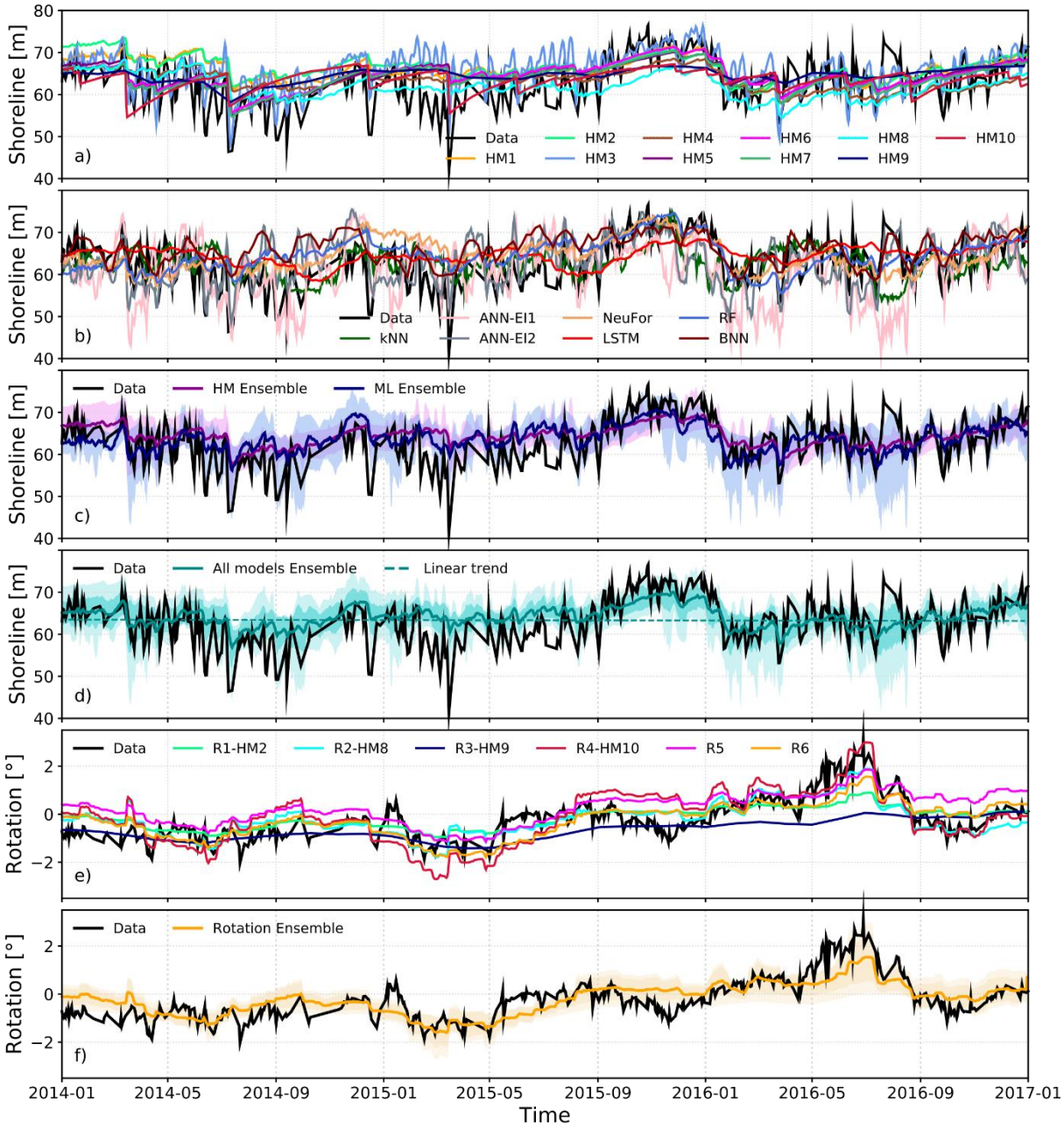

Figure 3. *Shorecast* predictions (2014-2017, blind test). Model outputs (see legends) compared to observations (black) a) Hybrid models b) Machine Learning models c) HM and ML ensemble d) Multi-model ensemble e) Rotation models f) Hybrid models ensemble for beach rotation. Dark shadows in the ensembles figures represent one standard deviation of the

models prediction. Light shadows represent maxima/minima envelope of the models predictions. See Methods section and Supporting Information for model details.

Model ensembles

Uncertainties due to the model limitations have been addressed through multi-model ensembles[20], such as, the global climate models [21,22]. An ensemble of the HM and ML models was created as a mean estimate of each type of model (HM and ML models, separately) to compare them (Figure 3c).

ML models displayed comparable skill to HM, suggesting they might be useful in describing multi-year variability at shorelines not well simulated by HM. In general, HM predictions do not capture the extremes events in shoreline position that occur over short time-scales (~monthly) exhibited by the Tairua beach data. In contrast, ML reproduced these fast oscillations and the more extreme events in the shoreline position (Figure 2b and Figure 4b). Overall, it appears that these two classes of model tend to focus on different timescales, even though time-scale is not explicitly controlled in many of the models used. Therefore, ML models and HM may play complementary role in estimating cross-shore shoreline position, due to their different approaches (inductive versus deductive). A multi-model ensemble is generated as a mean estimate of all the models (Figure 3d and dashed black line in Figure 4d,e). The total multi-model ensemble *Shorecast* often overlaps the shoreline data, showing capacity to predict seasonality and some extreme events, for instance, accretion (end of 2015) and erosion (beginning of 2016). When all models reproduce the measured shoreline position correctly, the average of the models converges (low standard deviation), while when some of the models diverge from the measured shoreline, the ensemble cancels out the possibility of a large error. In general, ensembles (HM, ML and all models ensemble) showed better performance than many individual models (Figure 3c,d, Figure 4 and Supporting Information-Table 1). Even though an ensemble approach may increase model complexity and might smooth the predictions, Figure 4d,e show that the total model ensemble captured extreme events (erosion/accretion) better than almost all HM and some of the ML models. Therefore, the ensemble approach improves the reliability of the predictions, and in effect reduces model-related uncertainty (Figure 3d,f).

Discussion

Assessment of models performance.

We calculated different metrics to assess predictive model performance (Figure 4 g-n and Supporting Information). We included a linear trend as a predictive model and, even though the linear trend does not follow the shoreline oscillations, metrics like R^2, RMSE or Skill were better than for some of the models (Figure 4 and Supporting Information Table 1).

The best metric for assessing model performance remains unclear and different model performance metrics favoured different models, highlighting the importance of considering multiple metrics and different approaches for a robust model evaluation. Model performance was also assessed in terms of quantile-quantile (predictions vs measurements) which provides information about extremes events, the direction of shoreline change (erosion/accretion) and

mean behaviour (Figure 4 a-f). In addition, we acknowledge that the usefulness of a model should not be expressed only in terms of metrics but also reproducibility and understanding that leads to scientific advancements.

In general, models showed a lower performance during the *Shorecast* period than during the calibration period, and lower performance compared to previous studies in different sites where the models were first presented and tested against a data set [12-18]. The exact reason is difficult to determine but it is evident that, despite data uncertainties, true predictions of unseen shoreline data remains a difficult task.

