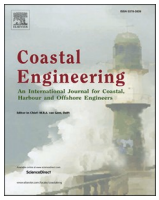
Coastal Engineering 199 (2025) 104738

Contents lists available at ScienceDirect

Coastal Engineering

journal homepage: www.elsevier.com/locate/coastaleng



Evaluating five shoreline change models against 40 years of field survey data at an embayed sandy beach

Oxana Repina [a,*], Rafael C. Carvalho [a,b], Giovanni Coco [c], José A.Á. Antolínez [d], Iñaki de Santiago [e], Mitchell D. Harley [f], Camilo Jaramillo [g], Kristen D. Splinter [f], Sean Vitousek [h], Colin D. Woodroffe [a]

[a] School of Science, University of Wollongong, Wollongong, NSW, Australia
[b] School of Environmental and Life Sciences, College of Engineering, Science and Environment, University of Newcastle, Callaghan, NSW, Australia
[c] School of Environment, University of Auckland, Auckland, New Zealand
[d] Department of Hydraulic Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands
[e] AZTI Marine Research, Basque Research and Technology Alliance, Pasaia, Spain
[f] Water Research Laboratory, School of Civil and Environmental Engineering, UNSW Sydney, Manly Vale, NSW, Australia
[g] Environmental Hydraulics Institute "IHCantabria", Universidad de Cantabria, Santander, Spain
[h] U.S. Geological Survey, Santa Cruz, CA, USA

ARTICLE INFO

Keywords:
Reduced-complexity models
Equilibrium models
Shoreline evolution
Cross-shore transport
Longshore transport

ABSTRACT

Robust and reliable models are needed to understand how coastlines will evolve over the coming decades, driven by both natural variability and climate change. This study evaluated how accurately five popular 'reduced-complexity' models replicate multi-decadal shoreline change at Narrabeen-Collaroy Beach, a sandy embayment in Sydney, Australia. Measured shoreline positions derived from approximately monthly field surveys were used for 20-year calibration and 20-year validation periods. The models performed similarly on average but with large variability between transects. The set-up of several models was modified to compensate for their sensitivity to imperfect input wave data, and further site-specific improvements were identified. Capturing interannual to decadal-scale variability in cross-shore and longshore dynamics at this site was challenging for all five models. Models appeared to aggregate key processes at this timescale into parameter values rather than representing them directly. This suggests time-varying parameters or changes to model structure may be necessary for decadal-scale simulations.

1. Introduction

Sandy beaches represent a third of the global coastline and are highly dynamic environments (Luijendijk et al., 2018), with shorelines capable of shifting tens or hundreds of metres over timescales of days to decades in response to storm events, climatic oscillations, and unbalanced sediment budgets (Harley et al., 2017; Reeve et al., 2019; Bishop-Taylor et al., 2021; Vos et al., 2023). Climate change over this century will add to this dynamic behaviour through sea-level rise as well as altered wave climates and fluvial sediment supply (Ranasinghe, 2016). Understanding the drivers of shoreline change over past decades and accurately modelling future shoreline change over the coming decades is needed for governments, developers, insurers and home-owners to manage risk to coastal assets and ecosystems, and adapt to climate change impacts (Ranasinghe, 2020; Toimil et al., 2020; Vousdoukas et al., 2020).

Over decadal timescales, shoreline change is driven by a combination of short-term and long-term processes, which may involve both cross-shore and longshore sediment exchange. At these scales, the role of gradients in longshore sediment transport becomes increasingly significant (Komar, 1971; Cowell et al., 1995), as does gradual sediment exchange between the upper and lower shoreface (Anthony and Aagaard, 2020; Hamon-Kerivel et al., 2020), exchange between beaches and nearby estuaries and inlets (Stive and Wang, 2003; Ranasinghe et al., 2013; Bamunawala et al., 2020), and the effects of mean sea-level change (Bruun, 1962; Cowell and Thom, 1994; Cowell and Kinsela, 2018). Short-term variability in wave conditions and water levels likewise contributes to shoreline variability over these timescales through storm erosion/recovery cycles (Coco et al., 2014; Brooks et al., 2017;

* Corresponding author.
E-mail address: orepina@uow.edu.au (O. Repina).

https://doi.org/10.1016/j.coastaleng.2025.104738
Received 5 December 2024; Received in revised form 14 February 2025; Accepted 25 February 2025
Available online 26 February 2025

Mortlock et al., 2017) and/or beach rotation (Ranasinghe et al., 2004; Harley et al., 2011a), in some cases driving most or all shoreline variability on beaches that are otherwise stable (Reeve et al., 2019; McLean et al., 2023). In addition, while event-scale processes have previously been discounted as noise about a mean trend (Cowell et al., 2003), recent work suggests that major storms may play a role in modulating the longer-term shoreline behaviour by facilitating sediment exchange between the upper shoreface and sources/sinks such as inlets (Sharples et al., 2020) or the lower shoreface (Harley et al., 2022). Seasonal or interannual wave climate variability can also influence longer-term sediment budgets by facilitating headland bypassing (Goodwin et al., 2013; Silva et al., 2021; Wishaw et al., 2021). Not all sites are sensitive to all processes and the most influential processes are site-dependent; however, shoreline change models applied to decadal-scale simulations will typically need to resolve multiple processes operating at multiple timescales, either directly or indirectly (Ranasinghe, 2020; Toimil et al., 2020; Hunt et al., 2023).

Models of shoreline change can be categorised on a spectrum of complexity according to whether they are physics/process-based (i.e., deductive) or data-driven (i.e., inductive) (Hunt et al., 2023). Models between the two ends of the spectrum balance the amount of process knowledge and input data required by aggregating individual processes into an overall behaviour or morphodynamic response. Labelled as reduced-complexity or semi-empirical models by Hunt et al. (2023) and hybrid models by Montaño et al. (2020), they often simulate a single process, such as cross-shore shoreline change in response to short-term variability in waves and/or water levels (Miller and Dean, 2004; Yates et al., 2009; Davidson et al., 2013), longshore sediment transport (Hanson, 1989; Ashton et al., 2001; Vitousek and Barnard, 2015), beach profile response to long-term sea-level rise (Bruun, 1962; Rosati et al., 2013; Atkinson et al., 2018), or planform beach rotation (Turki et al., 2013; Jaramillo et al., 2021a). Recent work has explored how such models can be coupled together in 'hybrid frameworks' that capture multiple drivers of shoreline change in a computationally efficient way (Vitousek et al., 2017b; Robinet et al., 2018; Antolínez et al., 2019; Hallin et al., 2019; Palalane and Larson, 2020; Tran and Barthélemy, 2020; Alvarez-Cuesta et al., 2021; de Santiago et al., 2021; Jaramillo et al., 2021b). At least in theory, these types of frameworks appear to be most viable for decadal-scale simulations at present (Reeve et al., 2016; van Maanen et al., 2016; Vitousek et al., 2017a; Ranasinghe, 2020; Toimil et al., 2020; Hunt et al., 2023).

However, justifying the use of a specific model for a given application requires that model performance is quantified and its relative strengths and limitations are understood (Jakeman et al., 2006; Bennett et al., 2013). While semi-empirical and reduced-complexity models may be simple, stable, and computationally efficient, this typically comes with the trade-off of reduced generality or transferability to different sites (Hunt et al., 2023), and testing performance at a range of morphodynamically distinct sites is necessary (e.g., Splinter et al., 2014). Where multiple models are coupled together, validation of the entire framework is necessary, as validation of the components does not necessarily prove the model framework itself is robust (Jakeman et al., 2006). A critical part of this is evaluating models' ability to replicate observed shoreline change, and in particular, their ability to do so on a separate period of data not used for model calibration (a blind validation) (Vitousek et al., 2017a; Montaño et al., 2020). Evaluating performance over the calibration period alone is not sufficient, as some of the apparently best-performing models may overfit to noise in the calibration data and have no forecasting skill (Aber, 1997; Jakeman et al., 2006). While performance has been rigorously evaluated this way for models over event-based or interannual timescales (e.g., Splinter et al., 2013; Montaño et al., 2020; Simmons and Splinter, 2022), few studies have tested whether models are capable of accurately replicating high frequency (e.g., monthly) measured shoreline change over a multi-decadal validation period. Recently, Ibaceta et al. (2022) tested a popular shoreline change model over a 14-year validation period at the Gold Coast, Australia. They showed that it was susceptible to large-magnitude errors due to long-term non-stationarity in the wave climate relative to the calibration period, highlighting the importance of these types of evaluations.

This study uses 40 years of beach profile survey data from Narrabeen-Collaroy Beach in Sydney, Australia (hereafter Narrabeen) to evaluate five shoreline change models, four of which are hybrid frameworks that capture both cross-shore and longshore processes. Shoreline variability at Narrabeen is driven by complex cross-shore and longshore dynamics that vary both spatially and temporally, providing a thorough test for such models. This site is also one of few globally where in-situ techniques have been used to consistently monitor beach and shoreline behaviour over multiple decades (Turner et al., 2016). This dataset has been previously used to evaluate these types of models but only over shorter, interannual scales using 5–15 year simulation periods, including only up to five years of data for model validation (Davidson et al., 2013; Splinter et al., 2013, 2014; Robinet et al., 2020; Tran and Barthélemy, 2020; Jaramillo et al., 2021a; Schepper et al., 2021). Recently, Jaramillo et al. (2021b) applied the hybrid framework model IH-MOOSE over a 35-year simulation period at Narrabeen but did not consider a separate validation period. In the present study, models are calibrated over a 20-year period and then evaluated over a 20-year validation period. Four of the five models initially simulated trends in shoreline positions that substantially diverged from observations. These initial results are briefly presented before modifications to these models are applied and all five models are re-evaluated together. This study highlights the challenges of accurately reproducing shoreline change over multiple decades, particularly at an embayed beach with complex morphodynamics, and recommends avenues for further improving the performance of reduced-complexity and semi-empirical shoreline change models.

2. Method

2.1. Model descriptions

Five shoreline change models were tested: ShoreFor (Davidson et al., 2013; Splinter et al., 2014), CoSMoS-COAST (Vitousek et al., 2017b), COCOONED (Antolínez et al., 2019), ShorelineEvol (de Santiago et al., 2021) and IH-MOOSE (Jaramillo et al., 2021b). The latter four models are hybrid frameworks that each couple: (1) a sub-model of cross-shore shoreline change driven by short-term variability in waves and/or water levels, based on the equilibrium principles of Wright and Short (1984), and (2) a sub-model of shoreline change driven by longshore processes (longshore transport or beach rotation). The first of the five models, ShoreFor, is a semi-empirical model of cross-shore shoreline change only. Initially, the hybrid framework of Tran and Barthélemy (2020) was included instead, which couples ShoreFor with an analytical longshore sediment transport sub-model. However, stable performance was not able to be achieved with the longshore sub-model over the validation period without the use of continuously time-varying parameters (these are discussed in 4.4 but outside the scope of this study to implement) and the sub-model was therefore excluded from further analysis, replacing the hybrid framework with ShoreFor as a standalone model.

Each of the five models evaluated here simulate shoreline positions only and assume a time-invariant equilibrium profile shape. The models are forced by bulk wave parameters (significant wave height, peak wave period, wave direction) and in some cases water levels, run at hourly to daily timesteps, and are computationally efficient for this timescale (with run-times between a few minutes and an hour for a 20-year calibration period at Narrabeen). In addition to the cross-shore and longshore transport sub-models, which are the focus of this study, several of the hybrid frameworks simulate additional drivers of shoreline change such as long-term sea-level rise, hard engineering structures, estuarine behaviour or storm-driven dune erosion. The five models and their sub-models are summarised in Table 1 and described briefly in turn below,

Table 1
Overview of five shoreline change models evaluated in this study.

Model	Cross-shore change	Longshore change	Mean sea-level rise	Other optional modules	First applied at
ShoreFor (Davidson et al., 2013; Splinter et al., 2014)	ShoreFor	–	–	Long-term trend for unresolved processes	Narrabeen-Collaroy Beach, NSW and Gold Coast, QLD, Australia
CoSMoS-COAST (Vitousek et al., 2017b, 2023)	Y09	One-line (CERC)	Bruun Rule	Long-term trend for unresolved processes	Southern Californian coast, USA
COCOONED (Antolínez et al., 2019)	MD04	One-line (CERC)	Bruun Rule	Foredune erosion (KD93)	North Beach, Washington State, USA
ShorelineEvol (de Santiago et al., 2021)	MD04	One-line (CERC)	Bruun Rule	Hard engineering, estuarine response to sea-level rise	Zarautz Beach, Basque coast, Spain
IH-MOOSE (Jaramillo et al., 2021b)	Y09	J21	–	Long-term trend for unresolved processes	Narrabeen-Collaroy Beach, NSW, Australia

Sub-model abbreviations refer to:
J21: Jaramillo et al. (2021a).
KD93: Kriebel and Dean (1993).
MD04: Miller and Dean (2004).
Y09: Yates et al. (2009).

with additional detail on their input data requirements and calibration parameters provided in 2.3. Full descriptions of the models, their underlying equations and their inputs are provided in the studies referenced in the first column of Table 1.

2.1.1. ShoreFor

ShoreFor (Davidson et al., 2013) is a popular semi-empirical model of cross-shore shoreline change, and a standalone model rather than a hybrid framework. The model is based on the equilibrium principles of Wright and Short (1984), with the simulated shoreline evolving at each timestep towards a time-varying equilibrium position at a rate depending on the present degree of disequilibrium and the wave energy available to transport sediment. This idea also underpins the Yates et al. (2009) model (hereafter Y09) and Miller and Dean (2004) model (hereafter MD04), which are used as cross-shore sub-models in the hybrid frameworks in this study. ShoreFor differs from Y09 and MD04 in that it is formulated explicitly rather than implicitly, so the shoreline position at the previous timestep is not used to calculate the position in the present timestep. Instead, only present and antecedent wave conditions are considered. Shoreline change is simulated independently at each modelled transect and sediment is not conserved between transects. The version of Splinter et al. (2014) is used here, where the accretion rate constant is a free parameter while the erosion rate constant is linearly related to the accretion constant and an "erosion ratio" term calculated from the wave forcing conditions (detailed further in 2.4.1).

2.1.2. CoSMoS-COAST

CoSMoS-COAST was developed and first applied to the coastline of southern California by Vitousek et al. (2017b). While not the first model to couple sub-models of cross-shore and longshore processes (e.g., Huxley, 2009), it motivated recent interest in hybrid frameworks of this type. CoSMoS-COAST discretises the coastline into transects. At each transect, the cross-shore response to varying wave conditions is simulated using Y09. Shoreline change due to longshore sediment transport between transects is captured using a one-line model (Pelnard-Considère, 1956), where sediment transport volumes are estimated using the CERC formula (Komar, 1971; US Army Corps of Engineers, 1984) and the gradient in longshore transport is derived numerically using a finite difference scheme (Vitousek and Barnard, 2015). Hereafter, this is referred to as the 'one-line (CERC)' sub-model.

In addition to the cross-shore and longshore sub-models, CoSMoS-COAST simulates the response to sea-level rise at each transect using the Bruun Rule (Bruun, 1962), and a linear trend term aggregates long-term influences such as fluvial or offshore sediment exchange that are not otherwise captured. The one-line (CERC), Bruun Rule and linear trend terms are added together to calculate 'long-term' shoreline change in the model, while the cross-shore shoreline change driven by Y09 is treated as 'short-term' change, and the two components evolve separately in the model before being linearly added together to produce the final output. CoSMoS-COAST also has data assimilation built in for automated calibration, meaning model parameter values (including the linear trend term) vary during the calibration period as the assimilation algorithm adjusts them each time an observation is available, but remain fixed at their final values during the validation period.

The original CoSMoS-COAST model was updated by Vitousek et al. (2021, 2023) to: (1) reformulate the cross-shore sub-model Y09, retaining identical model behaviour but re-arranging the free parameter terms to be expressed in easily-interpretable units such as distance or time; and (2) modify the data assimilation approach, with the extended Kalman Filter algorithm replaced with an ensemble Kalman Filter, additive noise included to control how rapidly the data assimilation converges to calibrated parameter values, and the algorithm allowed to consider the behaviour of all observations in a littoral cell when assimilating parameters at each transect (i.e., although different parameters may still be assimilated at each transect, the assimilation at a given transect is informed by observations at neighbouring transects). This study uses the updated version of Vitousek et al. (2023).