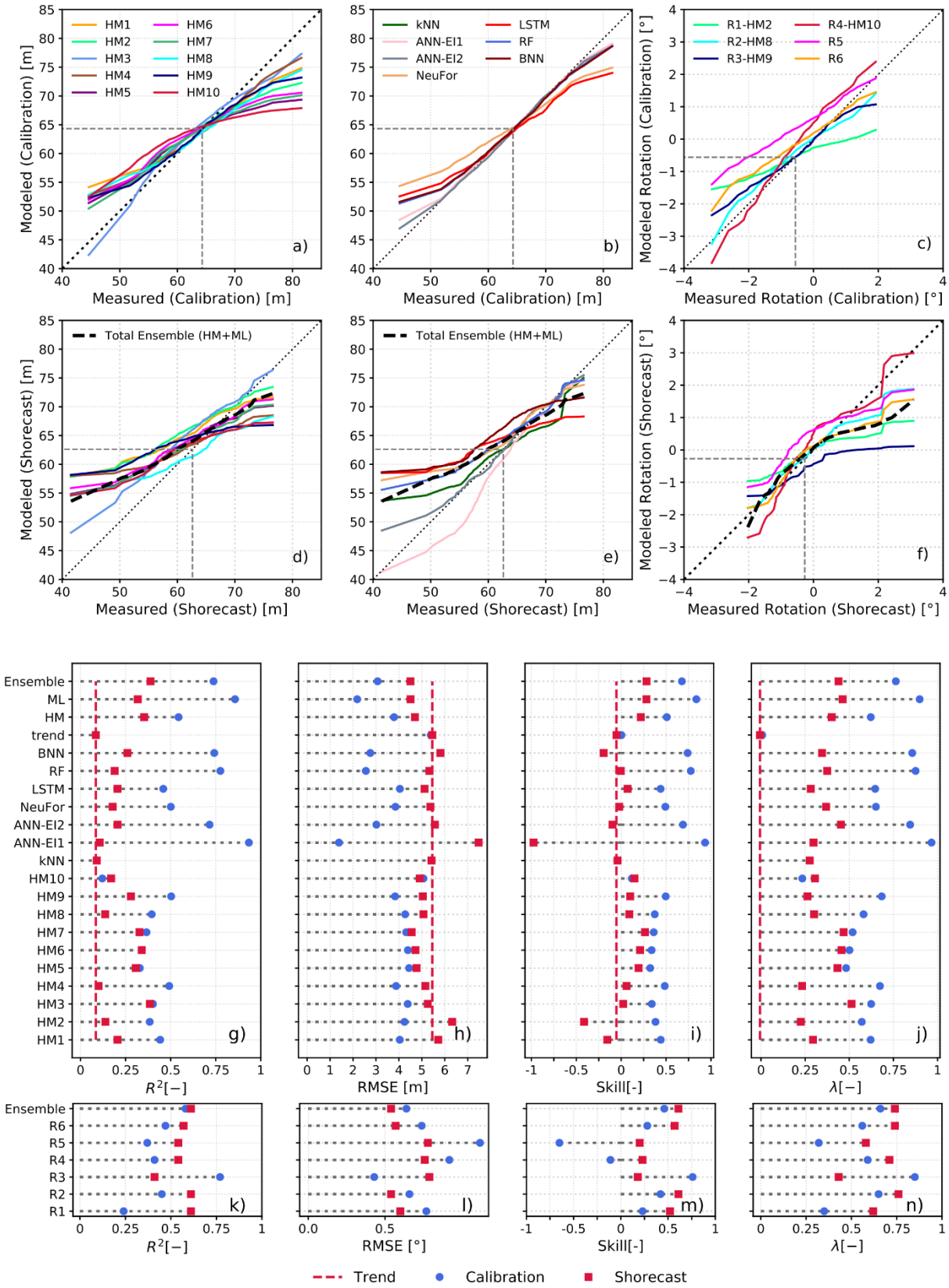

Figure 4. Models performance. Quantile-quantile plots of model behavior. Top 3 panels: Calibration period; bottom 3 panels: *Shorecast*. Model prediction vs measured shoreline position for a) and d) HM; b) and e) ML; c) and f) model prediction vs measured shoreline

rotation. Dashed grey line represent the average shoreline position during the calibration and the *Shorecast* period, respectively. R^2, RMSE, skill and λ for shoreline prediction for; g-j) averaged shoreline position; k–n) shoreline rotation. See supporting material for more information about the metrics and individual models.

Uncertainty

Uncertainties are a key component of any modelling study. During the *Shorecast*, uncertainty arises from both the shoreline position and wave characteristics. Wave characteristics have been obtained through a numerical model and so contain potential sources of error. This could affect models differently since some models only use wave height and others include also the wave period which in general is more difficult to reproduce. Despite the wave period being poorly reproduced when compared to reproducibility of the wave height (See Methods), this did not seem to affect models' performance during the calibration. Due to the stochastic nature of waves, probabilistic approaches using synthetic hydrodynamic forcing may give more realistic shoreline predictions that account for uncertainties, especially for long-term projections[23,24] . The shoreline positions we provided are likely to contain detection inaccuracies, although most of the errors occur over a scale shorter than a week[25]. Large shoreline changes, order of 30 days, were observed and these changes are only marginally affected by the faster-scale inaccuracies [25]. This is relevant since a number of models managed to reproduce instances of rapid and large shoreline retreat, but others completely missed the fast, order of 30 days, shoreline changes.

A different kind of uncertainly involves model structure and parameterizations. Such uncertainties in shoreline models arise because we use simplified models that may ignore some of the physical processes; for example, HM used during the *Shorecast* lack processes including, for example, overwash, beach-dune and/or beach-cliff interactions, influence of bars, human interventions which may play an important role in shoreline evolution. Also, we use a variety of ML techniques to find hidden processes and relations among drivers and response since, *a priori*, it is uncertain which ML model works best for the dataset provided. There is also an uncertainty related to parameters. While some models used only a minimal number of parameters (e.g., equilibrium models), others included many more (e.g., ML models) which brings up questions about model generality. Our test indicates that, even when model structure is similar, results may differ because models are highly dependent on a range of parameters that the scientist chooses.

Long-term predictions

Uncertainty in drivers (e.g., waves and SLR) and therefore shoreline response increases as longer time horizons are explored. During *"Shoreshop"*, we also attempted to predict Tairua shoreline evolution until 2100. Results are not presented here because uncertainties in the future wave climate were deemed too large and the model outputs diverged drastically. However, the exercise was educational in two ways. First, it highlighted that shoreline evolution models require calibration and often fail when implemented outside of the regime of calibration. Secondly, the processes causing long-term shoreline change can differ from those that produce seasonal to multi-annual oscillations in shoreline position or plan-view

shape (as addressed in the calibration and *Shorecast* periods). For an alongshore-restricted beach like Tairua, cross-shore processes related to SLR are likely to be the main cause of cumulative shoreline change. However, many of the models presented do not include the effects of SLR. We have thus far discussed creating ensembles for shoreline prediction, but ensembles are also needed for the drivers such as, SLR, wave conditions, storm surge where different scenarios are considered [8,23,24].

Comparisons between multiple models on additional datasets (e.g. longshore vs cross-shore transport or seasonal vs storm dominance) are recommended to address issues in long term predictions that include (a) downscaling an appropriate future wave climate using a probabilistic approach that allows uncertainties to be accounted for, (b) assessing morphodynamic implications of different scenarios of SLR, and (c) dealing with the effect of out-of-calibration parameters. This will help better differentiate model performance, including transportability to different wave climates and morphodynamic settings. We foresee an increase of this type of modelling competition as a way to accelerate knowledge exchange and dissemination, as well as fostering community interaction.

Methods

Study site

Tairua Beach is located in the Coromandel peninsula, on the east coast of the North Island of New Zealand (Figure 1a,b). Tairua is a pocket beach of 1.2 km long, with medium to coarse sand that exhibits intermediate beach states[25-27]. The lower shoreface slope is approximately 0.02, whereas the upper beach slope is steep ≈ 0.2 [28]. The beach is located in a microtidal environment with a tidal range varying between 1.2 and 2 m. Eighteen years of daily shoreline evolution (1999-2017) were obtained using a camera system located on a hill (elevation about 60 m) at the north end of the beach. During daylight, six hundred images were averaged over a period of 15 min every hour. The time-averaged images were then georectified and used to extract the shoreline position. To limit the influence of the tides, daily shoreline images with tidal level between 0.45 and 0.55 m were selected. Errors in the shoreline detection due to the footprint of the georectified images, standard deviation of the water levels, the influence of the tides, the uncertainty in wave setup, and other noise have been shown to affect the daily timescale but not the shoreline signal over weekly (and longer) timescales[25]. The wave characteristics (at 10 m water depth) were obtained using a hydrodynamic model (SWAN), validated with in situ measurements in 8 m water depth (indicated as S4_N in Figure 1b). The comparison was good in terms of wave height ($R^2 = 0.80$ and $RMSE = 0.31\ m$) but the wave period was poorly reproduced ($R^2 = 0.22$ and $RMSE = 2.9\ s$). The instrument used did not record wave direction so that a direct comparison could not be made.

Model classification

We have classified shoreline models as: process-based models (PBM), hybrid models (HM), and data-driven models (DDM). The PBM are simulation models or physics-based models, which include as many processes as practicable, and usually couple hydrodynamics, waves, sediment transport and morphology through mass and momentum conservation laws. In general, these models attempt to describe the faster (storm event) and smaller scale (<km)

processes and there is no conclusive evidence they can be successfully applied over large spatio-temporal scales. Also, in many cases, these models require input data that were not available for the present study.

We use the term "hybrid" to characterize shoreline models based on general principles (e.g., that a system is drawn towards an equilibrium configuration) that do not use detailed conservation of mass and momentum equations , and rely heavily on a data-driven approach to find the free parameters of the model. HM often base the prediction of the cross-shore position on the equilibrium concept [29], where the beach rate of change is governed by the difference between present and equilibrium conditions. Equilibrium conditions have been defined in terms of shoreline position [10,12] or wave history[13]. These models and similar variants have been applied successfully when addressing seasonal to interannual variability at many sites[14-17], however they may fail to simulate the shoreline evolution in environments where other processes such as alongshore sediment transport play an important role[3].

Additionally, the equilibrium concept has also been successfully applied to predict shoreline[18] and sandbar rotation[27] at pocket beaches. Other longshore transport models[11] have been applied to long term datasets but lacked the ability to reproduce cross-shore variations [30]. This issue was recently addressed[31–33], where alongshore sediment transport formulae such as, CERC[34,35] are combined with cross-shore equilibrium models. Table 1 provides a summary of the HM used during the Shoreshop.

Due to the surge of available measurements characterized by increasing spatial and temporal resolution from camera systems to satellite images[36,37] and novel approaches to modelling, DDM have become more popular. Examples of these models range from simple autoregressive models to machine learning (ML) techniques such as artificial neural networks. The use of ML techniques in a variety of coastal problems and settings has rapidly increased over the past few years[38], since ML algorithms can be highly effective predictors[39,40], can be used as part of larger models[41] and can provide physical insight[42]. Statistical models such as multiple linear regression or statistical downscaling[43,44], also fall in the category of data-driven, but have not been tested in the *Shorecast*. One of the drawbacks of DDM is that their performance depends highly on the quantity and quality of the data available. The ML techniques used during the Shoreshop are listed in Table 1. Details on *ML* models are provided in the Supporting Information.

Data availability.

All data provided to the participants of the "Shoreshop" is available at https://coastalhub.science/data. Upon acceptence of the manuscript this data will be archived in a FAIR repository (i.e., figshare), made public, and be given a DOI. The data will also be provided in a figshare deposit once the manuscript is accepted. Other data that support the findings of this study are available from the indicated sources or from the corresponding author upon reasonable request.

References

1. Church, J. A. & White, N. J. A 20th century acceleration in global sea-level rise. *Geophys. Res. Lett.* **33,** 94–97 (2006).

2. Nicholls, R. J. *et al.* Sea-level scenarios for evaluating coastal impacts. **5,** (2014).

3. Dodet, G. *et al.* Beach recovery from extreme storm activity during the 2013/14 winter along the Atlantic coast of Europe. *Earth Surf. Process. Landforms* (2018). doi:10.1002/esp.4500

4. Burvingt, O., Masselink, G., Scott, T., Davidson, M. & Russell, P. Climate forcing of regionally-coherent extreme storm impact and recovery on embayed beaches. *Mar. Geol.* **401,** 112–128 (2018).

5. Reguero, B. G., Losada, I. J. & Méndez, F. J. A recent increase in global wave power as a consequence of oceanic warming. *Nat. Commun.* **10,** 1–14 (2019).

6. Bruun. Sea-Level Rise as a Cause of Shore Erosion. *J. Waterw. Harb. Div.* **88,** 117–132 (1962).

7. Hanson, H. Genesis-A Generalized Shoreline Change Numerical Model. *J. Coast. Res.* **5,** 1–27 (1989).

8. Le Cozannet, G. *et al.* Quantifying uncertainties of sandy shoreline change projections as sea level rises. *Sci. Rep.* **9,** 1–11 (2019).

9. Kriebel, D. & Dean, R. G. Numerical simulation of time-dependent beach and dune erosion. *Coast. Eng.* **9,** 221–245 (1985).

10. Miller, J. K. & Dean, R. G. A simple new shoreline change model. *Coast. Eng.* **51,** 531–556 (2004).

11. Ashton, A., Murray, A. B. & Arnoult, O. Formation of coastline features by large-scale instabilities induced by high-angle waves. *Nature* **414,** 296 (2001).

12. Yates, M. L., Guza, R. T. & O'Reilly, W. C. Equilibrium shoreline response: Observations and modeling. *J. Geophys. Res. Ocean.* **114,** 1–16 (2009).

13. Davidson, M. A., Splinter, K. D. & Turner, I. L. A simple equilibrium model for predicting shoreline change. *Coast. Eng.* **73,** 191–202 (2013).

14. Splinter, K. D. *et al.* A generalized equilibrium model for predicting daily to interannual shoreline response. *J. Geophys. Res. F Earth Surf.* **119,** 1–23 (2014).

15. Castelle, B. *et al.* Equilibrium shoreline modelling of a high-energy meso-macrotidal multiple-barred beach. *Mar. Geol.* **347,** 85–94 (2014).

16. Ludka, B. C., Guza, R. T., O'Reilly, W. C. & Yates, M. L. Field evidence of beach profile evolution toward equilibrium. *J. Geophys. Res. Ocean.* **120,** 7574–7597 (2015).

17. Lemos, C. *et al.* Equilibrium modeling of the beach profile on a macrotidal embayed low tide terrace beach. *Ocean Dyn.* **68,** 1207–1220 (2018).

18. Turki, I., Medina, R., Coco, G. & Gonzalez, M. An equilibrium model to predict shoreline rotation of pocket beaches. *Mar. Geol.* **346,** 220–232 (2013).

19. Weigend, A. S. Paradigm change in prediction. *Philos. Trans. R. Soc. London. Ser. A Phys. Eng. Sci.* **348,** 405–420 (1994).

20. Limber, P. W., Barnard, P. L., Vitousek, S. & Erikson, L. H. A Model Ensemble for Projecting Multidecadal Coastal Cliff Retreat During the 21st Century. *J. Geophys. Res. Earth Surf.* 1566–1589 (2018). doi:10.1029/2017JF004401

21. Tebaldi, C. & Knutti, R. The use of the multi-model ensemble in probabilistic climate projections. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **365,** 2053–2075 (2007).

22. Buchanan, M. Ignorance as strength. *Nat. Phys.* **14,** 41567 (2018).

23. Ranasinghe, R., Callaghan, D. & Stive, M. J. F. Estimating coastal recession due to sea level rise: Beyond the Bruun rule. *Clim. Change* **110,** 561–574 (2012).

24. Davidson, M. A., Turner, I. L., Splinter, K. D. & Harley, M. D. Annual prediction of shoreline erosion and subsequent recovery. *Coast. Eng.* **130,** 14–25 (2017).

25. Blossier, B., Bryan, K. R., Daly, C. J. & Winter, C. Shore and bar cross-shore migration, rotation, and breathing processes at an embayed beach. *J. Geophys. Res. Earth Surf.* **122,** 1745–1770 (2017).

26. van Maanen, B., de Ruiter, P. J., Coco, G., Bryan, K. R. & Ruessink, B. G. Onshore sandbar migration at Tairua Beach (New Zealand): Numerical simulations and field measurements. *Mar. Geol.* **253,** 99–106 (2008).