2.1.3. COCOONED

COCOONED (Antolínez et al., 2019) likewise discretises the coastline into transects, and shoreline change driven by longshore transport gradients is simulated using the one-line (CERC) sub-model. Short-term cross-shore variability is captured using MD04. Unlike ShoreFor and Y09, MD04 uses storm surge as a forcing variable in addition to wave conditions, and the version in COCOONED is modified to also include astronomical tide. MD04 and the one-line (CERC) sub-models are coupled such that sediment gains or losses from longshore transport also alter the equilibrium position term in MD04, rather than linearly adding cross-shore and longshore change as separate contributions.

The Bruun Rule is used in COCOONED to simulate shoreline response to sea-level rise, but unlike CoSMoS-COAST, sea-level rise does not directly shift the shoreline but again alters the equilibrium term in MD04, meaning its contribution is simulated more gradually depending on wave conditions (D'Anna et al., 2021b). COCOONED also simulates foredune erosion following the approach of Kriebel and Dean (1993) and Mull and Ruggiero (2014). Profile change is not explicitly modelled; however, the volume of sediment lost from the foredune during erosion events offsets shoreline erosion by altering the equilibrium shoreline position in MD04. Finally, COCOONED is the only model considered here that internally propagates waves from the input nearshore depth to breaking point. This is achieved using Snell's law and linear wave theory, interactively adjusting wave refraction based on the shoreline orientation at each timestep while assuming shore-parallel bathymetric contours. The other models with longshore sub-models utilise waves at

fixed input depths (refer to Table 2) and only consider changes in shoreline orientation when calculating the relative angle between the shoreline and incident waves.

2.1.4. ShorelineEvol

ShorelineEvol (de Santiago et al., 2021) is a similar model to COCOONED and shares the MD04, one-line (CERC) and Bruun sub-models. Like COCOONED, the one-line (CERC) sub-model adjusts the shoreline position directly as well as shifting the equilibrium position term in MD04 (this is a modification of the original version of ShorelineEvol and allows the cross-shore equilibrium position to account for mean-trend shoreline change driven by longshore transport), while the Bruun Rule shifts the MD04 equilibrium term only. ShorelineEvol uses a slightly simpler implementation that does not internally propagate waves to breaking point, and breaking wave conditions must be pre-computed and input to the model. ShorelineEvol does not account for foredune processes; instead, it simulates the additional sediment demand from coastlines adjacent to estuaries created when sea-level rise increases accommodation space in the estuary (i.e., a sediment sink term) (Stive and Wang, 2003; Ranasinghe et al., 2013; Toimil et al., 2017). It also simulates the effect of hard engineering structures such as seawalls and groynes/breakwaters on cross-shore and longshore sediment transport. The estuary response and hard engineering modules were not activated at Narrabeen (elaborated in 2.3.3), but de Santiago et al. (2021) detail their implementation and provide example applications.

2.1.5. IH-MOOSE

IH-MOOSE, developed by Jaramillo et al. (2021b), differs from the other hybrid frameworks above in that it does not use the one-line (CERC) sub-model for longshore sediment transport or discretise the coastline into transects, and is only applicable to embayed beaches with a parabolic planform shape. IH-MOOSE calculates shoreline change across an embayment in four steps: (1) Y09 is used to calculate wave-driven cross-shore shoreline change at a single profile along the most wave-exposed section, termed the 'cross-shore control profile', (2) the parabolic bay shape equation of Hsu and Evans (1989) is used to model the embayment planform, and the timeseries of cross-shore change from the first step is extended to a timeseries of parabolic planforms, (3) the equilibrium model of planform beach rotation of Jaramillo et al. (2021a), hereafter J21, is used to simulate the mean orientation of the embayment based on offshore wave power and direction, and (4) the timeseries of orientations from the third step is used to rotate the timeseries of parabolic planforms from the second step to produce the final output. In the first step, a modification of Y09 proposed by Jaramillo et al. (2020) may be used at sites with mean-trend change driven by net gains or losses of sediment, where a linear trend term estimated by the user is incorporated into the modelled equilibrium shoreline position.

Table 2
Forcing data used by each model.

Model	Wave conditions		Water level
	Input depth	Required at	
ShoreFor	Nearshore	Model transects	–
CoSMoS-COAST	Nearshore	Model transects	Mean sea level
COCOONED	Nearshore (propagated internally to breaking depth)	Forcing transects	Mean sea level, astronomical tide, storm surge
ShorelineEvol	Breaking	Model transects	Mean sea level, astronomical tide, storm surge
IH-MOOSE	Nearshore	Cross-shore control profile (PF4)	–
	Offshore	Offshore from embayment	–

2.2. Study site description

Narrabeen-Collaroy Beach (Narrabeen) is an embayed sandy beach in Sydney, south-eastern Australia (Fig. 1). The embayment is 3.6 km long and bounded by Narrabeen Headland to the north and Long Reef Point to the south. Narrabeen is considered a stable, closed system with a near-perfectly balanced sediment budget at a decadal scale, as the headlands largely restrict sediment exchange with adjacent beaches and the low volumes of sediment lost to an intermittently open and closed coastal lagoon at the northern end of the beach are periodically dredged and used to replenish the southern end (Woodroffe et al., 2012; Carley et al., 2016). The sediment in the embayment is fairly uniform alongshore, a fine to medium quartz-carbonate sand with a median grain size of 0.3–0.4 mm (Short, 1984; Splinter et al., 2014; Turner et al., 2016). The foredunes reach up to 9 m high along the northern end while urban development has encroached onto much of the foredune at the southern end (Turner et al., 2016). The site is micro-tidal and tides are semi-diurnal, with a spring tidal range of 1.3 m (Turner et al., 2016). Mean sea level is correlated with the El Niño Southern Oscillation (ENSO), but excluding this interannual signal, has approximately followed the global mean over the past several decades (White et al., 2014).

The wave climate is dominated by mid-latitude cyclones which drive southeasterly swells (Short and Trenaman, 1992), with a mean offshore significant wave height of 1.6 m and a peak period of 10 s (Turner et al., 2016). This corresponds to a moderate-to-high energy embayment, but with significant variability about this mean. Storm waves can occur during any time of the year but are more common during the austral winter from mid-latitude cyclones and east coast lows, while summer tends to have lower easterly swell from anticyclonic highs and short northeasterly seas generated by regional sea breezes (Short and Trenaman, 1992). At an interannual scale, ENSO and Southern Annular Mode (SAM) modulate incoming wave energy and direction (Harley et al., 2010).

While the embayment is stable over a multi-decadal scale and shows minimal mean-trend change (e.g., Fig. 2), the energetic, highly variable wave climate interacts with the offshore morphology of the embayment to drive complex shoreline dynamics at shorter timescales. The central, wave-exposed section of Narrabeen can exhibit any of the six morphodynamic beach states (Wright and Short, 1984). The modal state straddles 'rhythmic bar and beach' and 'transverse bar and rip', a configuration that allows for rapid transitions between intermediate states on the order of days to weeks in response to variations in incident wave energy (Wright et al., 1985; Davidson et al., 2013). Wave energy exerts a stronger control over shoreline variability than water levels at these timescales due to a narrow continental shelf, low magnitudes of storm surge, the low tidal range, and typically long-duration storms that span several tidal cycles (Ibaceta and Harley, 2024).

In addition to this cross-shore variability, planform rotation of the embayment has been observed over both seasonal (Harley et al., 2011a) and interannual timescales (Ranasinghe et al., 2004; Short and Trembanis, 2004). This rotation is driven by a combination of cross-shore and longshore processes at Narrabeen, including storm-driven erosion/accretion cycles and sandbar migration (which vary in magnitude across the embayment due to the sheltering effect of the headlands and the resulting gradient in wave energy), as well as classic longshore sediment transport due to shifts in wave direction (Harley et al., 2011a, 2015). The relative dominance of cross-shore and longshore processes also varies significantly over time in response to interannual wave climate variability and large-magnitude nearshore changes following extreme storms (Ibaceta et al., 2023). Such complex morphodynamics, exhibited over a range of temporal scales, provide an opportunity to rigorously test shoreline change models incorporating both cross-shore and longshore processes.

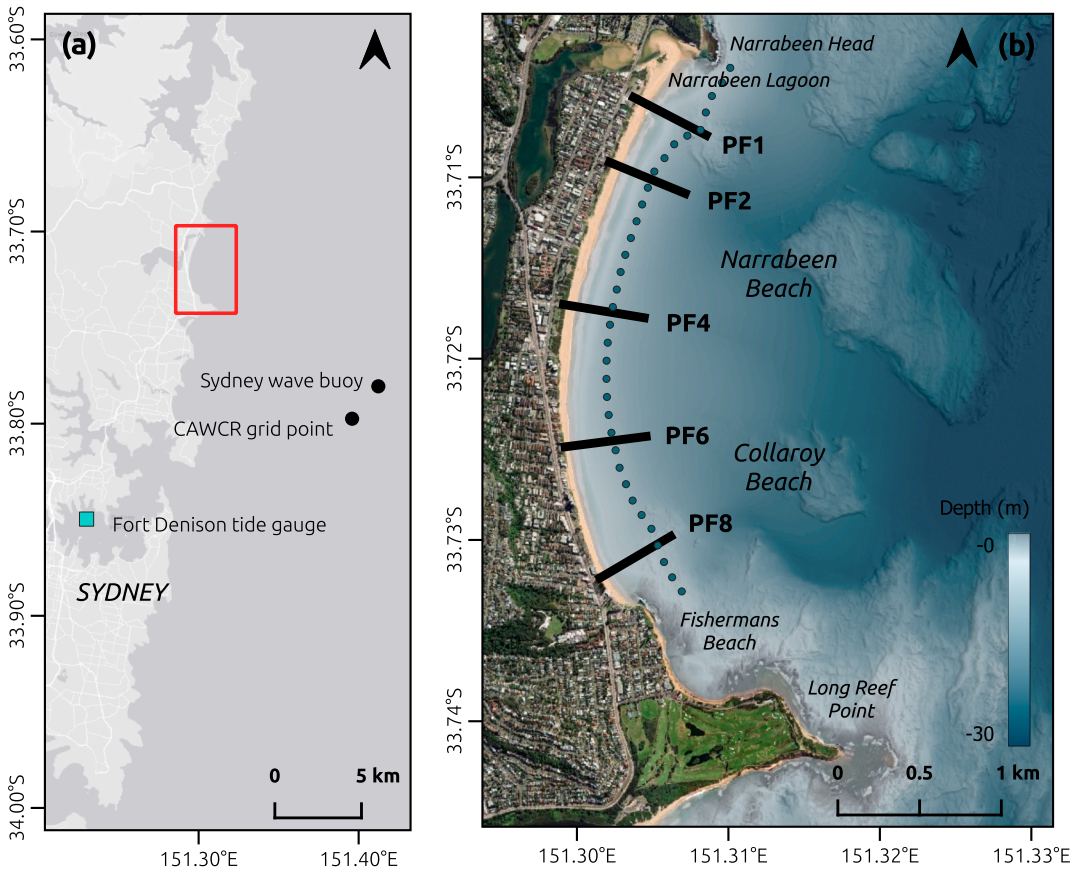


Fig. 1. The case study site, Narrabeen-Collaroy Beach. (a) The location of Narrabeen within Sydney, as well as the locations of data sources for waves (black dots) and water levels (blue square). (b) The embayment in detail, showing the five transects where beach profile surveys have been undertaken (labelled PF1–PF8), the virtual wave buoys where nearshore wave data were sourced (blue dots), and the complex nearshore bathymetry of the embayment. Bathymetric elevation data provided by The New South Wales Department of Climate Change, Energy, the Environment and Water (2019). Basemaps obtained from ESRI. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.3. Input data and model set-up

2.3.1. Measured shoreline positions and spatial discretisation

Beach profile surveys have been undertaken at Narrabeen along five transects (labelled PF1–PF8; Fig. 1b) approximately monthly since April 1976, and weekly to fortnightly in more recent years (Turner et al., 2016). Surveys were carried out using the Emery method (Emery, 1961) at 10 m cross-shore spacing until May 2005, after which RTK-GPS was used at ~0.1 m horizontal spacing (Turner et al., 2016). Harley et al. (2011b) compared concurrent Emery method and RTK-GPS surveys, finding that the earlier surveys showed a standard deviation of cross-shore horizontal errors between mean sea level and the +2 m contour (i.e., potential errors in shoreline positions) of 1.1 m. From each profile, the shoreline was extracted here (e.g., Fig. 2, top panels) as the +0.7 m contour relative to the Australian Height Datum (AHD), a datum which approximately corresponds to the average mean sea level around Australia. The +0.7 m contour represents mean high water springs (MHWS) at Narrabeen and is typically used for shoreline analysis and modelling at this site (e.g., Davidson et al., 2013; Vos et al., 2019a; Simmons and Splinter, 2022; Ibaceta et al., 2023) as it is strongly correlated with subaerial beach volume (Harley et al., 2011b). The initial shoreline position for model simulations was taken from the most recent survey date preceding the simulation start date, January 30, 1979 and February 1, 1979 respectively (the latter coinciding with the earliest available wave data).

ShoreFor simulates shoreline change at individual transects independently, so only the five surveyed transects were modelled. CoSMoS-COAST, COCOONED and ShorelineEvol discretise the planform shoreline into closely-spaced transects linked by longshore sediment transport. For these models, shore-normal transects were defined at approximately 100 m spacing between the five surveyed transects, for a total of 36 modelled transects across the 3.6 km embayment. An extra level of processing was required for ShorelineEvol, which simulates shoreline positions at fixed alongshore 'nodes' that are parallel to the average embayment orientation rather than shore-normal. Nodes were defined using the coordinates of the initial shoreline position along the 36 shore-normal transects. Modelled shorelines were then converted trigonometrically to distances along the five surveyed transects to compare against measured shorelines.

IH-MOOSE outputs continuous planform shorelines, which were intersected with the five surveyed transects after the model was run to compare modelled and measured shoreline positions. The central transect, PF4, was used as the 'cross-shore control profile' as it is the closest transect to the node or pivot point of beach rotation at Narrabeen (Short and Trembanis, 2004; Harley et al., 2011a; Jaramillo et al., 2021b). Following Jaramillo et al. (2021b), the embayment planform was parameterised using a parabolic bay shape equation (Hsu and Evans, 1989) south of PF4, and as a linear segment to the north of this transect.