27. Blossier, B., Bryan, K. R., Daly, C. J. & Winter, C. Nearshore sandbar rotation at single-barred embayed beaches. *J. Geophys. Res. Ocean.* 1063–1084 (2016). doi:10.1002/2015JC010796.Received

28. Smith, R. K. & Bryan, K. R. Monitoring Beach Face Volume with a Combination of Intermittent Profiling and Video Imagery. *J. Coast. Res.* **234,** 892–898 (2007).

29. Wright, L. D., Short, A. D. & Green, M. O. Short-term changes in the morphodynamic states of beaches and surf zones: An empirical predictive model. *Mar. Geol.* **62,** 339–364 (1985).

30. Ruggiero, P., Buijsman, M. & Kaminsky, G. M. Gelfenbaum, G. Modeling the effects of wave climate and sediment supply variability on large-scale shoreline change. *Mar. Geol.* **273,** 127–140 (2010).

31. Vitousek, S., Barnard, P. L., Limber, P., Erikson, L. & Cole, B. A model integrating longshore and cross-shore processes for predicting long-term shoreline response to climate change. *J. Geophys. Res. Earth Surf.* 1–25 (2017). doi:10.1002/2016JF004065

32. Robinet, A., Idier, D., Castelle, B. & Marieu, V. A reduced-complexity shoreline change model combining longshore and cross-shore processes: The LX-Shore model. *Environ. Model. Softw.* **109,** 1–16 (2018).

33. Antolínez, J. A. A., Méndez, F. J., Anderson, D., Ruggiero, P. & Kaminsky, G. M. Predicting climate driven coastlines with a simple and efficient multi-scale model. *J. Geophys. Res. Earth Surf.* 2018JF004790 (2019). doi:10.1029/2018JF004790

34. USAGE. *Shore protection manual*. (1984). doi:doi.org/ 10.5962/bhl.title.47829

35. Kamphuis, J. W. Alongshore Sediment Transport Rate. *J. Waterw. Port, Coastal, Ocean Eng.* **117,** 624–640 (1991).

36. Luijendijk, A. *et al.* The State of the World's Beaches. *Sci. Rep.* **8,** 6641 (2018).

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

38. Goldstein, E., Coco, G. & Plant, N. G. A review of machine learning applications to coastal sediment transport and morphodynamics. *Earth-Science Rev.* **194,** 97–108 (2019).

39. Passarella, M., Goldstein, E. B., De Muro, S. & Coco, G. The use of genetic programming to develop a predictor of swash excursion on sandy beaches. *Nat. Hazards Earth Syst. Sci.* **18,** 599–611 (2018).

40. Beuzen, T. *et al.* Bayesian Networks in coastal engineering: Distinguishing descriptive and predictive applications. *Coast. Eng.* **135,** 16–30 (2018).

41. Goldstein, E. B. & Coco, G. Machine learning components in deterministic models: hybrid synergy in the age of data. *Front. Environ. Sci.* **3,** 1–4 (2015).

42. Tinoco, R., Goldstein, E. . & Coco, G. A data-driven approach to develop physically sound predictors: Application to depth-averaged velocities on flows through submerged arrays of rigid cylinders. *Water Resour. Res.* 1247–1263 (2015). doi:10.1002/2014WR016380.Received

43. Callaghan, D. P., Ranasinghe, R. & Roelvink, D. Probabilistic estimation of storm erosion using analytical, semi-empirical, and process based storm erosion models. *Coast. Eng.* **82,** 64–75 (2013).

44. Anderson, D., Ruggiero, P., Antolínez, J. A. A., Méndez, F. J. & Allan, J. A Climate Index Optimized for Longshore Sediment Transport Reveals Interannual and Multidecadal Littoral Cell Rotations. *J. Geophys. Res. Earth Surf.* **123,** 1958–1981 (2018).

Acknowledgments

Funding from the Auckland Hazard Hub and a project awarded to GC (Climate change impacts on weather-related hazards) is gratefully acknowledged. Thanks to Waikato Regional

Council and NIWA for providing the video images, to R. Bell (NIWA) for the tide and SLR data, and to MetOcean for the wave hindcast. Thanks to the Centre for eResearch of the University of Auckland for providing computational resources through the Nectar Research cloud. Shoreline detection was based on scripts provided by B. Blossier and C. Daly. Thanks also for the ANR-Carnot funding for the BRGM contribution. We thank Joe Long for providing a USGS internal review of this manuscript

Author Contributions

J.M and G.C. conceived the competition, organized the workshop and prepared the first draft of the manuscript. JM performed all data analysis and prepared all figures. All co-authors participated in the Shoreshop running their models and participated to the writing of the manuscript.

Additional information

The authors declare no competing interests.

Supplementary Information is provided

Figure 1. Study site and input conditions. a) Location of Tairua, New Zealand North Island b) Detail of Tairua Beach . Pressure sensor (S4_N) location used for SWAN model validation c) Alongshore shoreline position at Tairua beach. Red represents shoreline advance and blue shoreline retreat over time. d) Daily alongshore-averaged position e) Shoreline rotation (orientation) with positive values representing southward accretion (anti-clockwise rotation) and negative values representing northward accretion (clockwise rotation). f) Significant wave height g) Peak period h) Wave direction. Grey shading show the data that was hidden from modelers (*Shorecast* period, 2014-2017).

Figure 2. Three years of the entire calibration period (1999-2014). Examples of model outputs (see legend) compared to three years (2001-2004) of calibration data (black): a) Hybrid models; b) Machine Learning models; c) Shoreline rotation models. See Methods section and Supporting Information for model details.

Figure 3. *Shorecast* predictions (2014-2017, blind test). Model outputs (see legends) compared to observations (black) a) Hybrid models b) Machine Learning models c) HM and ML ensemble d) Multi-model ensemble e) Rotation models f) Hybrid models ensemble for beach rotation. Dark shadows in the ensembles figures represent one standard deviation of the

models prediction. Light shadows represent maxima/minima envelope of the models predictions. See Methods section and Supporting Information for model details.

Figure 4. Models performance. Quantile-quantile plots of model behavior. Top 3 panels: Calibration period; bottom 3 panels: *Shorecast*. Model prediction vs measured shoreline position for a) and d) HM; b) and e) ML; c) and f) model prediction vs measured shoreline rotation. Dashed grey line represent the average shoreline position during the calibration and the *Shorecast* period, respectively. R^2, RMSE, skill and λ for shoreline prediction for; g-j) averaged shoreline position; k–n) shoreline rotation. See supporting material for more information about the metrics and individual models.

Table 1. Models used during the "Shoreshop"

Supporting Information for

Blind testing of shoreline evolution models

Jennifer Montaño[1,*], Giovanni Coco[1], Jose A.A. Antolínez[2], Tomas Beuzen[3], Karin R. Bryan[4], Laura Cagigal [1,2], Bruno Castelle[5], Mark A. Davidson[6], Evan B. Goldstein[7], Raimundo Ibaceta[3], Déborah Idier[8], Bonnie C. Ludka[9], Sina Masoud-Ansari[1], Fernando J. Méndez[2], A. Brad Murray[10], Nathaniel G. Plant[11], Katherine M. Ratliff[10], Arthur Robinet[5,8], Ana Rueda[2], Nadia Sénéchal[5], Joshua A. Simmons[3], Kristen D. Splinter[3], Scott Stephens[12], Ian Townend[13], Sean Vitousek[14, 15,] Kilian Vos[3]

[1] School of Environment, Faculty of Science, University of Auckland, Auckland, 1142, New Zealand
[2] Departamento de Ciencias y Tecnicas del Agua y del Medio Ambiente, Universidad de Cantabria, Santander, Spain,
[3] Water Research Laboratory, School of Civil and Environmental Engineering, UNSW Sydney, NSW, 2052, Australia
[4] School of Science, University of Waikato, Private Bag 3105, Hamilton, New Zealand
[5] UMR EPOC, University of Bordeaux/CNRS, Bordeaux, France
[6] Coastal Processes Research Group, School of Biological and Marine Sciences, Plymouth University, Drake Circus, PL4 8AA Plymouth, UK
[7] Department of Geography, Environment, and Sustainability, University of North Carolina at Greensboro, NC 27412, USA
[8] BRGM, 3 avenue Claude Guillemin, 45060 Orléans cédex, France
[9] Scripps Institution of Oceanography, University of California, San Diego, United States
[10] Division of Earth and Ocean Sciences, Nicholas School of the Environment, Center for Nonlinear and Complex Systems, Duke University Durham, NC, USA
[11] U.S. Geological Survey St. Petersburg Coastal and Marine Science Center, 600 4th Street South, St. Petersburg, FL, USA
[12] National Institute of Water and Atmosphere, Hamilton, New Zealand
[13] University of Southampton, Southampton SO17 1BJ, UK
[14] Pacific Coastal and Marine Science Center, U.S. Geological Survey Santa Cruz, CA, USA
[15] Department of Civil and Materials Engineering, University of Illinois at Chicago, IL, USA

Here we provide a brief description of the models used during the *Shoreshop* and a summary of the models' performance assessment based on different metrics.