2.3.2. Forcing data

The forcing variables used by each model are summarised in Table 2. Hourly bulk wave parameters (Fig. 2) were extracted from the Centre for Australian Weather and Climate Research (CAWCR) wave hindcast (Durrant et al., 2014), beginning in February 1979, at the grid point closest to the Sydney wave buoy (Fig. 1a). Refer to Turner et al. (2016) for a validation of this hindcast against Sydney wave buoy

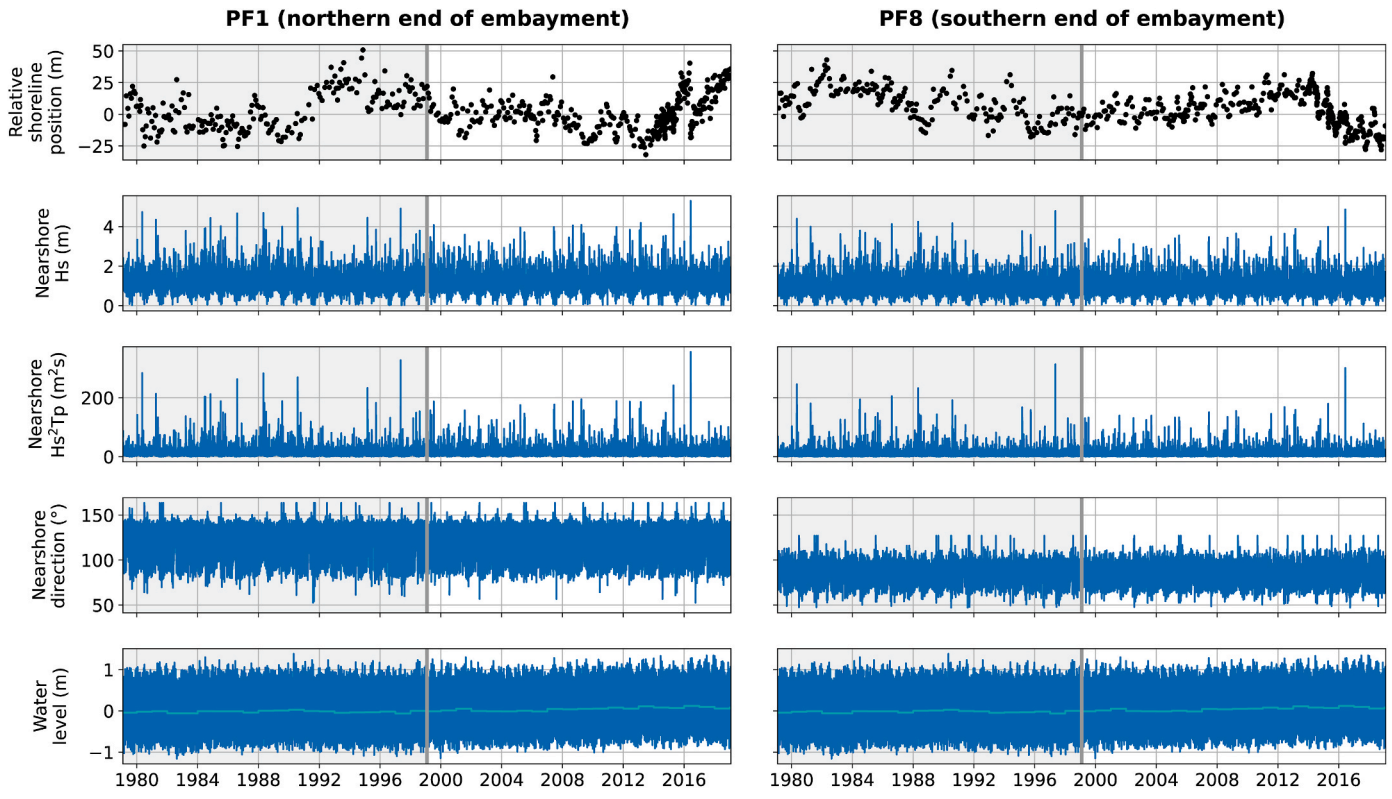


Fig. 2. Measured shoreline positions and forcing data characteristics, shown for two contrasting transects at opposite ends of the embayment. From top to bottom, panels show: measured shoreline positions, significant wave height Hs, a proxy for incident wave power Hs^2Tp (visualised for reference, noting that peak period Tp was input to the models), nearshore wave direction from geographical north, and measured water levels relative to the Australian Height Datum, AHD (with annual mean sea level in light blue). A single water level timeseries was used for the embayment so the lower-most panels are identical for both transects. The 20-year calibration period is shaded and the subsequent 20-year validation period is unshaded. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

measurements. These offshore wave conditions were propagated to nearshore nodes at approximately 10 m depth using parametric transforms developed for the Sydney wave buoy within the NSW State Wide Nearshore Wave Transformation Tool (NWTT) (Baird Australia, 2017) (Fig. 1b). Nearshore wave conditions were then linearly interpolated alongshore to the modelled transect locations. For ShorelineEvol, nearshore waves were propagated to breaking point using Snell's Law and the conservation of energy flux as simultaneous equations under linear wave theory (the same approach is used by COCOONED to propagate waves internally, but COCOONED also accounts for variations in the shoreline orientation during the simulation period while calculating wave refraction). For the J21 rotation sub-model of IH-MOOSE, offshore CAWCR conditions were extracted for the grid point closest to Narrabeen (33° 44′S, 151° 20′E) rather than to the Sydney wave buoy.

Water level data (Fig. 2, bottom panels) were obtained for Narrabeen from the Fort Denison tide gauge located inside Sydney Harbour (Fig. 1a), the nearest gauge with a record extending back to February 1979 to match the start of the wave forcing data. Hourly water levels at this gauge were extracted from the Australian National Collection of Homogenised Observations of Relative Sea Level (ANCHORS) dataset (Hague et al., 2021) and gaps were filled using ANCHORS data from the nearest gauges at Port Kembla or Newcastle. Water levels were then separated into three components using the 'annual means' method (Hague and Taylor, 2021), as follows: (1) annual mean sea level was calculated, (2) astronomical tide was modelled as oscillating around mean sea level using the Matlab package UTide (Codiga, 2011) following Viola et al. (2021), and (3) storm surge was taken as the leftover residual after subtracting mean sea level and astronomical tide from the measured water level.

2.3.3. Model constants and additional model-specific inputs

Hydrodynamic and morphodynamic characteristics required by the models were calculated as follows. The wave height exceeded 12 h per year and the associated period, H_{12} and T_{12}, were calculated from the offshore CAWCR hindcast data over the simulation period and were therefore the same for all modelled transects (5.1 m and 12.9 s respectively). The average annual depth of closure was calculated from the H_{12} and T_{12} values using Hallermeier's (1981) formula, and was likewise the same for all transects (10.5 m). The berm height was set to a constant of 3 m for all transects as the approximate average berm contour at Narrabeen (Harley et al., 2016). The median sediment grain size D_{50} was set at a uniform 0.3 mm across the embayment (Short, 1984; Turner et al., 2016). For the one-line (CERC) sub-models, Dirichlet boundary conditions of zero longshore sediment transport at the edges of the embayment were imposed as Narrabeen is considered an effectively closed sediment compartment (Harley et al., 2011a; Carley et al., 2016).

For CoSMoS-COAST, the nearshore slope tan β was taken as the mean slope between mean sea level and the depth of closure contour at the five surveyed transects using bathymetry data provided by Turner et al. (2016). These slopes were then interpolated to the 36 modelled transects. For COCOONED, dune toe and crest elevations were automatically extracted from each beach profile along the five surveyed transects using the Python package pybeach (Beuzen, 2019; van IJzendoorn et al., 2021). Values that were too low (<2 m above mean sea level) were filtered out as pybeach had difficulty correctly distinguishing between the berm contour and dune contour at the two southernmost transects where much of the foredune has been lost to urban development. The mean dune toe and crest elevations at the five surveyed transects were interpolated to obtain elevations for the 36 modelled transects.

The estuary sub-model of ShorelineEvol was not activated at

Narrabeen because, as described in 2.2, Narrabeen Lagoon is periodically dredged and the sediment used to replenish the beach, effectively negating its role as a sediment sink over the hindcast period. Hard engineering structures were also not implemented in ShorelineEvol. While ad hoc discontinuous armouring is present at the southern end of Narrabeen, this is not typically modelled at this site. The modern vertical seawall was also not modelled as construction began in 2019, after the simulation period used here. The linear trend term of ShoreFor (refer to Equation (4)) was set to zero. The term is intended to represent net sediment gains or losses that are not otherwise captured by the model, but can contribute to unrealistic behaviour over multi-decadal simulations as trends are likely to vary over time, and the cumulative trend over long periods can overwhelm the shoreline behaviour simulated by the rest of the model (D'Anna et al., 2020, 2021a). Narrabeen has also been stable over the long term, with a near-zero average trend over the simulation period (e.g., Fig. 2).

2.3.4. Calibration of model parameters

The simulation period was limited to 40 years by the availability of wave forcing data (from February 1979 onwards) and water level and beach survey data (up to December 2019 at the time it was accessed). The first 20-year period 1 February 1979–1 February 1999 was used for calibration, and the subsequent 20-year period to February 1, 2019 was set aside for validation. ShoreFor was calibrated over ~14.5 years of data (24 July 1984–1 February 1999) as it requires up to 2000 days of wave data prior to simulations starting to calculate the equilibrium condition term (Splinter et al., 2017).

Model performance was evaluated during calibration using Mielke's index λ (Duveiller et al., 2016), given by:

$$\lambda = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2}{\sigma_x^2 + \sigma_y^2 + (\overline{x} - \overline{y})^2 + \frac{\kappa}{n}} \tag{1}$$

where x_i and y_i are the observed and modelled data points, σ_x^2 and σ_y^2 are their population variances, $\overline{x}$ and $\overline{y}$ their means, n is the number of data point pairs, and κ depends on Pearson's correlation coefficient R according to:

$$\kappa = \begin{cases} 0, & \text{if } R \geq 0 \\ 2\left|\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})\right|, & \text{otherwise.} \end{cases} \tag{2}$$

This means λ is equal to zero when $R < 0$, equal to R when there is no bias, and smaller than R in the presence of either additive or multiplicative bias (i.e., offset between the modelled and measured timeseries or consistent differences in the magnitude of oscillations). The relationship between λ and R is given by:

$$\lambda = \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\overline{x} - \overline{y})^2} R \tag{3}$$

Compared to root mean squared error (RMSE), Mielke's index is a better measure of the degree that measured shoreline variability is captured by the model (i.e., whether the standard deviation in both timeseries is similar) and helps to avoid "flat" calibrations that fit the mean trend of the observations but fail to capture oscillations. A point of reference for this index is available in Montaño et al. (2020), who found 10 reduced-complexity shoreline models achieved Mielke's index values of around 0.2–0.7 over a 15-year calibration period using alongshore-averaged shoreline position data at Tairua Beach, New Zealand (their results are compared to those of the present study in 4.1). It is acknowledged that calibration against different metrics may result in different 'optimal' parameters and therefore different model performance; however, models were calibrated using the same metric and later evaluated against multiple metrics to provide a balanced assessment of performance (these are described in 2.5).

Model parameters were calibrated using automated approaches (Table 3). CoSMoS-COAST was calibrated using its built-in data assimilation module, with data assimilation turned on during the calibration period and then off during the validation period. ShoreFor, ShorelineEvol and IH-MOOSE were calibrated using the optimisation algorithm Shuffled Complex Evolution (SCE-UA) (Duan et al., 1992, 1994). SCE-UA is popular in hydrological modelling as it efficiently converges to the global optimum and is relatively robust against imperfect algorithm settings or 'hyperparameters' (Tolson and Shoemaker, 2007; Matott et al., 2013; Arsenault et al., 2014). The long run-time of COCOONED precluded the use of SCE-UA; instead, Dynamically Dimensioned Search (DDS) was used (Tolson and Shoemaker, 2007). DDS was developed to find good (but not necessarily optimal) parameter sets for computationally expensive models within a limited computational budget. Access to parallel computing infrastructure allows multiple parallel instances of DDS to search more of the parameter space and identify better parameter sets over shorter periods of serial compute time (Tolson et al., 2014). COCOONED was calibrated using DDS with 60 serial runs (DDS-60) on 15 parallel compute cores, which required close to a week of serial compute time (total 900 model runs). However, as uniform random sampling rather than log-uniform sampling was used in the implementation of DDS here, the algorithm had difficulty converging to good values of the erosion and accretion rate parameters, whose sampling ranges span several orders of magnitude. This likely contributed to the sub-optimal calibration of the cross-shore sub-model of COCOONED, which is identified in later sections. Additional information on the model free parameters and calibration settings for all five models is provided in the Supplementary Material.

Within each model, all parameters (i.e., for both cross-shore and longshore sub-models) were calibrated simultaneously. Preliminary tests with ShorelineEvol (not shown) indicated that calibrating the sub-models separately produced similar or worse mean performance across transects over the calibration period as simultaneous tuning, and the result depended on the order of calibration (i.e., whether the cross-shore or longshore sub-model was calibrated first), with the behaviour of the sub-model tuned first being over-exaggerated. The exception to this was IH-MOOSE, where the two sub-models were calibrated separately due to the computational expense of the coupling steps and the fact that each sub-model is calibrated against different measured data (cross-shore shoreline change at one central transect for the Y09 sub-model and mean shoreline orientation across the embayment for J21).

2.4. Modifications applied following initial simulations

2.4.1. ShoreFor and re-calculating the erosion ratio parameter

The performance of ShoreFor was found to be acceptable over the calibration period, but the model initially simulated large accretionary trends that strongly diverged from observations over the validation period (presented in 3.1.1 below). This behaviour was found to be resolved by re-calculating ShoreFor's erosion ratio term r independently for the validation period. In ShoreFor, shoreline change $\partial Y/\partial t$ is calculated according to Equation (4), and r controls the rate of shoreline response during erosive forcing conditions (F^-) compared to accretionary conditions (F^+) by scaling the accretion rate parameter c:

$$\frac{\partial Y}{\partial t} = c(F^+ + rF^-) + b \tag{4}$$

where b is the linear trend term, set to zero in this study (refer to 2.3.3). Unlike in the Y09 and MD04 cross-shore sub-models, where no specific relationship is assumed between the accretion and erosion rate constants and both are free parameters, r in ShoreFor is not a free parameter but is calculated internally from the wave forcing according to:

$$r = \left|\frac{\sum_{i=0}^{N} < F_i^+ >}{\sum_{i=0}^{N} < F_i^- >}\right| \tag{5}$$

Table 3
Free parameters, calibration approach and run time of each model.

Model	Calibration method	Programming language (for implementation used here)	Approx. run time over calibration period[a]	Number of free parameters calibrated[b]
ShoreFor	Optimisation (SCE-UA)	R	0.01 s (once φ is set) 10 s–3 min (total, depending on φ) [c]	2 per transect
CoSMoS-COAST	Data assimilation (ensemble Kalman filter)	Matlab	1 h [d]	7 per transect
COCOONED	Optimisation (DDS)	Python	1 h [e]	5 embayment-wide
ShorelineEvol	Optimisation (SCE-UA)	Fortran	3 min [f]	4 embayment-wide
IH-MOOSE	Optimisation (SCE-UA)	Matlab	0.01 s for Y09 or J21 20 min for whole model	8 embayment-wide

[a] On a 4-core 1800 MHz processor for a single run through the 20-year calibration period at an hourly timestep.
[b] ShoreFor and CoSMoS-COAST use alongshore-varying parameters; that is, different parameters were calibrated at each modelled transect. In the remaining models, alongshore-fixed parameters were used (one set for the whole embayment). Model parameters are described in the Supplementary Material.
[c] Run time per transect; each transect is calibrated separately.
[d] With data assimilation turned on.
[e] With internal propagation of waves from nearshore to breaking turned on.
[f] Run-time includes the step of interpolating the output shorelines at fixed nodes to the shore-normal transects.

where | | are absolute value operators, the < > operators remove the linear trend but preserve the mean, and N is the total number of time-steps in the simulation period (for full definitions of the forcing terms F^+ and F^-, refer to Splinter et al., 2014). This formulation is intended to produce no trend in the modelled shoreline position over the simulation period if there is no trend in the balance of accretionary and erosive wave forcing conditions (Splinter et al., 2017). Typically, r is calculated over the calibration period and kept constant for validation or forecasting like other model parameters (Davidson et al., 2013; Montaño et al., 2020), as was done here initially. However, it appears that wave conditions were not consistently balanced over the multi-decadal simulation period used here, and re-calculating r was applied as a reasonable first-pass solution, using one value for the calibration period and another for the validation period. This issue and more sophisticated solutions are discussed in 4.3.2.

2.4.2. One-line (CERC) sub-models and wave direction bias

CoSMoS-COAST, COCOONED and ShorelineEvol were found to simulate a large trend of planform shoreline realignment over both calibration and validation periods during initial simulations, as presented in 3.1.2 below. These three models use the one-line (CERC) sub-model for longshore transport. In this sub-model, the rate of shoreline, $\partial Y/\partial t$, is proportional to the gradient in longshore sediment transport, $\partial Q/\partial x$, according to:

$$\frac{\partial Y}{\partial t} = -\frac{1}{d_B + d_c}\frac{\partial Q}{\partial x} \tag{6}$$

where d_B is the berm height above mean sea level, d_c is the depth of closure below mean sea level and Q is the longshore sediment transport rate, given by:

$$Q = K\frac{\rho}{(\rho_s - \rho)p}\frac{H_b^2 C_{g,b}}{16}\sin[2(\theta_b - \alpha_s)] \tag{7}$$

K is a free parameter controlling the rate of sediment transport, ρ is the sea water density, ρ_s is the sediment density, p is the sediment porosity, H_b is the breaking wave height and $C_{g,b}$ is the group velocity of waves at breaking. CoSMoS-COAST uses a slightly simpler formulation $Q = K_1 H_s^2 \sin[2(\theta_b - \alpha_s)]$, where K_1 is a free parameter that aggregates breaking wave group velocity $C_{g,b}$ with the constants and the nearshore significant wave height is used instead of the wave height at breaking point.

In all three models, erroneous planform realignment was simulated unless the rate parameter K (or K_1 in CoSMoS-COAST) was set to a very small value such that the modelled rate of longshore transport was negligible. Similar modelled realignment at Narrabeen was recently identified by Chataigner et al. (2022) using a standalone one-line longshore transport model, and they showed that it was driven by mean biases in input wave direction data and the high sensitivity of one-line models to such biases.

To evaluate model performance without the influence of potentially poor-quality input data, the realignment trend was corrected for using a 'virtual equilibrium shoreline' approach based on Anderson et al. (2018), Antolínez et al. (2019) and Robinet et al. (2020), as follows. First, each of the three models were run over the calibration period and the simulated shoreline was allowed to realign. The value of K was manually selected to produce a sufficiently high rate of longshore transport such that by the end of the calibration period, the average shoreline position had stabilised, and no further realignment occurred. The shoreline position over the final five years of this simulation was then averaged to represent a virtual 'long-term equilibrium' position relative to the biased wave conditions. Next, each model was re-started from the beginning of the calibration period using the virtual equilibrium shoreline as the initial shoreline. Model performance was then evaluated by comparing modelled shoreline change relative to the virtual equilibrium shoreline, and measured shoreline change relative to the measured initial shoreline (i.e., rather than comparing absolute shoreline positions). This approach is approximately equivalent to removing the mean of the wave direction timeseries (e.g., Tran and Barthélemy, 2020) but was found to produce better results in preliminary tests (not shown). This issue is discussed further in 4.2.1.

2.5. Evaluating model performance

2.5.1. Overall performance

Model performance was evaluated over the validation period using Mielke's index, the same metric used for calibration (Equation (1)). As model performance and relative rankings are known to depend on the choice of metric (Montaño et al., 2020) and different metrics evaluate different aspects of model skill (Bennett et al., 2013; Liemohn et al., 2021), three additional metrics were visualised for the validation period using Taylor diagrams (Taylor, 2001): (1) the correlation between measured and modelled shoreline positions, (2) the standard deviation of modelled shoreline positions compared to measured positions (i.e., how closely models matched the measured shoreline variability), and (3) centred root mean squared error (CRMSE), that is, RMSE with bias removed. As the theoretical design of Taylor diagrams requires bias to be excluded, the complementary metrics of bias (mean error) and uncentred RMSE were also visualised.

2.5.2. Performance over different timescales

To evaluate each model's ability to capture processes operating over different timescales within the multi-decadal period, measured and modelled shoreline positions over the 20-year validation period were separated into a 6-month rolling mean and residual. The 6-month rolling mean was considered to represent seasonal to interannual-scale shoreline variability, while the residual component was attributed to shorter-term, storm-driven erosion/accretion cycles. The models' ability to capture each of these components was evaluated using Mielke's index. Additive bias or mean error was removed from the modelled rolling mean shoreline position so that Mielke's index only measured correlation and multiplicative bias (i.e. whether shoreline variability was over- or under-estimated) for both components. In other words, as additive bias is inherently excluded from the residual component, it was also removed from the 6-month rolling mean component to directly compare model performance scores across the two timescales.

2.5.3. Performance in capturing cross-shore and longshore processes

Finally, to evaluate each model's ability to capture the processes underlying shoreline dynamics at Narrabeen and their spatiotemporal variability, an empirical orthogonal function (EOF) decomposition was performed on both measured and modelled shoreline positions following Harley et al. (2011a), Antolínez et al. (2019) and Robinet et al. (2020). Also known as a principal component analysis (PCA), this technique may be used to separate a temporally and spatially varying signal (in this case, the shoreline position) into a linear series of distinct modes, termed eigenfunctions or EOFs, that each explain an independent component of the variability. The percentage of temporal variability explained by each spatial EOF is given by a corresponding weighting factor or eigenvalue in the linear series. Prior to EOF decomposition, measured and modelled shoreline positions at the five transects were subset to concurrent timesteps, de-meaned, converted to a log-spiral coordinate system to express positions in purely cross-shore and alongshore coordinates (Harley and Turner, 2008), and interpolated to 4 m alongshore (Harley et al., 2011a; Ibaceta et al., 2023). As the technique is very sensitive to the time period used (Ibaceta et al., 2023), the calibration and validation periods were both separately analysed.

Using this approach, Harley et al. (2011a, 2015) and Ibaceta et al. (2023) linked the first two EOFs of measured shoreline change at Narrabeen to: (1) sediment exchange operating consistently either onshore or offshore across the embayment at any one time, represented by a spatial EOF with a uniform sign alongshore (positive or negative), and (2) sediment transport occurring in opposite directions at opposite ends of the embayment, represented by a spatial EOF with a sign reversal near the centre of the embayment. These are termed the "cross-shore coherent" and "longshore coherent" modes respectively. However, EOF decomposition is a purely statistical technique and the modes do not explicitly identify physical cross-shore and longshore processes. At Narrabeen, the cross-shore coherent mode does appear to correspond to cross-shore processes, while the longshore coherent mode appears to be a mixture of mechanisms combining both cross-shore and longshore transport (discussed further in 4.1). Therefore, although this decomposition does not perfectly measure whether the models are correctly partitioning cross-shore and longshore behaviour, it does to some extent indicate their ability to capture multiple spatially and temporally varying processes operating concurrently at Narrabeen.

3. Results

3.1. Initial model performance prior to modifications

3.1.1. ShoreFor (without re-calculating the erosion ratio term)

ShoreFor showed acceptable performance over the calibration period (mean Mielke's index, λ, of 0.38) but initially simulated large accretionary trends that caused the modelled shoreline to diverge from observations by 50–100 m by the end of the validation period (Fig. 3). Similar behaviour was recently observed by Ibaceta et al. (2022), who applied ShoreFor to a 14-year calibration period and used a total of 14 years for validation at a site on the Gold Coast, Australia. They found that ShoreFor simulated large mean-trend erosion when initialised outside of the calibration period, which resulted in an RMSE of 254.3 m over their total 28-year simulation. They attributed this to the model's sensitivity to wave climate non-stationarity between the two simulation periods, and this behaviour is discussed further in 4.3.2 below.

3.1.2. Models with one-line (CERC) sub-models (without a virtual equilibrium shoreline)

The initial output of COCOONED and ShorelineEvol showed a large net clockwise rotation of the embayment planform over both the calibration and validation periods, diverging strongly from measured shoreline positions despite automated calibration using optimisation algorithms (Fig. 4). By the end of the 40-year simulation period, the modelled shorelines showed ~25–75 m of mean-trend accretion at the northern and central transects (PF1–PF4) and ~100–200 m erosion at the southernmost transect (PF8) relative to measured shoreline positions.

CoSMoS-COAST was calibrated using built-in data assimilation. During preliminary tests with assimilation turned off for the full simulation (not shown), CoSMoS-COAST simulated a similar trend of planform shoreline realignment over the calibration period to the two models above. In the initial results (Fig. 4), where data assimilation was turned on for calibration, CoSMoS-COAST replicated measured shoreline changes almost perfectly during the calibration period (mean $\lambda > 0.99$). However, this is a consequence of data assimilation, which adjusts modelled shoreline positions towards measured positions at each timestep for which a measurement is available, and is not necessarily indicative of forecasting performance. During the validation period, where data assimilation was turned off, CoSMoS-COAST simulated similar mean-trend realignment as COCOONED and ShorelineEvol, with accretion at PF1–PF4 and erosion at PF8, although lower in magnitude (Fig. 4).

These three models simulate longshore sediment transport using the one-line (CERC) sub-model, which is responsible for this behaviour in the presence of biased input wave direction data as noted in 2.4.2 above (Chataigner et al., 2022). The pattern of planform realignment simulated here closely matched the model behaviour observed at Narrabeen by Robinet et al. (2020) using LX-Shore (a hybrid model using a grid-based one-line sub-model coupled with ShoreFor), as well as by Chataigner et al. (2022), who used a standalone one-line (CERC) model. The reasons for and implications of this issue are discussed further in 4.2.1.

3.2. Model performance with modifications applied

3.2.1. Modelled shoreline positions and Mielke's index performance scores

Model performance improved substantially once modifications were applied (Figs. 5 and 6). ShoreFor showed much more stable performance during the validation period once the erosion ratio term r was re-calculated (Mielke's index $\lambda = 0.33$ compared to $\lambda = 0.04$). For the three CERC-based models affected by wave direction bias, the mean Mielke's index across transects improved by 0.25–0.42 when re-initialised with a virtual equilibrium shoreline, and planform shoreline realignment was no longer evident in the output timeseries. IH-MOOSE, not shown in the initial results above, achieved a reasonable performance across both calibration and validation periods without additional modifications.

With modifications applied, mean performance across the five transects for all models was clustered within a Mielke's index of 0.3–0.5 over both the calibration and validation periods. The exception was CoSMoS-COAST over the calibration period, showing near-perfect performance at all transects due to data assimilation (mean $\lambda = 0.99$). Of the models without assimilation, IH-MOOSE achieved the best performance

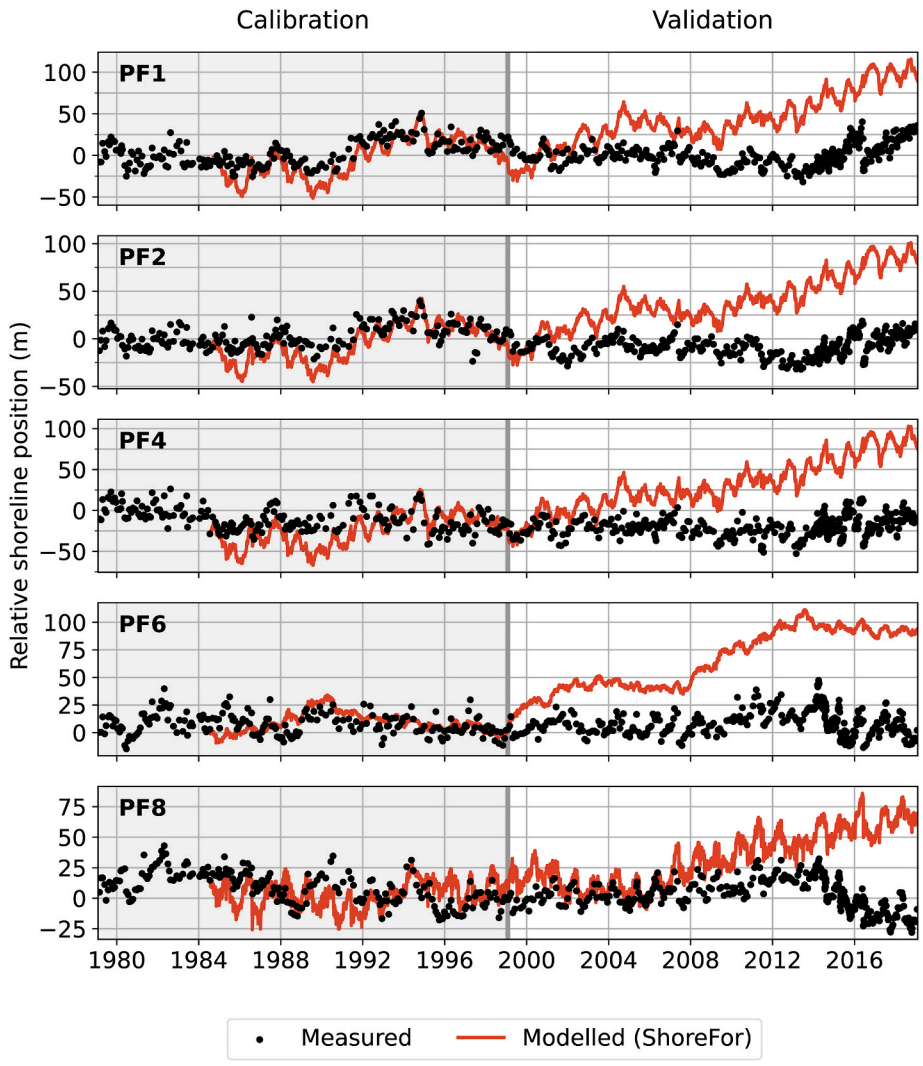


Fig. 3. Measured versus modelled shoreline positions for ShoreFor, with the erosion ratio term calculated over the calibration period and kept constant over the validation period. Note that y-axis scales differ between plot panels but grid spacing is consistent.

over the calibration period (mean $\lambda = 0.45$) without requiring modifications, although it was closely followed by ShorelineEvol ($\lambda = 0.41$) and ShoreFor ($\lambda = 0.38$). Over the validation period, the best mean performance was achieved by ShorelineEvol ($\lambda = 0.49$), but with similar scores from COCOONED ($\lambda = 0.43$) as well as CoSMoS-COAST and IH-MOOSE ($\lambda = 0.39$ for both, with data assimilation now off in CoSMoS-COAST). The relative performance of ShoreFor is notable as it is a standalone cross-shore model designed for event to interannual-scale simulations, not a hybrid framework with a sub-model of long-term behaviour, and was expected to be disadvantaged for this reason over a 40-year simulation period.

Despite similar results on average, performance varied markedly across transects for each model. Models appeared to have difficulty at different transects and in different ways. For example, ShoreFor achieved the second-highest score of the five models over the validation period at PF4, but the lowest score at PF6, where it simulated a positive (seaward) offset and underestimated event to seasonal-scale variability. CoSMoS-COAST was the best-performing model at PF8 but achieved the lowest score at PF1 due to a negative (landward) offset. IH-MOOSE achieved the highest score at PF1 and PF4, but the lowest score at PF2, where the modelled shoreline position showed a large seaward offset. In addition, all three models re-initialised with a virtual equilibrium shoreline showed poor performance at PF4 ($\lambda < 0.2$), where modelled shoreline positions were offset seaward and seasonal to interannual-scale oscillations were not captured.

Interestingly, the Mielke's index scores were higher over the validation period than the calibration period for several models at individual transects, as well as for the mean performance of COCOONED and ShorelineEvol over all five transects. This is unusual as models are typically better able to reproduce the data they are calibrated on rather than unseen data. This may be the result of: (1) a possible improvement in the quality of the hindcast wave dataset in more recent years (i.e., lower scatter compared to measured conditions), as the underlying wind and sea-ice concentration reanalysis product was able to assimilate increasing numbers of observations; and/or (2) more frequent beach profile surveys in recent years allowing for potentially higher Mielke's index scores, as Mielke's index is a correlation-based metric and may increase as more timesteps are added to the calculation.

3.2.2. *Performance across other metrics*

Taylor diagrams of correlation, centred RMSE and measured versus modelled standard deviations (i.e., magnitudes of shoreline variability) are shown in Fig. 7. The complementary metrics of bias and uncentred RMSE are shown in Fig. 8. These metrics show similar trends as noted above, although they quantify different aspects of model behaviour and highlight that different metrics favour different models. On average across all transects, model performance was again very similar, with all models achieving a mean RMSE of 15–19 m and most achieving a

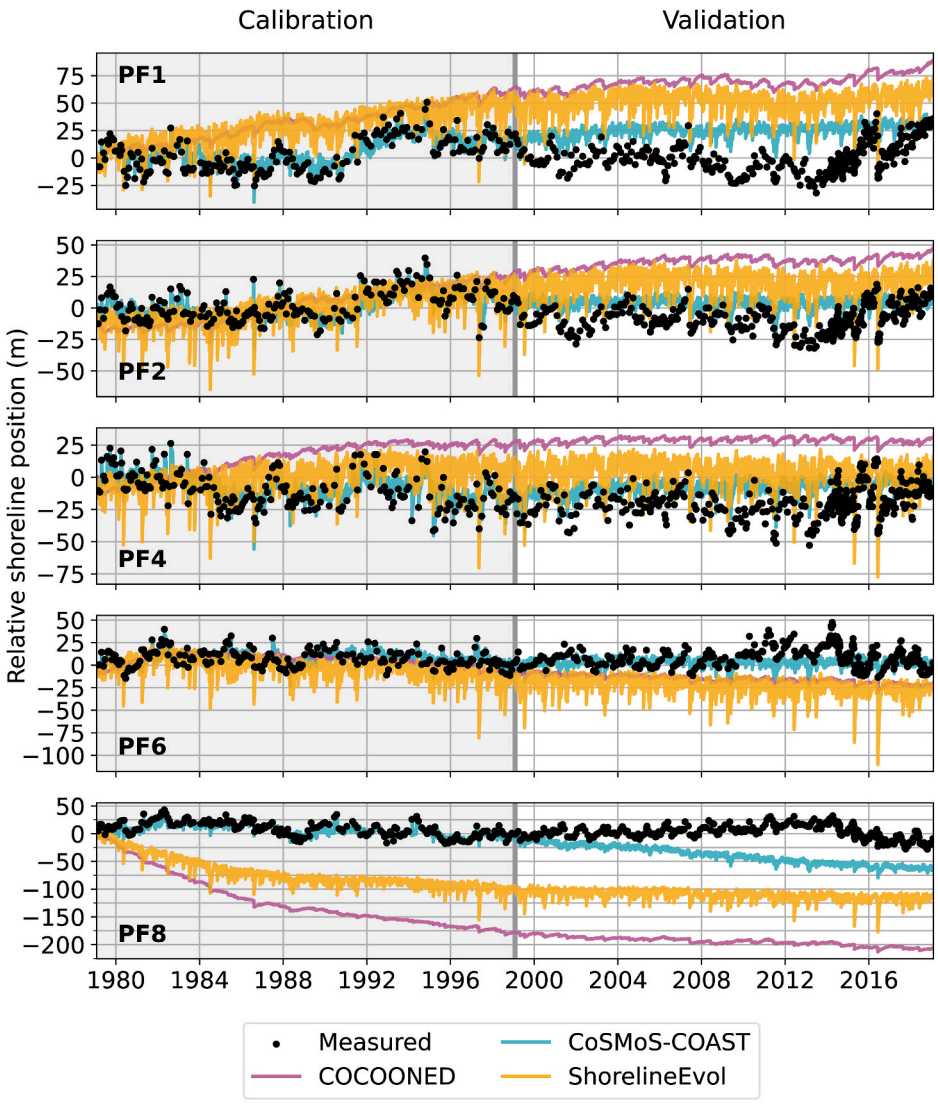


Fig. 4. Measured versus modelled shoreline positions for three models using the one-line (CERC) sub-model and affected by wave direction bias at Narrabeen. Note that y-axis scales differ between plot panels but grid spacing is consistent. In CoSMoS-COAST, data assimilation was turned on for the calibration period and off during the validation period.

correlation of ~0.6. ShoreFor achieved the lowest correlation score ($R = 0.36$) but the best mean absolute bias (5.2 m compared to 7.8–13.2 m for the other models), while CoSMoS-COAST achieved the best mean correlation ($R = 0.65$).