Hybrid(HB) Models

ShoreFor (HM1)

ShoreFor is an equilibrium-based cross-shore model first presented in Davidson et al.[1]. The model formulation used in this work follows the modifications of Splinter et al. [2] allowing for a more general equilibrium model with inter-site variability of model coefficients. The model formulation follows:

$$\frac{dY}{dt} = c(F^+ + r\,F^-) + b$$

Where $\frac{dY}{dt}$ is the rate of shoreline change, dependent on the magnitude of wave forcing F defined as:

$$F = P^{0.5}\,(\Omega_{\emptyset} - \Omega)/\sigma,$$

where P is the breaking wave energy flux and Ω is the dimensionless fall velocity. The model includes two coefficients. The first one, c which is the rate parameter accounting for the efficiency of cross-shore sediment transport and ∅ which defines the window width of a filter function, performing a weighted average of the antecedent dimensionless fall velocity and is a proxy for the 'beach memory'. The model contains two constants, $r = \sum F^+ / \sum F^-$ and σ which is the standard deviation of $(\Omega_{\emptyset} - \Omega)$, both computed over the calibration segment of the wave data. The linear trend parameter, b, has been included to simplistically account for longer-term processes (e.g. longshore sediment transport, sediment supply, etc) not explicitly accounted in the model. The model is calibrated by choosing the minimum normalized mean square error (NMSE) of a least-squares regression solving for c, and b for different values of ∅ in the range of 5 to 1000 days.

ShoreFor_LX (HM2-R1)

The ShoreFor_LX model presented here aims to provide a more physical representation of the impact of longshore sediment transport gradients on shoreline change than ShoreFor and is a simple variation of the ShoreFor model described above with the inclusion a CERC-type rotational term in place of the trend term (b):

$$\frac{dY}{dt} = c_1(F^+ + r\,F^-) + c_2\frac{dF_2}{dx}$$

Here x is the longshore direction and F2 is the longshore forcing term defined as $P\,sin(2\alpha)$, where α is the angle of wave incidence. This version of the ShoreFor model has an equivalent number of free parameters to the original version (c1, c2 and ∅). The model equation is integrated using a standard one-line algorithm[3].

Y09-HF(HM3)

The model is a modification of Yates et al.[4] to predict faster shoreline changes. The Ensemble Empirical Mode Decomposition (CEEMD) method[5,6], designed to identify

oscillations in non-linear and non-stationary time-series, was applied to the time-series of waves and shoreline position. The Yates model was then applied only to modes with a timescale larger than 60 days, while faster oscillations (between 10-60 days) were considered adding a new term to the Yates model (which involves adding a new free parameter):

$$\frac{dY}{dt} = C^{\pm}E^{0.5}\left(E - E_{eq}\right) + C_{hf}E_{hf}^{0.5}$$

Where the subscript hf represents the high frequency oscillations obtained with the CEEMD method. During the calibration, coefficients were selected by minimizing RMSE.

Shorefor + uKF (HM4)

This model uses a joint unscented Kalman Filter[7] within the ShoreFor model (HM1) to assess the time-varying free parameters (b, c, φ) that best fit the observed shorelines during the calibration period. A state vector, here represented by the observed shorelines and the ShoreFor paramaters are sampled through sigma points in a way that their mean and covariance are maintained[8]. The nonlinear system (here given by the original ShoreFor formulation) is used to propagate each sigma point at the prediction step, weighting these propagations to obtain an accurate estimation of the state vector mean and covariance. Then, an update step combines the propagated non-linearities with observed shoreline measurements to optimally estimate the state vector when observed shorelines become available.

Following[9], we define the total shoreline position $Y(t)$ by the contribution of short-term processes Y_{st} and long-term (or unresolved) processes Y_{lt}, thus linking Y_{st} with the equilibrium formulation and Y_{lt} with the b coefficient.

Yates et al., 2009 - *3 Coefficients* (HM5)

HM5 uses the formulation of Yates model, but with a single scalar value for the rate coefficient C. (Yates model used a different value of C for erosion and accretion, $C^{\mp}$)

Based on the equilibrium framework suggested by [10,11] and similar to Miller and Dean model [12], the shoreline change rate depends on both the incident wave energy and the difference between the incident wave energy and the equilibrium energy

$$\frac{dY}{dt} = CE^{0.5}(E - E_{eq})$$

where the equilibrium energy is defined using a linear relationship with the present shoreline position,

$$E_{eq} = aY + b$$

The three free parameters (a, b, and C) are optimized by minimizing RMSE, using simulated annealing. This model attempts to minimize RMSE while minimizing the number of free parameters, minimizing the number of independent variables (e.g. only wave height is used and not wave period), and minimizing equation complexity (e.g. uses a linear relation for equilibrium energy rather than a quadratic).

Yates et al, 2009-*4 Coefficients* (HM6, HM7)

The fitting and hindcasting of shoreline position were done using a derivative of the Kriebel and Dean model[13], as proposed by Yates et al.[4]. An offset to the initial shoreline was included as an additional fit parameter, as proposed by [14]. The model relates nearshore wave energy, (represented by $\frac{H_s^2}{16}$), to shoreline position, Y. Inshore wave data were obtained using plane bed refraction to account for refraction and shoaling to the closure depth of the beach profile, taking account of water levels variations due to the tide.

Fitting of the parameters is sensitive to the initial guess and search bounds. Two of the model parameters (a and b) are estimated by finding the line for $dY/dt = 0$ from the energy between surveys and the shoreline anomaly (dY) and this was modified to use only the points that were close to $dY/dt = 0$. The accretion and erosion rates, C^+ and C^-, were estimated from the mean positive and negative values of dY/dt respectively, with an initial offset of zero.

Parameter fitting was done using derivative free, constrained non-linear optimisation. Matlab functions for particle swarm, simulated annealing and simplex optimisation were used. Statistically, there was little to choose between them and the results reported used the simplex method. Selection of the best fit was based on minimising the RMSE with the slope of the modelled v observed regression line as close to 1 as possible. In all model runs the large erosional anomalies were better represented than the accretional anomalies, suggesting that the model is not capturing the full extent of beach recovery.

LX-Shore (HM8, R2)

LX-Shore is a two-dimensional plan-view cellular-based one-line shoreline change model for wave-dominated sandy coasts presented first in Robinet et al.[15]. The model can simulate shoreline change resulting from gradients in total longshore sediment transport and/or from cross-shore transport driven by the variability in incident wave energy. LX-Shore can handle complex shoreline geometries (e.g. sand spits, islands), including non-erodible areas such as coastal defences and headlands, and is coupled with the spectral wave model SWAN[16] to cope with complex nearshore wave fields. For the present application, the longshore sediment transport along the coastline was computed using the formula of [17] with a calibration multiplication factor (fQl) tuned to minimize the RMSE of beach orientation at Tairua over 1999-2013. The cross-shore transport is resolved using an adaptation of the ShoreFor model[1,2] where the disequilibrium term is computed from offshore wave conditions instead of breaking wave conditions. The three ShoreFor free parameters (φ, c and b) were optimized using a simulated annealing procedure minimizing the RMSE of alongshore-averaged cross-shore shoreline position at Tairua over 1999-2013. During the simulation, waves were propagated onshore using SWAN with default processes and parameter values. Overall, only 4 calibration parameters were tuned to simulate shoreline change along the entire embayment.