For individual transects, most models were generally able to achieve a CRMSE of 10–15 m and an uncentred RMSE of 10–20 m but showed significant bias or over/under-estimation of shoreline variability at specific transects, as noted above. At PF4, the three models re-initialised using a virtual shoreline notably underestimated shoreline variability and showed a 10–20 m seaward bias, as identified above from the shoreline position timeseries (Fig. 5). ShoreFor tended to capture the right degree of variability and show low bias except at PF6, where it had the largest bias of the five models (11.7 m). CoSMoS-COAST showed fairly large biases, notably at PF1 and PF2 (up to −26.5 m), but had the best correlation and closely matched the magnitude of measured shoreline variability at these transects. COCOONED and ShorelineEvol performed well at PF2 and PF6 but tended to overestimate shoreline variability at the outermost transects PF1 and PF8. IH-MOOSE performed well at PF4 but overestimated shoreline variability at the two northern transects, underestimated at the two southern transects, and showed a 35.4 m seaward bias at PF2.

3.2.3. *Performance over different timescales*

The 6-month rolling mean of measured and modelled shoreline positions (with bias removed) is shown in Fig. 9. The Mielke's index scores are also listed to quantify the ability of the models to capture this rolling mean (representing seasonal-scale and longer-term behaviour), as well as its residual component (representing shorter-term storm-driven erosion/accretion cycles). In terms of capturing shoreline behaviour over seasonal scales and longer, CoSMoS-COAST performed best, followed by ShorelineEvol and COCOONED. All three of these models use the one-line (CERC) sub-model of longshore transport. Each of these three models also scored better in capturing longer-term behaviour than they did in capturing event-scale oscillations. The notable exception to this was at the transect PF4 in the centre of the embayment, where they achieved a score of near-zero in capturing longer-term behaviour. IH-MOOSE, which uses the semi-empirical sub-model of beach rotation, J21, did not score as well as the top three CERC-based models (except at PF4) for longer-term behaviour. However, it out-performed ShoreFor, the standalone cross-shore model, at all transects except PF6. Interestingly, the EOF analysis in 3.2.4 shows that PF6 is the most dominated by cross-shore processes over the validation period.

In terms of capturing event-scale erosion/accretion cycles, ShorelineEvol and IH-MOOSE tended to show the best performance, with

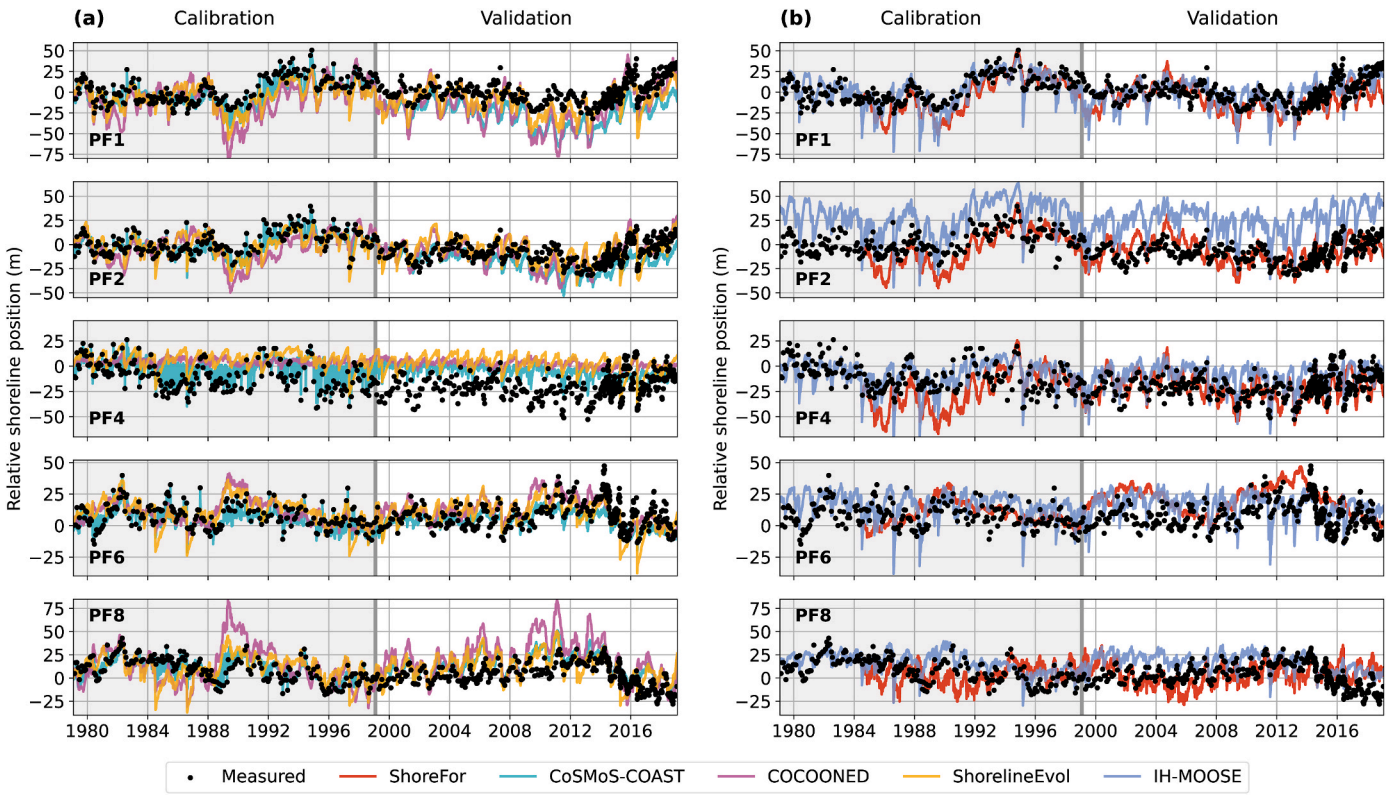


Fig. 5. Measured versus modelled shoreline positions for the five models, separated for clarity between panels: (a) models using the one-line (CERC) sub-model, initially affected by wave direction bias; (b) other models. Note that y-axis scales differ between transects but grid spacing is consistent. In CoSMoS-COAST, data assimilation was turned on for the calibration period and off during the validation period.

ShorelineEvol achieving relatively consistent scores across transects ($\lambda \sim 0.6$) while the performance of IH-MOOSE was best at the central transect PF4 ($\lambda = 0.74$) and decreased towards the outer transects at the edges of the embayment. This may relate to the calibration approach: the cross-shore sub-model parameters of IH-MOOSE were tuned at PF4, while those of ShorelineEvol were tuned based on average performance across the five transects. Surprisingly, neither CoSMoS-COAST nor ShoreFor performed as well as these two models (except at PF8), despite having alongshore-varying parameters (i.e., parameters calibrated separately at each transect) compared to the single set of parameters tuned for the whole embayment in ShorelineEvol and IH-MOOSE. COCOONED generally achieved lower scores in capturing short-term oscillations than the other models, which is attributed to the sub-optimal calibration of its cross-shore parameters noted in 2.3.4. Interestingly, the three models with one-line (CERC) sub-models were able to capture short-term oscillations at PF4 with a similar degree of success as they each did at other transects, despite near-zero scores for capturing longer-term behaviour, indicating that poor performance at this transect is caused by the longshore sub-models rather than the cross-shore sub-models as initially hypothesised. This is further supported by the EOF analysis in the next section (3.2.4), which shows that this transect has a large contribution from longshore coherent processes over the validation period.

Comparing ShoreFor against IH-MOOSE at PF4 provides an opportunity to compare ShoreFor directly against Y09, the cross-shore sub-model of IH-MOOSE. IH-MOOSE was set up with Y09 calibrated at PF4, and with PF4 acting as the pivot point or node of the rotation sub-model, meaning the output of IH-MOOSE at PF4 is driven entirely by Y09 alone. IH-MOOSE (Y09) performed better than ShoreFor in capturing both timescales of variability over the 20-year period. IH-MOOSE (Y09) also performed better than all other models at this profile.

More generally, the scores of ShoreFor were somewhat surprising across the five transects. At PF1–PF4, the model captured both longer-term and event-scale behaviour with similar scores of $\lambda \sim 0.4$. At PF6, which is shown in the following section to be the most cross-shore dominated transect over this period, ShoreFor achieved a score of $\lambda = 0.60$ for the longer-term behaviour, on par with the three models using one-line (CERC) sub-models, and better than IH-MOOSE. However, ShoreFor only scored $\lambda = 0.13$ for the event-scale behaviour. Conversely, at PF8, ShoreFor achieved $\lambda = 0.07$ for longer-term shoreline behaviour, but $\lambda = 0.73$, the best score of the five models, in capturing event-scale oscillations. This suggests ShoreFor may have had difficulty capturing multiple concurrent processes that drive shoreline change over different timescales within the multi-decadal period. Although this may be expected for a standalone cross-shore model applied to a long simulation period compared to the hybrid framework models (which include multiple sub-models to capture multiple timescales of behaviour), this has also been identified as a specific shortcoming of ShoreFor by previous authors over shorter simulation periods (e.g., Schepper et al., 2021; Tran et al., 2021) and is discussed in 4.3.2.

3.2.4. Performance in replicating EOF decomposition of measured shoreline positions

The first two modes of the EOF decomposition for both measured and modelled shoreline positions are shown in Fig. 10. Over the calibration period 1979–1999, the EOF decomposition of measured shoreline positions (Fig. 10a, black lines) is similar to the long-term average at Narrabeen identified in time-invariant analyses by Harley et al. (2011a) over the period 1976–2010 and Ibaceta et al. (2023) over 1976–2019. The cross-shore coherent mode is dominant, explaining 54% of shoreline variability, while the longshore coherent mode accounts for 32% of variability. The spatial EOF of the cross-shore coherent mode shows a peak at PF4 and higher values at the northern end of the embayment compared to the southern end, while the longshore-coherent mode shows a sign reversal just north of PF4.

The corresponding decomposition of modelled shoreline positions

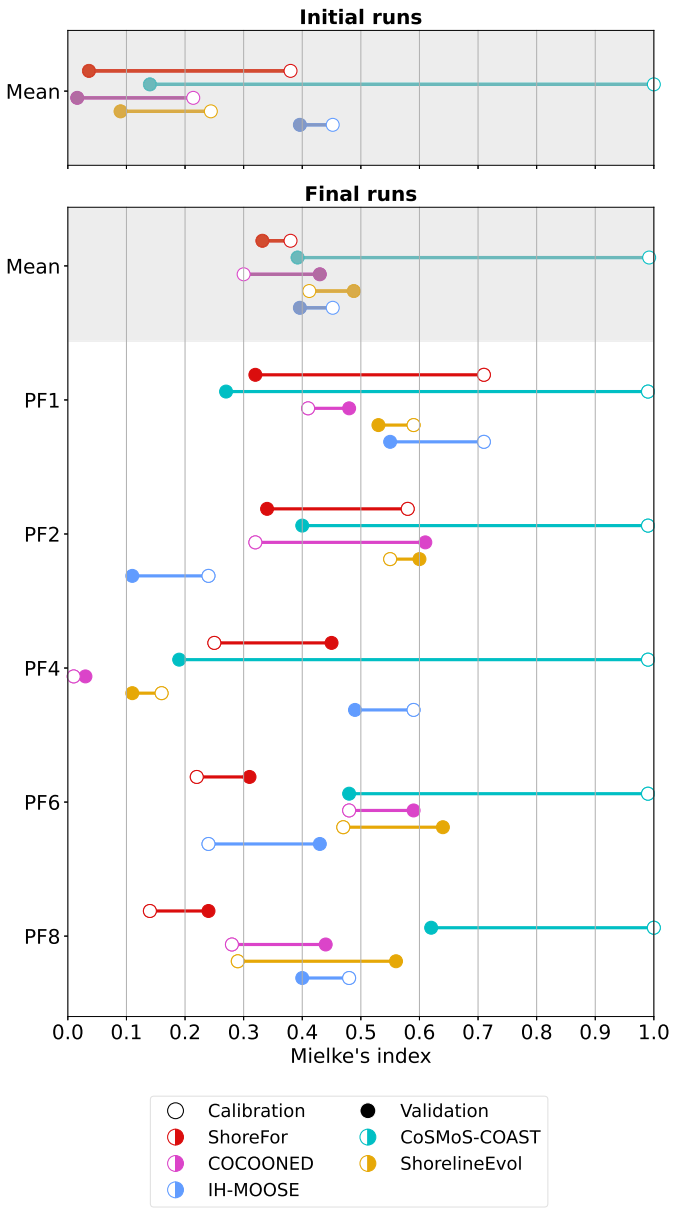


Fig. 6. Mielke's index values of each model at each transect, as well as the mean across transects (shaded). Open circles show calibration performance and filled circles show validation performance. In CoSMoS-COAST, data assimilation was turned on for the calibration period and off during the validation period. The initial runs refer to the model outputs described in 3.1, i.e., before a virtual equilibrium shoreline was used for ShorelineEvol, COCOONED and CoSMoS-COAST, and before the erosion ratio term in ShoreFor was re-calculated over the validation period.

over the calibration period showed notable differences between models. With data assimilation turned on for calibration, CoSMoS-COAST replicated all components of the measured EOF decomposition almost perfectly over this period, as expected from its near-perfect replication of the measured timeseries itself (Fig. 5). In terms of spatial EOFs, ShoreFor roughly replicated the cross-shore coherent mode but not the longshore coherent mode. However, it was calibrated over a 15-year rather than 20-year period (1984–1999) and cannot be directly compared to the 1979–1999 decomposition. The remaining three models, with alongshore-fixed parameter values (see Table 3), were able to replicate the longshore coherent spatial EOF fairly closely, but not the cross-shore coherent EOF. In terms of the percentage of variance explained by each mode, ShorelineEvol replicated the measured percentages reasonably well. Conversely, the variability of COCOONED's output was overwhelmingly dominated by the longshore-coherent mode (90%), attributed to the poor calibration and resulting underperformance of its cross-shore sub-model. The reverse was true for IH-MOOSE and ShoreFor, with 90% and 87% of variability explained by the cross-shore coherent mode, respectively. These two models do not explicitly simulate longshore sediment transport (at all in the case of ShoreFor, and only semi-empirically through an equilibrium beach rotation sub-model in the case of IH-MOOSE).

For the validation period 1999–2019 (Fig. 10b), the EOF decomposition of measured shoreline positions indicates very different shoreline dynamics compared to the calibration period, consistent with the analysis by Ibaceta et al. (2023) showing high variability in these dynamics at Narrabeen over interannual timescales. Over the validation period, the longshore coherent mode is dominant over the cross-shore coherent mode, explaining 52% of variability compared to 41%. The spatial EOFs show a reversal in their alongshore magnitudes (e.g., the cross-shore coherent mode is now greater in magnitude at the southern end of the embayment) as well as changes to their shape. In particular, the spatial EOF of the longshore-coherent mode is more complex than the relatively linear shape it showed over the calibration period, with a high magnitude at PF4 and the node further south, closer to PF6.

The EOF decompositions of modelled shorelines over the validation period are intriguing. The only model that replicated both spatial EOFs with reasonable accuracy was ShoreFor, which was unexpected as it does not explicitly resolve longshore processes or consider wave direction as an input. However, ShoreFor was also the only model with at least one time-varying parameter, with the erosion ratio r re-calculated as described in 2.4.1. The parameters of CoSMoS-COAST varied during data assimilation but converged to fixed values towards the latter half of the calibration period and were fixed at these values for the validation period. This model showed similar, but not identical, spatial EOFs over the validation period compared to its spatial EOFs over the calibration period. The spatial EOFs of the remaining models, with entirely time-invariant parameters, appeared to be almost identical to their calibration period EOFs. The longshore coherent spatial EOFs of these models did show a reversal in their alongshore magnitude, consistent with measurements, but not a change in their shape, while the shape of the cross-shore coherent EOFs remained entirely unchanged. As wave direction is used as an input in the longshore sub-models but not the cross-shore sub-models, we hypothesise the alongshore reversal in the observed EOFs may have been driven by a shift in prevailing wave direction between the two time periods. In the case of ShoreFor, recalculating r (independently at each transect) appeared to allow it to capture the shift in shoreline dynamics without explicitly accounting for longshore processes or considering wave direction as an input.

Comparing the proportion of variability explained by each mode between the calibration and validation periods shows a similar pattern (right-most panels of Fig. 10). For the models with entirely time-invariant parameters, COCOONED, ShorelineEvol and IH-MOOSE, these proportions appear to have been set during the calibration period and then maintained into the validation period. Conversely, CoSMoS-COAST and ShoreFor did show a reversal in the dominant mode, consistent with measurements, although the proportions were skewed too far towards the longshore coherent mode (surprisingly for ShoreFor, a model of only cross-shore processes).

Finally, comparing measured and modelled temporal EOFs (middle panels of Fig. 10) shows that models typically achieved higher correlations with the measured longshore coherent mode than the cross-shore coherent mode for both time periods. COCOONED in particular showed good correlations with the longshore coherent temporal EOFs over both time periods but notably low correlations with the cross-shore coherent EOFs, again attributed to the poor calibration of its cross-shore sub-model. The exception was ShoreFor, which showed a higher correlation to the cross-shore coherent EOF, as expected for a standalone cross-shore model. However, it still showed a reasonable correlation to

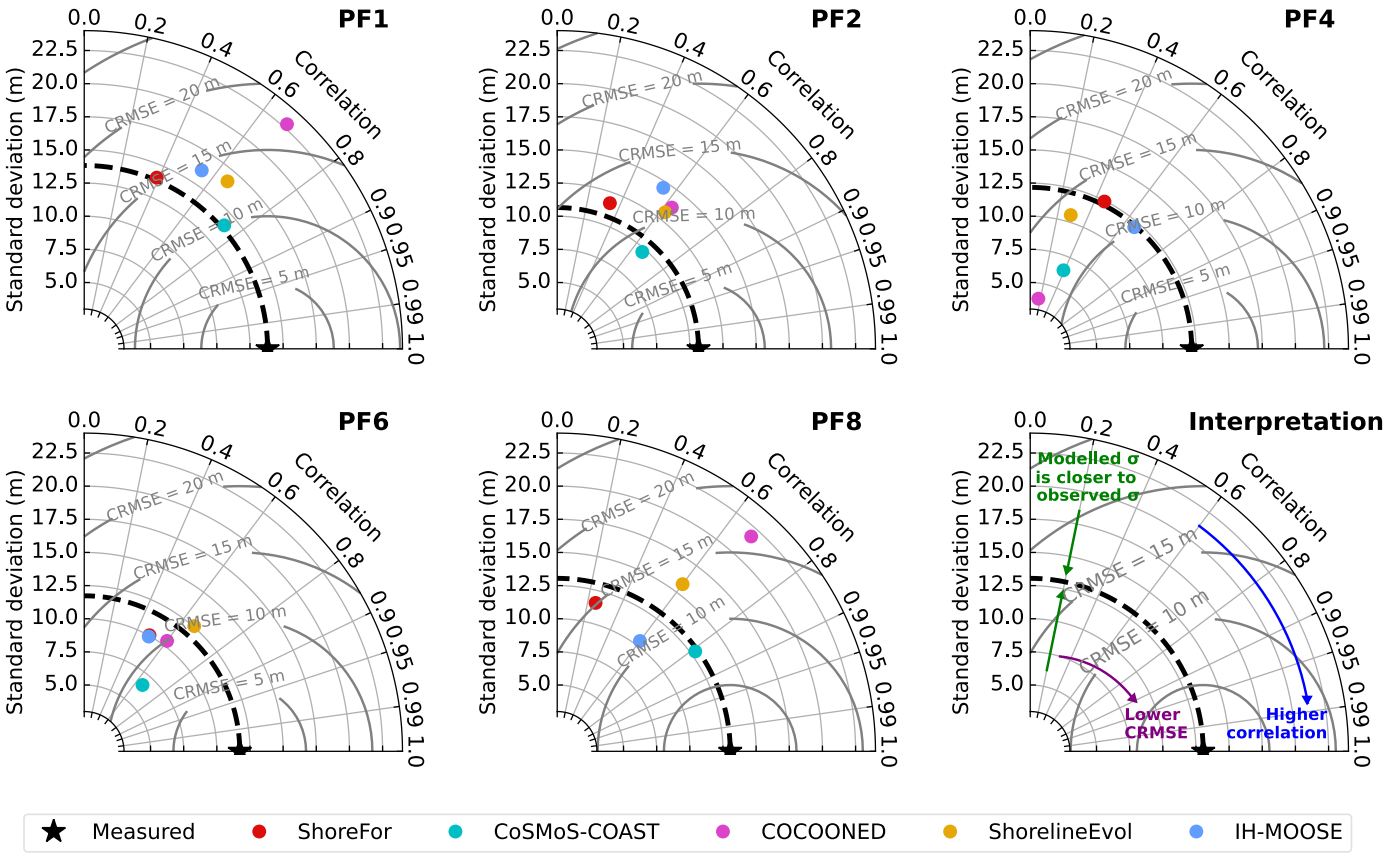


Fig. 7. Taylor diagrams of model performance by transect over the validation period 1999–2019. The correlation between modelled and measured shoreline positions is shown by the angle from the vertical y-axis. Radial axes show the standard deviation of modelled or measured shoreline positions moving outwards from the bottom left. Additional contours show centred root mean squared error (CRMSE), i.e., RMSE with bias removed. Models with better performance are located towards the bottom of each plot (i.e., higher correlation with measurements) and closer to the dashed black line (i.e., the standard deviation or modelled degree of shoreline variability is closer to the measured variability).

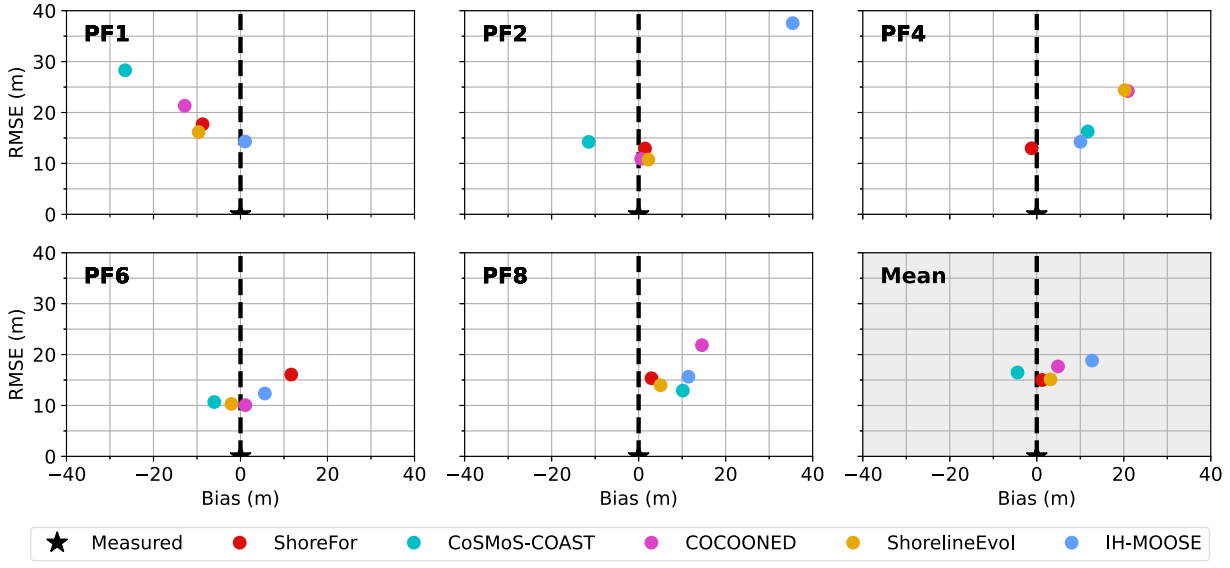


Fig. 8. Complementary metrics to those shown in the Taylor diagrams (Fig. 7): bias and (uncentred) root mean squared error (RMSE).

the longshore coherent EOF, particularly over the validation period. This supports the view that the longshore coherent EOF at Narrabeen is driven by a mixture of cross-shore as well as longshore transport (Harley et al., 2011a, 2015), but also suggests that the alongshore-varying and time-varying parameters used in ShoreFor here may allow it to replicate longshore processes to some extent even though the model is not explicitly formulated to capture these.

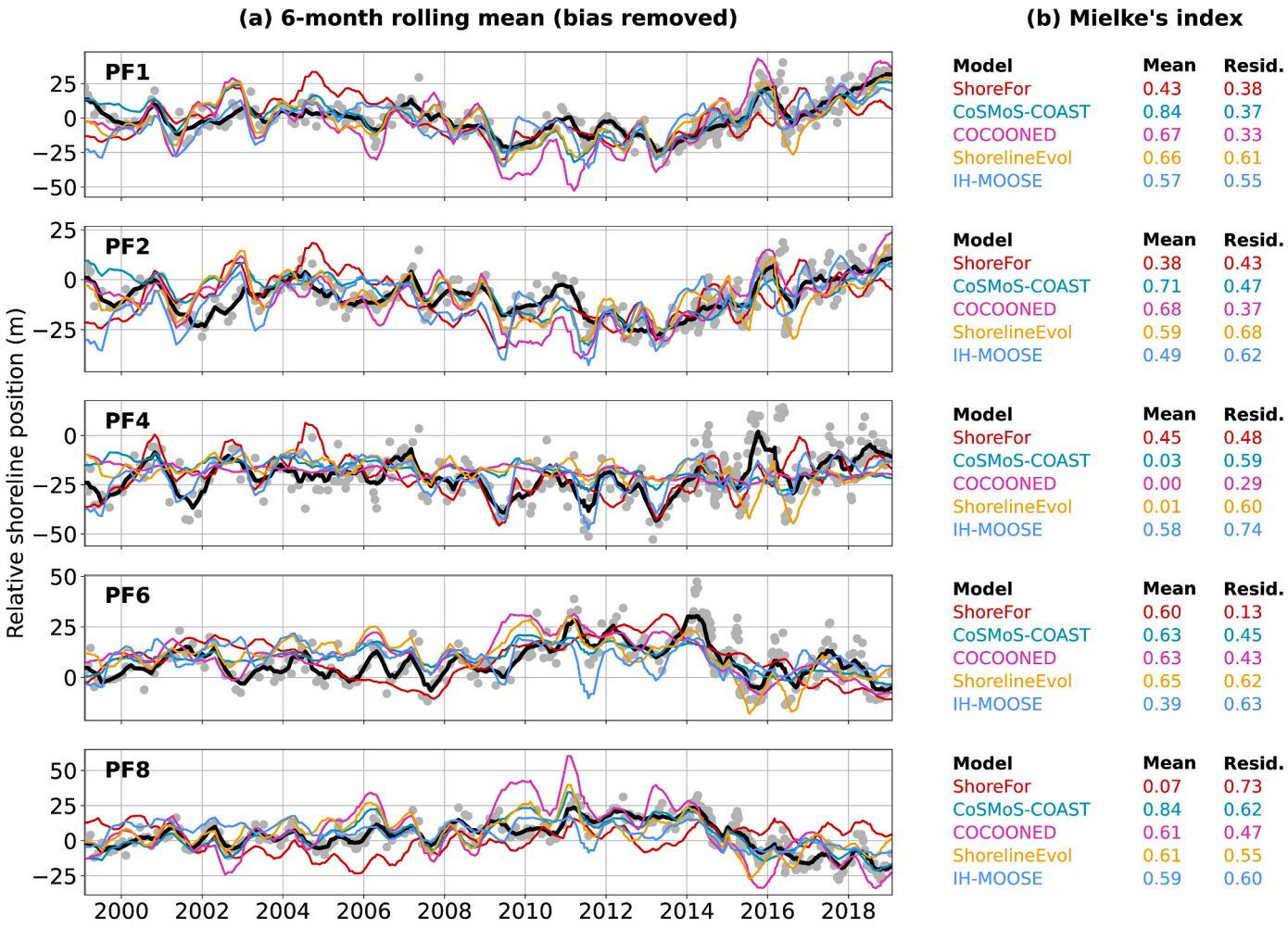


Fig. 9. Six month rolling mean of measured (black) and modelled (coloured) shoreline positions. The individual (i.e., un-averaged) measured shoreline positions are shown in grey. The modelled positions were subset to timesteps concurrent with measurements, and bias (mean error) was removed before calculating the rolling mean. The timeseries of corresponding residuals are not shown, but Mielke's index values for both the rolling mean and residuals are listed on the right.

4. Discussion

4.1. Overall performance of models at Narrabeen

The overall performance achieved by the five models was promising given the simplicity of the models compared to the complexity of shoreline dynamics at Narrabeen. For example, all five models assume an equilibrium beach profile shape that remains constant in time and simply translates landwards or seawards as the shoreline position changes, and none of the models explicitly consider sandbar dynamics or storage of sediment in the nearshore following storms. In contrast, the profile shape of the upper beach and surf zone is very dynamic at Narrabeen, with rapid onshore-offshore movement of the sandbar and correspondingly rapid transitions through the intermediate morphodynamic beach states (Wright et al., 1985; Davidson et al., 2013). The position of the bar has been shown to modulate shoreline variability by controlling wave energy dissipation and therefore erosion during storms (Harley et al., 2009), as well as sediment availability and accretion rates during subsequent recovery periods (Phillips et al., 2017).

Additional simplifying assumptions were made in setting up the models. A constant berm height and closure depth were assumed to define the landward and seaward limits of modelled sediment erosion or deposition (with COCOONED also accounting for dune erosion). However, Harley et al. (2022) found that extreme storms appear to mobilise sediment from the lower shoreface, contributing to significant net accretion over the following 12 months compared to pre-storm volumes. The models were also set up with a uniform grain size across the embayment (refer to 2.3.3), meaning that the equilibrium beach profile shape was also assumed to be constant alongshore. However, due to the sheltering effect of the headlands bounding the embayment (Fig. 1b) and the resulting gradient in wave energy exposure, the modal beach state also varies alongshore, with the more sheltered southern end tending towards more reflective profiles (Wright and Short, 1984) and also exhibiting differing sandbar behaviour to the more exposed northern end (Harley et al., 2015). Finally, urban development and ad hoc armouring at the southern end of the embayment were not simulated, although these do truncate the maximum inland extent of profile fluctuations during severe storms at PF6 (Short and Trembanis, 2004). Nonetheless, the models did not show an obvious degradation in performance at PF6 compared to the other transects, with three of five models achieving their best or second-best Mielke's index score at this transect (Fig. 6).