CosMos-Coast (HM9, R3)

CoSMoS-COAST is a "one-line" model that integrates longshore and cross-shore transport processes to predict long-term coastal change presented in Vitousek et al.[18]. The model synthesizes several popular process-based shoreline change models including (1) a "one-line" model for longshore transport[19], (2) a wave-driven cross-shore equilibrium shoreline

change model[4], (3) a cross-shore equilibrium beach profile change model due to sea-level rise[20–22], and a residual term obtained via data assimilation. The governing equation of CoSMoS-COAST, based on conservation of sediment in the alongshore direction, is given by

$$\underbrace{\frac{\partial Y}{\partial t}}_{\text{shoreline change}} = \underbrace{-\frac{1}{d}\frac{\partial Q}{\partial X}}_{\text{longshore transport}} + \underbrace{CE^{1/2}\Delta E}_{\text{cross-shore transport}} - \underbrace{\frac{c}{\tan\beta}\frac{\partial S}{\partial t}}_{\substack{\text{shoreline migration}\\ \text{due to sea-level rise}}} + \underbrace{v_{lt}}_{\substack{\text{long-term shoreline trend;}\\ \text{unresolved processes}}}$$

The original CoSMoS-COAST model applied an extend Kalman filter data-assimilation technique following[9]. The current version of CoSMoS-COAST applies an ensemble Kalman filter data assimilation technique (with $N = 100$ ensembles) following [23].

COCOONED(HM10, R4)

COCOONED model presented by Antolínez et al.[24] is a transect-based process-driven coastal change model loaded by waves and varying water levels that solves: nearshore wave propagation with a hybrid scheme based on SWAN[16]; cross-shore transport and equilibrium shoreline change using a modified version of Miller and Dean model[12]; longshore transport and shoreline change with a one line approach similar to[25]; foredune erosion using the model proposed by [26] based on [13] and the inclusion of profile adjustment by sediment supply. In Tairua, we do not compute dune erosion and we solve the equation:

$$\frac{\partial \mathrm{Y}}{\partial \mathrm{t}} = \frac{-1}{\mathrm{d}}\frac{\partial \mathrm{Q_L}}{\partial \mathrm{x}} + \mathrm{K_C}\left(\mathrm{Y_{eq}} - \mathrm{Y}\right) + \frac{-1}{\mathrm{d}}\left(\mathrm{q_x} + \mathrm{q_y}\right)$$

where Y represents the shoreline position, t is time, $\frac{\mathrm{q_x}}{\mathrm{d}}$ and $\frac{\mathrm{q_y}}{\mathrm{d}}$ are the alongshore and cross-shore sediment sources per unit shoreline and unit time. $\mathrm{Q_L}$ is the alongshore transport rate,

$$\mathrm{Q_L} = \mathrm{Q_{0sin}}\left(2(\theta_\mathrm{b} - \alpha_{\mathrm{shoreline}})\right)$$

where $\mathrm{Q_0}$ is computed applying the CERC[27] sediment transport formula. The cross-shore component of the model is based upon the general observation that the shoreline tends to approach an equilibrium position, $\mathrm{Y_{eq}}$.

The model is set-up with transects every 100 m and is run with a time step of 6h. There are no sediment sources at the boundaries ($\mathrm{q_x} = 0,\ \mathrm{q_y} = 0$), and we use the waves characteristics are provided. The model is loaded with waves and water levels. Three calibration parameters are set up to reproduce the observational period provided in terms of RMSE, bias and correlation coefficient (alongshore transport rate, $\mathrm{K_L}$; erosional cross-shore transport rate, $\mathrm{K_{C,e}}$; accretional cross-shore transport rate, $\mathrm{K_{C,a}}$) at every transect.

Rotation R5, R6

These models use the alongshore wave energy flux to predict the orientation of the shoreline, where the orientation is the slope of the regression line fit to the shoreline for each time step. The models are essentially the same as the one-line diffusion models

(reviewed in [28]), which have been applied to beach rotation in [29],with the addition that the beach rotates around its alongshore equilibrium position (which is the shoreline that would evolve if the alongshore wave energy flux was zero.) As the alongshore wave energy flux deviates from zero, there is a time-delay over which the beach responds to the change. [29] use the beach volume and sediment characteristics to inform estimates of the delay. We use the version in [30] where the rotation rates are coefficients C that are set by fitting the observations to the model.

$$\frac{\partial \alpha}{\partial t} = CE(\sin\theta\cos\theta - \alpha\cos 2\theta)$$

Where E is wave energy, θ is wave approach angle and α is the shoreline orientation. In R5, constant coefficients were used, with a different one for negative and positive wave energy flux. In R6, coefficients were linearly dependant on the wave energy flux, with the slope and intercept fitted by matching model output to observations (this should allow better fitting of large events). The optimal values of coefficients were obtained by RSME using boot-strapping for optimisation. The mean was removed from the observations and mean beach orientation (54.2 degrees) was removed from θ.

Machine Learning (ML) Models

k Nearest Neighbor (kNN)

The k-nearest neighbor (kNN) model used the past as a 'catalog' to predict the future system dynamics. Here, the 'catalog' is a set of 5 dimensional vectors, where each vector represents a day. The 5 dimensions are the daily significant wave height, Hs, daily shoreline position, two past shoreline positions (from previous days) and the future shoreline positions (which can be used to compute daily shoreline change). A prediction involves searching the catalog to find the 'k' nearest vectors in the 4 dimensional Euclidean space. The 'k' past instances are used to compute a weighted average (using inverse distance) of the daily shoreline change. The shoreline position can then be updated, and the whole process repeats for the following day's prediction. To optimize parameters, the known historic data were split into two parts — the first was used as a catalog, and the second was treated as validation data to find optimized parameter values. After the optimization, the entire catalog of known data were used for the 'Shorecast'.

Autoregressive neural network (ANN) with exogenous inputs

The input dataset consisted of a multidimensional dataset combining the modes obtained through out the Complete Ensemble Empirical Mode Decomposition method [5], with daily wave conditions (H_s, mean wave period T_{02}, wave direction θ) and the shoreline position of the previous day Y_{i-1}. The two algorithms proposed used a different number of nodes (7 and 2) to connect the variables in a single hidden layer (a sigmoid function was used to connect different layers). The calibration was performed by minimizing the RMSE and differed between the two algorithms proposed. In the second algorithm, white noise (+/- 5 m) was added to the shoreline signal to minimize overfitting.

Long-Short Term Memory (LSTM)

LSTMs are recurrent neural network models capable of learning long-term dependencies in sequence modelling problems [31]. A stateless model was created using six input features ($H_s, T_{01}, T_{02}, T_p,\ \theta$ and the previous shoreline position Y_{i-1}). Features were normalised and training samples were generated adding sequence subsampling with a window size of 10 time-steps The LSTM model was implemented using Python with Keras and Tensorflow. The model architecture consisted of a single LSTM unit fully connected to a single output neuron using a linear activation. Mean Squared Error was used as the loss function and the Adam[32] optimizer was used to tune network weights. Nested cross-validation was performed using 5 partitions with a batch size of 4 and 10 epochs per training cycle. In all cases the random number generator was initialised to 42. Daily forecasts were generated iteratively, using the forecasted value as part of the next input series.