The complexity of shoreline dynamics at Narrabeen can also be considered in terms of the physical processes underlying the first two EOF modes, proposed by Harley et al. (2011a, 2015) and Ibaceta et al. (2023). The cross-shore coherent mode is considered to correspond to cross-shore shoreline oscillations during storm erosion/recovery cycles, which vary in magnitude alongshore due to the alongshore gradient in wave energy exposure (Harley et al., 2011a, 2015). The longshore coherent mode has been linked to at least three possible physical

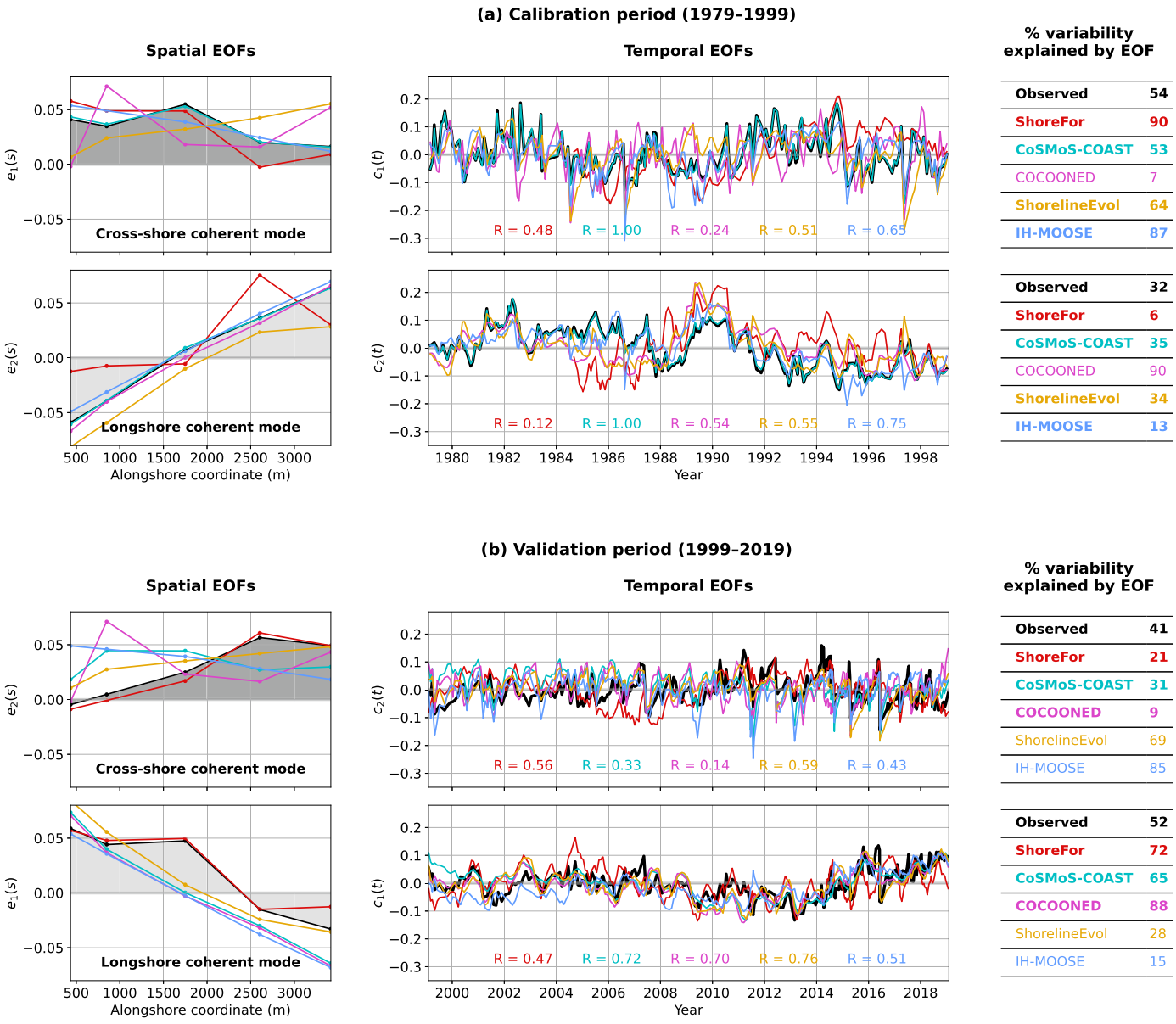


Fig. 10. The first two empirical orthogonal functions (EOFs) for measured (black) and modelled (coloured) shoreline positions at Narrabeen over the calibration period (a) and validation period (b). For each period, the shoreline position p at alongshore coordinate s and time t is given by $p(s,t) = \sum_{k=1}^{N} w_k e_k(s) c_k(t)$, where w_k is a weighting factor, $e_k(s)$ is the kth spatial EOF and $c_k(t)$ is the kth temporal EOF. The EOFs are visualised so that the cross-shore coherent mode always appears in the top panel and the longshore coherent mode appears in the bottom panel, allowing the shapes to be compared easily. This means for some models, the first modelled EOF (k = 1) is plotted against the second measured EOF (k = 2) and vice versa. Models with the two EOFs in the correct order (i.e., matching observations) are bolded in the table on the right.

processes, including both longshore and cross-shore sediment transport: (1) seasonal to interannual-scale reversals in longshore transport direction ('classic' beach rotation), (2) opposing sandbar behaviour at each end of the embayment, and (3) longshore redistribution of sediment across the upper shoreface during extreme storms from oblique angles, and subsequent redeposition of sand in a different part of the embayment to that from which it was eroded (Harley et al., 2011a, 2015; Ibaceta et al., 2023). The models tested here are designed to resolve the mechanism underlying the cross-shore coherent mode, and only the first of the three longshore coherent mechanisms (and ShoreFor only resolves the former, not the latter). Despite this, the models performed surprisingly well in reproducing both modes, particularly over the calibration period.

The overall performance scores of the models at Narrabeen are comparable to those of Montaño et al. (2020), who evaluated 10 reduced-complexity or hybrid models (including ShoreFor, CoSMoS-COAST and COCOONED) against alongshore-averaged shoreline position data at Tairua Beach, New Zealand. The 10 models achieved Mielke's index values ranging from 0.23 to 0.62 over a 15-year calibration period (excluding two models with data assimilation turned on during this period) and 0.23–0.51 over a 3-year validation period. In the present study, models achieved Mielke's index values of 0.30–0.45 averaged over five transects during the 20-year calibration period (excluding the near-perfect score of CoSMoS-COAST, with data assimilation turned on). Over the subsequent 20-year validation period, the corresponding range across the five models was 0.33–0.49. Unlike the present study, models in Montaño et al. (2020) were calibrated using RMSE or normalised mean squared error (NMSE), and their Mielke's index scores would be expected to be higher if it was used to calibrate models directly. Nonetheless, the similar range of scores observed over

the much longer validation period considered here (i.e., 20 years rather than 3 years) does provide some confidence in the ability of these types of models to maintain consistent performance over multi-decadal simulations (although the different sites mean that the scores are not directly comparable, and testing performance over multiple simulation period durations at a single site remains an avenue for future work).

However, overall performance scores remain relatively low from the perspective of using the models to forecast shoreline change for coastal management. This remained the case even after modifications to improve performance were applied to four of the five models. Model performance also varied substantially between the five transects evaluated in the present study, with models replicating measured shoreline variability well at some transects (up to $\lambda = 0.64$ and RMSE = 10.1 m over the validation period) and very poorly at others ($\lambda = 0.01$, RMSE = 37.6 m). Although the simplifications described above are expected to have played some role, the performance of the different models appeared to be particularly limited by different factors that depended on their cross-shore and longshore sub-models. These are discussed in turn below.

4.2. Sub-models of longshore processes

Comparing the results of the standalone cross-shore model ShoreFor to the hybrid framework models suggests that coupling sub-models of both cross-shore and longshore processes does improve performance at Narrabeen over a multi-decadal simulation period, but the additional complexity introduces additional sources of error. ShoreFor alone performed surprisingly well compared to the hybrid frameworks, particularly over the calibration period (Fig. 6). This likely reflects the strength of a cross-shore model at a site that has been shown to be dominated by cross-shore processes on average (Harley et al., 2011a; Ibaceta et al., 2023), but also that additional issues limited the performance of the more complex models. The hybrid frameworks generally outperformed ShoreFor based on mean Mielke's index values (i.e., overall performance), their ability to capture >6-monthly shoreline variability (Fig. 9) and their correlation with the longshore-coherent mode of shoreline variability (Fig. 10). However, ShoreFor tended to show more consistent performance across transects, while all three CERC-based models performed very poorly at PF4 and IH-MOOSE showed a large bias at PF2. These issues are respectively attributed to the difficulty of correcting for wave direction bias, and coupling rotation and cross-shore sub-models, as elaborated below.

4.2.1. One-line (CERC) sub-model of longshore sediment transport

The outputs of the three hybrid framework models using the one-line (CERC) sub-model, CoSMoS-COAST, COCOONED and ShorelineEvol, were initially dominated by large-magnitude planform shoreline realignment (Fig. 4), similar to the one-line model behaviour observed by Robinet et al. (2020) and Chataigner et al. (2022) at Narrabeen. This issue has also been noted at other sites such as the U.S. Pacific Northwest, where erroneous modelled shoreline realignment was observed unless a virtual equilibrium shoreline correction was applied (Anderson et al., 2018; Antolínez et al., 2019). Recently, Chataigner et al. (2022) undertook a sensitivity analysis of the one-line (CERC) model and showed that it is very sensitive to small biases in mean direction. They found that increasing the standard deviation of direction bias by only 1° increased the standard deviation of shoreline position errors by 5 m at Narrabeen.

The results here extend the work of Chataigner et al. (2022) in two ways. First, this issue dominates the behaviour of one-line (CERC) models where biases are present, including when the models are used as part of a hybrid model framework. Coupling additional shoreline change sub-models does not resolve or compensate for this issue. This was also observed by Robinet et al. (2020), who applied the hybrid framework LX-Shore (a one-line model coupled with ShoreFor) at Narrabeen. Second, the issue cannot be "calibrated out" by tuning only the free parameters, regardless of whether optimisation or data assimilation is used, unless the longshore transport rate parameter K (Equation (7)) is set to very low values. However, doing so would suppress the contribution of the one-line model and prevent real longshore change from being captured, and in the case of hybrid frameworks, would likely result in non-physical parameter values calibrated in the cross-shore sub-models in compensation. In the case of data assimilation, realignment is supressed during the calibration period as the modelled shoreline position is continuously corrected towards observations. However, once data assimilation is turned off for validation or forecasting, shoreline realignment resumes.

Prior to these recent studies, the CERC formula had received criticism but critiques focused largely on uncertainty around the model free parameter K (e.g., Pilkey and Cooper, 2002). Alternative formulae proposed used different combinations of constants (Kamphuis, 1991; Mil-Homens et al., 2013) or included wind and tide effects as well as wave conditions (Bayram et al., 2007). However, the sensitivity to wave direction bias arises because the CERC formula assumes the rate of shoreline change depends on the angle between incident wave fronts and the present shoreline (the $\sin[2(\theta_b - \alpha_s)]$ term in Equation (7)). This term is maintained in alternative formulae, meaning the sensitivity to bias would be maintained. This was shown to be the case by Robinet et al. (2020), who used the Kamphuis (1991) approach rather than the CERC formula in their one-line sub-model. Therefore, this sensitivity remains a concern as one-line models (with CERC or similar formulae) are widely used, both in hybrid frameworks and standalone models of longshore change such as GENESIS (Hanson, 1989; Hanson and Kraus, 1989), CEM (Ashton and Murray, 2006) and ShorelineS (Roelvink et al., 2020).

To resolve the issue, either the input wave direction data must be corrected, as proposed by Chataigner et al. (2022), or the output modelled shorelines must be corrected. The latter approach was implemented here using a virtual equilibrium shoreline, similar to Anderson et al. (2018), Antolínez et al. (2019) and Robinet et al. (2020). While model performance substantially improved, the results were not perfect. All three models corrected with this approach showed poor performance at PF4 (e.g., Fig. 6). The 6-month rolling mean and EOF decompositions (Figs. 9 and 10) indicated this was driven by the longshore sub-models, rather than the cross-shore sub-models having difficulty with a long simulation period as was initially hypothesised. The virtual equilibrium shoreline also introduced a large offset in the mean shoreline position at this transect relative to observations. Additionally, in the case of CoSMoS-COAST, the manually-generated equilibrium shoreline may have over-compensated for wave direction bias and contributed to some offset in the opposite direction (this may have also occurred for the other two models to a lesser degree, at the transects at the edges of the embayment). However, the method proposed by Chataigner et al. (2022) to correct the wave data directly using Monte Carlo simulations was not computationally feasible for the models used here. Identifying the best way to correct for this issue remains an avenue for future work.

However, if corrections are made for wave direction bias, it appears that the high sensitivity of one-line models to mean wave direction allows them to successfully reproduce beach rotation on embayed coastlines, even when rotation is driven by subtle shifts in wave climate. The three hybrid frameworks using this approach were able to replicate Narrabeen's seasonal planform rotations with a reasonable degree of success, which Harley et al. (2011a) showed are driven by relatively small seasonal shifts involving only 24% of the wave climate together with some degree of cross-shore processes. Likewise, Anderson et al. (2018) found a one-line (CERC) model was able to replicate the gradual rotation of embayments in the Pacific Northwest over multiple decades, which the authors suggested were driven by phase shifts of the Pacific Decadal Oscillation. This supports the use of one-line models to assess how shorelines will respond to shifts in mean wave direction under climate change (Ranasinghe, 2016), although given the large

uncertainty associated with wave climate projections (e.g., Zarifsanayei et al., 2022), their sensitivity may result in unreasonably large uncertainty bounds on shoreline change projections.

4.2.2. J21 sub-model of beach rotation

The longshore sub-model J21 of the hybrid framework IH-MOOSE represents a simpler approach than the one-line (CERC) sub-model discussed above. Rather than explicitly simulating longshore sediment transport volumes, it links wave forcing directly to morphological change (Hunt et al., 2023). Interestingly, this sub-model did not require modifications to produce a stable output, achieved the best Mielke's index score of the non-assimilated models during the calibration period, and performed comparably over the validation period overall to the more complex models without requiring modifications for wave direction bias. This highlights the potential value of simpler models, particularly where adequate data are not available (Cowell et al., 1995; van Maanen et al., 2016).

However, the simplicity comes with a trade-off of reduced generality (Hunt et al., 2023), with J21 (and IH-MOOSE) being applicable only to embayed beaches. Additionally, converting the mean planform orientations simulated by J21 to absolute shoreline positions as part of a hybrid framework requires the embayment planform to be parameterised into a specific shape. IH-MOOSE assumes a parabolic planform, following the equation developed by Hsu and Evans (1989). Fitting this equation to a site requires manual calibration effort together with additional nearshore wave data beyond bulk parameters at a single depth (González and Medina, 2001; Jaramillo et al., 2021b). This also means the performance of IH-MOOSE depends on the accuracy of the planform parameterisation as well as the performance of its two sub-models. The parameterisation used here, a parabolic planform south of PF4 combined with a linear section north of PF4, followed Jaramillo et al. (2021b). While reasonable, it produced poor performance at PF2, where the modelled shoreline showed a ~35 m seaward offset on average over the validation period (Fig. 5), similar to the result observed by Jaramillo et al. (2021b). This may have been resolved with a different planform parameterisation.

4.3. Sub-models of cross-shore processes

It appears that the cross-shore sub-models contributed more to producing relatively low overall performance scores than the longshore sub-models, as the hybrid frameworks tended to capture the longshore coherent mode better than the cross-shore coherent mode. This is to be expected, as the cross-shore sub-models were designed for seasonal to interannual-scale simulations and have rarely been evaluated over a multi-decadal period. Additionally, the models had difficulty capturing the extremes of storm-driven erosion/accretion cycles (frequently over- or underestimating shoreline variability at each transect; Fig. 7), and the hybrid frameworks scored lower in capturing <6-monthly oscillations than seasonal and longer-term behaviour (Fig. 9). This supports recent work indicating that capturing the extremes of short-term shoreline oscillations remains a challenge for cross-shore models based on equilibrium principles, and machine learning models may present a promising alternative (Montaño et al., 2020; Gomez-de la Peña et al., 2023).

4.3.1. Alongshore-varying parameters in cross-shore sub-models

Alongshore-fixed parameters in the cross-shore sub-models (i.e., one set of parameters for the entire embayment; Table 3) may have contributed to poor performance. The wave energy exposure varies alongshore at Narrabeen, together with the magnitude and dominance of cross-shore processes (Harley et al., 2011a, 2015), and optimal cross-shore parameters may be expected to vary likewise. Although this variability should in theory be captured by the wave forcing terms of the models rather than the parameter values, previous studies suggest that this is not the case in practice. Splinter et al. (2014) found the optimal parameters of ShoreFor to vary across the five transects at Narrabeen, with the optimal value of the response rate parameter c in particular 4–6 times larger at PF8 compared to PF1. More generally, the authors found model parameters to strongly depend on the mean dimensionless fall velocity across sites, which varies alongshore at Narrabeen. Similarly, Yates et al. (2009, 2011) found the optimal parameters of Y09 to vary alongshore at multiple sites in California, USA.