NeuFor

A neural network (multilayer perceptron) was used to predict shoreline change (dY_t) at each timestep (t) and integrated to model shoreline position (Y_t) over time. The neural network takes as input the normalised hydrodynamic forcing (H_s and T_{02}) over a daily shoreline change timestep. The model has an explicit feedback whereby the previous prediction of overall shoreline position (Y_{t-1}) forms an additional normalised input variable to the network. The hydrodynamic forcing variables are provided to the model at three-hour timesteps making a total of 17 input variables including the previous shoreline position. An exhaustive search using 5-fold cross-validation was to determine the optimum model inputs, architecture and hyperparameters. The model was developed and trained using the Python package PyTorch to optimise mean squared error and the coefficient of determination when predicting shoreline position over time. The final model contained two hidden layers (of 40 and 20 neurons), with rectified linear unit activation functions and a dropout of 40% on each of the hidden layers to prevent overfitting.

Random Forest (RF)

Random Forests (RFs) are ensembles of decision trees[33]. Training time series of shoreline positions were linearly interpolated and smoothed with a three day moving average. The input variables used are H_s, T_p and the disequilibrium in dimensionless fall velocity (Ω_{dis}) defined as the difference between the instantaneous and weighted average of the antecedent dimensionless fall velocity[1] and using a 'memory decay' (ϕ) of 30 days. H_s, T_p and Ω were defined at 12 and 24 hours prior to time *t*, resulting in a total of seven input variables. We used the Python toolkit SciKit-Learn[34]. The final model was an ensemble of 50 decision trees where each tree was constructed using a bootstrap sample of the training dataset. The samples were the same size as the training data but drawn with replacement. Each tree was also constrained to have a maximum depth of eight and a minimum of 20 data samples was required to form a split in the tree. The hyperparameter values were selected following an exhaustive 10-fold cross-validation grid search with the RF performance measured as the RMSE from 10-fold cross-validation.

Bayesian Networks (BN)

Following [35] we forecast the alongshore averaged shoreline position using as input variables the position at an earlier time (Y_{i-1}), H_s, T_{02}, and wave direction. Model time-

step was 5 days, and faster variations were removed by time-averaging over this interval. Time-averaged inputs were used to train the BN[36]. Input variables were discretized into 3 bins, each having equal prior probability based on the training data. The output bin (Y_{i+1}), was discretised into 5 bins. A 10-fold calibration/validation testing approach [37] ensured that forecast skill was maximized. During calibration, the RMSE was minimized but results indicate that the calibration included some overfitting. We cast the forecast into a Bayesian-mean value, x, where $x = \sum p(i)x(i)$ over the bins, $p(i)$ is the predicted forecast probability that the shoreline lies in the *ith* bin and $x(i)$ is the shoreline position at the center of the *ith* bin. The forecast uncertainty can be computed as the variance using the same approach. The results obtained at the model time step (5 days) were linearly interpolated to the sample time step (1 day) for comparison with the other models.

Models performance assessment.

A brief description of the metrics used is provided below. The following symbols are used:

- $o =$ observed shoreline position /rotation angle
- $m =$ modelled shoreline position /rotation angle.

Also,the use of a top bar ($\overline{x}$) indicate the average, σ_x denotes the standard deviation, σ_x^2 is the variance, $|x|$ is the absolute value and N the length of the time series. We used the following metrics:

- *Squared error or coefficient of determination:* Indicates how well model fit the observed data in terms of their correlation. A disadvantage R^2 is that the series may be well correlated but have large residuals[1].

$$R^2 = \left(\frac{1}{N-1}\sum_{i=1}^{N}\frac{(o_i-\bar{o})(m_i-\bar{m})}{\sigma_o\sigma_m}\right)^2.$$

- *Root Mean Square Error:* Summarize the mean differences in units of o and m.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(o_i - m_i)^2}$$

- *Skill:* Compares the error variance with the observed variance by normalizing RMSE. In general, skill can range as: a poor model performance (0-0.2), Fair(0.2-0.4) , Good(0.4-0.7), Excellent(> 0.7).

$$Skill = 1 - \frac{\sum_{i=1}^{N}(o_i - m_i)^2}{\sigma_o^2}$$

- *Mean Absolute Error:* As the RMSE, this metric does not consider the direction of the errors and additionally it is unbounded.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|o_i - m_i|.$$

Index of Agreement d[38]*:* The index of agreement can detect additive and proportional differences in the observed and simulated means and variances. However, it is overly sensitive to extreme values due to the squared differences.

$$d = 1 - \frac{\sum_{i=1}^{N}(o_i - m_i)^2}{\sum_{i=1}^{N}(|m_i - \bar{o}| + |o_i - \bar{o}|)^2}$$

where $0 \leq d \leq 1$, with values close to zero for poor performance and values close to 1 indicating excellent performance.

- *Mielke's modification* $\boldsymbol{\lambda}$ [39]*:* It is a symmetric index, dimensionless and bounded as the index of agreement.

$$\lambda = 1 - \frac{N^{-1}\sum_{i=1}^{N}(o_i - m_i)^2}{\sigma_o^2 + \sigma_m^2 + (\bar{o} - \bar{m})^2}$$

Table S3 displays the above metrics for each model. Each model is represented by a colour shown in the legend. Models that obtained a good performance based on each metric are in the upper part of the table.

Table 1. Summary of different metrics used to assess models performance. Upper panels display values for the averaged alongshore shoreline position and bottom panels metrics for the rotation predictions during the calibration period (left) and Shorecast (right).

Calibration period (Shoreline position)					
r2	RMSE	Skill	MAE	d	λ
0.93	1.39	0.93	1.01	0.98	0.97
0.86	2.18	0.84	1.67	0.95	0.90
0.78	2.56	0.78	1.89	0.94	0.88
0.74	2.76	0.74	2.07	0.92	0.86
0.74	3.02	0.69	2.40	0.92	0.85
0.72	3.07	0.68	2.44	0.87	0.76
0.54	3.79	0.51	2.92	0.82	0.68
0.50	3.83	0.50	2.98	0.82	0.67
0.50	3.85	0.49	2.99	0.80	0.65
0.49	3.88	0.49	3.01	0.80	0.65
0.46	4.04	0.44	3.08	0.79	0.62
0.44	4.05	0.44	3.16	0.78	0.62
0.40	4.25	0.38	3.33	0.78	0.62
0.40	4.27	0.37	3.35	0.76	0.58
0.38	4.32	0.36	3.43	0.75	0.57
0.37	4.38	0.34	3.49	0.71	0.52
0.34	4.39	0.34	3.52	0.70	0.50
0.33	4.45	0.32	3.53	0.68	0.48
0.12	5.07	0.12	4.07	0.47	0.23
	5.39	0.00	4.29	0.08	0.01
(-)	(-)	(-)	(-)	(-)	(-)

Shorecast (Shoreline position)					
r2	RMSE	Skill	MAE	d	λ
0.39	4.50	0.28	3.42	0.73	0.51
0.39	4.51	0.28	3.45	0.69	0.47
0.35	4.56	0.27	3.52	0.69	0.46
0.34	4.70	0.22	3.57	0.68	0.46
0.33	4.73	0.21	3.60	0.67	0.45
0.32	4.77	0.19	3.62	0.67	0.44
0.31	4.91	0.15	3.88	0.66	0.43
0.28	5.04	0.10	3.91	0.64	0.40
0.26	5.07	0.09	4.01	0.64	0.38
0.21	5.12	0.07	4.02	0.63	0.37
0.21	5.16	0.06	4.05	0.62	0.35
0.21	5.26	0.02	4.07	0.58	0.31
0.19	5.33	0.00	4.15	0.57	0.30
0.18	5.38	-0.02	4.16	0.56	0.30
0.17	5.43	-0.04	4.18	0.56	0.30
0.14	5.46	-0.05	4.27	0.56	0.28
0.14	5.57	-0.10	4.48	0.55	0.28
0.11	5.72	-0.16	4.53	0.54	0.26
0.10	5.81	-0.20	4.56	0.54	0.23
0.09	6.32	-0.41	5.11	0.51	0.23
0.09	7.48	-0.98	5.91	0.24	-0.01

Shoreline Position
HM1
HM2
HM3
HM4
HM5
HM6
HM7
HM8
HM9
HM10
kNN
ANN-EI1
ANN-EI2
NeuFor
LSTM
RF
BNN
Trend
HM
ML
Ensemble

Models color legend

Callibration period (Shoreline rotation)					
r2	RMSE	Skill	MAE	d	λ
0.77	0.43	0.76	0.33	0.92	0.85
0.58	0.64	0.46	0.5	0.81	0.66
0.47	0.66	0.42	0.53	0.8	0.65
0.45	0.74	0.28	0.58	0.77	0.59
0.41	0.77	0.23	0.64	0.75	0.56
0.37	0.92	-0.11	0.71	0.6	0.35
0.24	1.12	-0.65	0.94	0.56	0.32

Shorecast (Shoreline rotation)					
r2	RMSE	Skill	MAE	d	λ
0.61	0.54	0.61	0.44	0.87	0.76
0.61	0.54	0.61	0.45	0.86	0.74
0.61	0.57	0.57	0.47	0.86	0.74
0.57	0.6	0.52	0.48	0.84	0.71
0.54	0.76	0.23	0.57	0.78	0.62
0.54	0.78	0.2	0.64	0.77	0.58
0.41	0.79	0.18	0.66	0.64	0.43

Shoreline Rotation
R1-HM2
R2-HM8
R3-HM9
R4-HM10
R5
R6
Ensemble

References

1. Davidson, M. A., Splinter, K. D. & Turner, I. L. A simple equilibrium model for predicting shoreline change. *Coast. Eng.* **73,** 191–202 (2013).
2. Splinter, K. D. *et al.* A generalized equilibrium model for predicting daily to interannual shoreline response. *J. Geophys. Res. F Earth Surf.* **119,** 1–23 (2014).
3. Kamphuis, J. W. *Introduction to coastal engineering and management*. (World Scientific, 2010).
4. Yates, M. L., Guza, R. T. & O'Reilly, W. C. Equilibrium shoreline response: Observations and modeling. *J. Geophys. Res. Ocean.* **114,** 1–16 (2009).
5. Huang, N. E. *et al.* The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **454,** 903–995 (1998).
6. Torres, M., Colominas, M. A., Schlotthauer, G. & Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. in *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)* 4144–4147 (2011).
7. Wan, E. A. & van der Merwe, R. The Unscented Kalman Filter. in *Kalman Filtering and Neural Networks* 221–280 (2001).
8. Julier, S. J. & Uhlmann, J. K. Unscented filtering and nonlinear estimation. in

Proceedings of the IEEE 92.3 401–422 (2004).
9. Long, J. W. & Plant, N. G. Extended Kalman Filter framework for forecasting shoreline evolution. *Geophys. Res. Lett.* **39,** 1–6 (2012).
10. Wright, L. D. & Short, A. D. Morphodynamic variability of surf zones and beaches: a synthesis. *Mar. Geol.* **56,** 93–118 (1984).
11. Wright, L. D., Short, A. D. & Green, M. O. Short-term changes in the morphodynamic states of beaches and surf zones: An empirical predictive model. *Mar. Geol.* **62,** 339–364 (1985).
12. Miller, J. K. & Dean, R. G. A simple new shoreline change model. *Coast. Eng.* **51,** 531–556 (2004).
13. Kriebel, B. D. L. & Dean, R. G. Convolution method for time-dependent beach-profile response. *J. Waterw. Port, Coastal, Ocean Eng.* **119,** 204–226 (1993).
14. Castelle, B. *et al.* Equilibrium shoreline modelling of a high-energy meso-macrotidal multiple-barred beach. *Mar. Geol.* **347,** 85–94 (2014).
15. Robinet, A., Idier, D., Castelle, B. & Marieu, V. A reduced-complexity shoreline change model combining longshore and cross-shore processes: The LX-Shore model. *Environ. Model. Softw.* **109,** 1–16 (2018).
16. Booij, N. R. R. C., Ris, R. C. & Holthuijsen, L. H. A third-generation wave model for coastal regions: 1. Model description and validation. *J. Geophys. Res. Ocean.* **104,** 7649–7666 (1999).
17. Kamphuis, J. W. Alongshore Sediment Transport Rate. *J. Waterw. Port, Coastal, Ocean Eng.* **117,** 624–640 (1991).
18. Vitousek, S., Barnard, P. L., Limber, P., Erikson, L. & Cole, B. A model integrating longshore and cross-shore processes for predicting long-term shoreline response to climate change. *J. Geophys. Res. Earth Surf.* 1–25 (2017). doi:10.1002/2016JF004065
19. Pelnard-Considère, R. Essai de théorie de l'évolution des formes de rivage en plages de sable et de galets. *Les Energies de la Mer: Compte Rendu Des Quatriemes Journees de L'hydraulique* 13 (1956).
20. Bruun. Sea-Level Rise as a Cause of Shore Erosion. *J. Waterw. Harb. Div.* **88,** 117–132 (1962).
21. Davidson-Arnott, R. G. D. Conceptual Model of the Effects of Sea Level Rise on Sandy Coasts. *J. Coast. Res.* **216,** 1166–1172 (2005).
22. Anderson, T. R., Fletcher, C. H. ., Barbee, M. M., Frazer, L. N. & Romine, B. M. Doubling of coastal erosion under rising sea level by mid-century in Hawaii. *Nat. Hazards* **78,** 75–103 (2015).
23. Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res. Ocean.* **99,** 10143–10162 (1994).
24. Antolínez, J. A. A., Méndez, F. J., Anderson, D., Ruggiero, P. & Kaminsky, G. M. Predicting climate driven coastlines with a simple and efficient multi-scale model. *J. Geophys. Res. Earth Surf.* 2018JF004790 (2019). doi:10.1029/2018JF004790
25. Vitousek, S. & Barnard, P. L. A nonlinear, implicit one-line model to predict long-term shoreline change. in *Proceedings of the Coastal Sediments 2015* (2015).
26. Mull, J. & Ruggiero, P. Estimating Storm-Induced Dune Erosion and Overtopping along U.S. West Coast Beaches. *J. Coast. Res.* **298,** 1173–1187 (2014).

27. USAGE. *Shore protection manual*. (1984). doi:doi.org/ 10.5962/bhl.title.47829
28. Ruggiero, P., Buijsman, M., Kaminsky, G. M. & Gelfenbaum, G. Modeling the effects of wave climate and sediment supply variability on large-scale shoreline change. *Mar. Geol.* **273,** 127–140 (2010).
29. Turki, I., Medina, R., Coco, G. & Gonzalez, M. An equilibrium model to predict shoreline rotation of pocket beaches. *Mar. Geol.* **346,** 220–232 (2013).
30. Blossier, B., Bryan, K. R., Daly, C. J. & Winter, C. Shore and bar cross-shore migration, rotation, and breathing processes at an embayed beach. *J. Geophys. Res. Earth Surf.* **122,** 1745–1770 (2017).
31. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9,** 1735–1780 (1997).
32. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. in *arXiv preprint arXiv:1412.6980* 1–15 (2014).
33. Breiman, L. Random forests. *Mach. Learn.* **45,** 5–32 (2001).
34. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12,** 2825–2830 (2011).
35. Plant, N. G. & Stockdon, H. F. Probabilistic prediction of barrier-island response to hurricanes. *J. Geophys. Res. Earth Surf.* **117,** 1–17 (2012).
36. Corp, N. S. Netica 5.05.
37. Fienen, M. N. & Plant, N. G. A cross-validation package driving Netica with python. *Environ. Model. Softw.* **63,** 14–23 (2015).
38. Willmott, C. J. & Willmott, C. J. Some Comments on the Evaluation of Model Performance. *Bull. Am. Meteorol. Soc.* **63,** 1309–1313 (1982).
39. Duveiller, G., Fasbender, D. & Meroni, M. Revisiting the concept of a symmetric index of agreement for continuous datasets. *Sci. Rep.* **6,** 1–14 (2016).