In the present study, the role of alongshore-fixed cross-shore parameters was inconclusive. The models with alongshore-fixed parameters, ShorelineEvol and IH-MOOSE, tended to outperform the models with alongshore-varying parameters, CoSMoS-COAST and ShoreFor, in capturing <6-monthly oscillations (Fig. 9). ShorelineEvol (alongshore-fixed) also showed the highest correlation to the cross-shore coherent temporal EOF, while CoSMoS-COAST (alongshore-varying) showed the lowest (Fig. 10). However, these may be spurious results driven by other factors. In the case of CoSMoS-COAST, a single rate parameter ΔT is used for both erosion and accretion events rather than separate C^+ and C^- parameters which may lower performance (although Yates et al. (2011) found this decreased model R^2 by only <10% for Californian sites). In the case of ShoreFor, lower performance scores may be a model-specific issue relating to the long simulation period (discussed below). Conversely, the performance of IH-MOOSE (alongshore-fixed) at capturing <6-monthly oscillations was best at PF4, the transect that its cross-shore parameters were calibrated against, and performance scores decreased towards the edges of the embayment, suggesting alongshore-varying parameters may have improved performance. Additionally, only models with alongshore-varying parameters were able to accurately replicate the shape of the cross-shore coherent spatial EOF along the embayment (Fig. 10). However, it is noted that unless data assimilation is used (e.g., CoSMoS-COAST), implementing alongshore-varying parameters substantially increases the computational cost and difficulty of calibration.

4.3.2. ShoreFor compared to Y09 and MD04

All three (sub-)models of short-term cross-shore processes considered here, Y09, MD04 and ShoreFor, are based on the equilibrium principles of Wright and Short (1984). At each timestep, the shoreline evolves towards an equilibrium position at a rate that depends on the present degree of disequilibrium and the wave energy available to transport sediment. The three sub-models differ conceptually in how disequilibrium is calculated, discussed in detail by Vitousek et al. (2021) and briefly summarised as follows. In both Y09 and MD04, the rate of shoreline change depends on wave forcing as well as the present shoreline position, which acts as a 'damping' term and limits the effect of wave forcing. This means that as a beach is increasingly eroded (accreted), it becomes more difficult to erode (accrete) further. Conversely, the rate of shoreline change in ShoreFor depends only on present and antecedent wave conditions, not the shoreline position.

This means that the shoreline position modelled by Y09 and MD04 tends to oscillate around a fixed value, while ShoreFor allows for large fluctuations of the shoreline and its output only oscillates around a fixed position if erosive and accretionary conditions maintain a consistent balance over time. Therefore, if long-term trends in wave energy are present, ShoreFor simulates mean-trend erosion or accretion, while Y09 and MD04 simulate short-term oscillations of a greater magnitude but not mean-trend change. This behaviour was demonstrated analytically by Vitousek et al. (2021) and through simulations by D'Anna et al. (2021a, 2022). However, ShoreFor's potential for unstable behaviour over multi-decadal simulations in the presence of a non-stationary wave climate was not illustrated until the recent work of Ibaceta et al. (2022). They applied the model to a 14-year calibration and a 14-year validation period using satellite-derived shorelines (Vos et al., 2019b) from the Gold Coast, Australia. They found that ShoreFor diverged substantially from observations with large mean-trend erosion, with the model being unable to maintain stable shoreline behaviour as the balance of erosive and accretionary wave conditions shifted between the calibration and validation periods. In the present study, similar mean-trend divergence

was observed during initial simulations with ShoreFor (but not Y09 or MD04), suggesting similar non-stationarity was present in Narrabeen's wave climate between the calibration and validation periods used here. Notably, Ibaceta et al. (2022) used the form of the model initially proposed by Davidson et al. (2013), where separate erosion and accretion constants were both calibrated as free parameters, rather than the version of Splinter et al. (2014) with one free parameter and the erosion ration term r, as used in the present study. This implies the issue is not strictly related to r itself, but a more fundamental aspect of the model.

In the present study, this behaviour was resolved using a first-pass solution of re-calculating ShoreFor's erosion ratio parameter r over the validation period, rather than using the value obtained during calibration. Although re-calculating r produced a stable output, the resulting performance scores of ShoreFor were not as high as the other cross-shore sub-models. In particular, ShoreFor and Y09 can be compared directly at PF4, where the output of IH-MOOSE was driven exclusively by Y09 (its cross-shore sub-model). At this profile, IH-MOOSE (Y09) outperformed ShoreFor in capturing both >6-monthly mean behaviour and shorter-scale oscillations (Fig. 9) and achieved better Mielke's index and Pearson's correlation scores. Additionally, ShorelineEvol (with MD04 as its cross-shore sub-model) and IH-MOOSE (Y09) tended to outperform ShoreFor in capturing <6-monthly oscillations across most transects, despite having alongshore-fixed parameters, unlike ShoreFor. Better performance is expected from ShoreFor using approaches such as re-calculating r over moving 5-year windows (D'Anna et al., 2022) or explicitly relating all of ShoreFor's free parameters to wave climate covariates and allowing them to vary accordingly at each timestep (Ibaceta et al., 2022).

However, more generally, ShoreFor appeared to have difficulty concurrently capturing different timescales of shoreline change over the multi-decadal period, with the moving mean decomposition suggesting that it attempted to focus on either short-term oscillations or longer-term variability at different transects but was not able to capture both well at the same transect (Fig. 9). This is in line with previous studies which have shown that ShoreFor can have difficulty with different timescales when shoreline behaviour is forced by different drivers (e.g., storm events versus ENSO variability) (Montaño et al., 2021; Schepper et al., 2021; Tran et al., 2021), or where interannual-scale shifts in wave climate cause shifts in the dominant timescale of shoreline response (Splinter et al., 2017; Ibaceta et al., 2022). Therefore, performance may also be improved by incorporating additional terms into the model structure to represent the influence of short-scale processes on longer-term shoreline behaviour and vice versa, directly capturing multiple interacting timescales of variability in forcing conditions and shoreline change (Schepper et al., 2021).

Conversely, Y09 appeared to be able to reasonably reproduce both short-term oscillations and >6-monthly behaviour over the multi-decadal period at PF4, where it was the sole contributor to simulated shoreline change by IH-MOOSE. This is somewhat contrary to the results of D'Anna et al. (2022), who found that Y09 simulated little interannual variability compared to ShoreFor. The formulation of the equilibrium condition in Y09 links shoreline positions to wave energy through two free parameters over the calibration period, meaning the model cannot account for wave climate variability on longer timescales than those present during calibration (D'Anna et al., 2021a, 2022; Vitousek et al., 2021). In this case, it appears that providing a multi-decadal period for calibration does allow the model to capture interannual-scale variability over subsequent multi-decadal simulation periods. Finally, the performance of MD04 as a cross-shore sub-model is more difficult to evaluate as it was not calibrated as a standalone model to any transect. However, it appeared to perform similarly to Y09 at Narrabeen, with similar scores of ShorelineEvol (MD04) and IH-MOOSE (Y09) at capturing <6-monthly oscillations across the five transects.

4.4. Time-varying parameters

The results of this study support recent recommendations to move towards new approaches of capturing long-term variability in wave climate and shoreline behaviour for multi-decadal simulations (Splinter and Coco, 2021), particularly the use of time-varying parameters (Ibaceta et al., 2020, 2022; D'Anna et al., 2022). Time-varying parameters are increasingly being used in hydrological models to account for large shifts in simulated systems driven by climate change or land use change, to reduce model sensitivity to the calibration period used, and to overcome the inability of models to adequately simulate all magnitudes of streamflow (i.e., low flows to flood pulses) (e.g., Pathiraja et al., 2016, 2018; Zeng et al., 2019; Zhang and Liu, 2021). Zeng et al. (2019) noted, "Although the time-invariance of model parameters is one of the basic criteria of a high-quality hydrologic model, very few (if any) models can achieve this due to their inherent limitations." It is increasingly becoming apparent that this is also true of shoreline change models, particularly in the presence of interannual wave climate variability.

Interannual wave climate variability has been linked to variations in the relative contributions of cross-shore and longshore processes (Ibaceta et al., 2023), the dominant timescale of shoreline variability (Splinter et al., 2017), the envelope of cross-shore shoreline excursions (McLean et al., 2023), and the equilibrium planform of embayed beaches (Mortlock and Goodwin, 2016). Over the historical periods considered, this variability has been driven by teleconnections such as ENSO, but further wave climate change is projected this century for many parts of the world (Morim et al., 2019). Although wave climate variability should be accounted for in models' wave forcing terms, it appears the models are not fully able to capture some aspects of the resulting variability in shoreline behaviour. The EOF analysis (Fig. 10) indicated that the alongshore magnitude of cross-shore and longshore processes, as well as their relative dominance, was notably different between the 20-year calibration and 20-year validation periods used here. While ShoreFor appeared to be most sensitive to this shift (requiring at least one time-varying parameter to produce a stable output), the EOF analysis indicated that the performance of all models was affected. This included CoSMoS-COAST, the model calibrated using data assimilation, although to a lesser degree as its parameters varied somewhat over the calibration period. The modelled alongshore magnitudes of cross-shore and longshore processes and their relative dominance appeared to be established in models during the calibration period and then maintained into the validation period, implying that this aspect of the wave climate–shoreline behaviour connection is being captured by free parameters rather than explicitly simulated by the models.

Similarly, it is well established that optimal parameter values depend on the specific calibration period selected and/or its duration, with models being tuned to the characteristics of wave climate over this period (e.g., Splinter et al., 2013, 2017; D'Anna et al., 2020; Ibaceta et al., 2020; Tran and Barthélemy, 2020). Splinter et al. (2014) explicitly showed that the optimal parameters of ShoreFor are strongly correlated to wave climate variables, particularly the interannual (~5 year) mean antecedent dimensionless fall velocity. This finding was in the context of parameters varying alongshore and between different sites, but a similar result was subsequently found by Ibaceta et al. (2022) while testing how the optimal parameters of ShoreFor varied through time at a single location. While alongshore-varying parameters may be needed to compensate for variations in profile or sediment characteristics, or varying errors in nearshore wave conditions, the link to interannual wave climate in optimal parameter choices implies that key processes which translate wave forcing conditions to shoreline behaviour are being missed by reduced-complexity models at these longer timescales. The results in the present study suggest that exploring time-varying parameters not just for ShoreFor but also MD04, Y09 and the longshore process sub-models may be a promising avenue for future work. The need for time-varying parameters may be used to indicate specific shortcomings in model structure and identify potential improvements

(Smith et al., 2008); in this case, seeking to capture more of the processes that drive interannual-scale variability in modes of shoreline behaviour.

4.5. Limitations

This study tested models at only one site, Narrabeen-Collaroy Beach in southeast Australia, which represents only one certain set of geomorphological and hydrodynamic characteristics. Shoreline behaviour is controlled by complex cross-shore and longshore processes that vary both spatially (Harley et al., 2011a) as result of geological controls by large headlands and offshore reefs (Daly et al., 2014; Robinet et al., 2020), and temporally (Ibaceta et al., 2023), driven by interannual wave climate variability (Harley et al., 2010; Mortlock and Goodwin, 2016; Vos et al., 2023). These complex dynamics provided a rigorous test for the cross-shore and longshore sub-models of the hybrid framework models, and the evaluation of model performance focused on these aspects. Other aspects such as the sediment demand of estuaries in response to sea-level rise or the effect of hard engineering structures were not set up (in ShorelineEvol), and foredune erosion (in COCOONED) appeared to contribute only minimally. Narrabeen is micro-tidal and storm surge does not play a big role in driving shoreline change due to the narrow continental shelf (Ibaceta and Harley, 2024), so the site also did not rigorously test the role of tides and storm surge as model forcing variables (in MD04 in COCOONED and ShorelineEvol). Nonetheless, this study builds on previous work (e.g., Montaño et al., 2020) in testing multiple models at a site where this was not previously undertaken, and in particular, over a multi-decadal validation period.

The ability of the models to simulate shoreline response to sea-level rise was also not rigorously tested, although the sea-level rise sub-model was turned on in models that had this capability (CoSMoS-COAST, COCOONED and ShorelineEvol). Over the hindcast period studied here, sea-level rise has not appeared to produce a measurable response in the measured shoreline behaviour at Narrabeen, consistent with other sites with long-term monitoring programs on the southeast Australian coastline (McLean et al., 2023). Testing sea-level rise response models such as the Bruun Rule (Bruun, 1962) has not been possible to date except in laboratory studies (Atkinson et al., 2018; Beuzen et al., 2018) or on rapidly subsiding coastlines (e.g., List et al., 1997; Leatherman et al., 2000; Zhang et al., 2004), and sea-level rise is generally expected to emerge as a dominant driver of shoreline behaviour for most coastlines after the middle of the century (Dean and Houston, 2016; Vitousek et al., 2017b; Le Cozannet et al., 2019; Thiéblemont et al., 2021; D'Anna et al., 2022). However, some 'early responder' sites may provide an opportunity to test models over the coming years (Sharples et al., 2020).

It is emphasised that this contribution does not necessarily reflect the best possible performance that each of the five models could theoretically achieve. Although automated approaches were used to calibrate model parameters, each model required manual set-up and/or site-specific modifications, and several additional modifications were possible but not implemented (suggested in the discussion above). The relative rankings of the models therefore do not represent an absolute measure of their performance. Better performance scores and/or different rankings may be achieved with additional tuning at this site, and different rankings and performance scores are likely to be achieved at other sites with different morphodynamic characteristics. Rather, this study identifies potential limitations and improvements to reduced-complexity or semi-empirical shoreline modelling approaches more generally, and highlights the value of testing models at a range of morphodynamically different sites beyond those where they were first developed. Finally, although other authors have shown that ensembles of models typically perform better than any individual model (e.g., Montaño et al., 2020; Simmons and Splinter, 2022), this study highlights that addressing site-specific complexities in each model can make implementing an ensemble challenging in practice, particularly for multiple sites at regional scales.

5. Conclusions

Five shoreline change models (ShoreFor, CoSMoS-COAST, COCOONED, ShorelineEvol and IH-MOOSE) were run over a 20-year calibration and 20-year validation period at Narrabeen-Collaroy Beach, Australia. Four models were hybrid frameworks coupling sub-models of cross-shore and longshore processes, and one (ShoreFor) was a standalone model of wave-driven cross-shore change. Initially, three models that used the one-line (CERC) sub-model to simulate longshore sediment transport showed large-magnitude planform shoreline realignment in response to biased nearshore wave data (Chataigner et al., 2022). This issue dominated model behaviour regardless of the other sub-models in the hybrid framework, or whether optimisation or data assimilation was used to calibrate the models. These three models were re-initialised using a 'virtual equilibrium shoreline', where the modelled shoreline was allowed to realign until equilibrium was reached before re-starting simulations from this equilibrium position. This substantially improved performance and the longshore transport sub-models were subsequently able to capture seasonal-scale rotations of the embayment. Nonetheless, this solution was imperfect and the best way to correct for this issue, which is not limited to this site, remains an avenue for future work.

The standalone cross-shore model ShoreFor showed reasonable performance over the calibration period but simulated large accretionary trends over the validation period that diverged from measured positions. Similar behaviour was recently observed by Ibaceta et al. (2022) and attributed to the sensitivity of the model to wave climate non-stationarity. In the present study, this was resolved with a first-pass solution of re-calculating the model's 'erosion ratio' term r, derived from wave forcing conditions, independently for the validation period.

With these modifications applied, the performance of all five models over the validation period was similar on average but with substantial variability between transects and across performance metrics, as different models were limited by different factors. These included: shortcomings in the virtual equilibrium shoreline approach implemented here (for the models CoSMoS-COAST, COCOONED and ShorelineEvol using the one-line (CERC) sub-model), difficulty capturing multiple timescales of shoreline variability concurrently over a long simulation period (ShoreFor), and the parameterisation of the embayment planform (IH-MOOSE). These results suggest that no model or approach was clearly better or worse than the others; rather, accurately simulating shoreline change over multiple decades at a morphodynamically complex site poses an ongoing challenge.

Overall, coupling sub-models of longshore processes (longshore sediment transport or beach rotation) to cross-shore sub-models did appear to improve performance, despite the additional complexity and potential for error described above. The cross-shore sub-models of the hybrid framework also appeared to 'let down' the performance of hybrid frameworks more than the longshore sub-models. This may have been improved with alongshore-varying parameter values for the cross-shore sub-models.

Additionally, while the measured shoreline behaviour showed a shift in the spatial variability and relative dominance of cross-shore and longshore processes over the two 20-year periods, modelled shoreline dynamics appeared to be established during the calibration period and remained fixed into the validation period unless parameters were allowed to vary over time in some way. These results indicate that the current model structures are not capturing all aspects of the connection between interannual wave climate variability and shoreline behaviour, with key processes being aggregated into model free parameters rather than being directly accounted for by the models' wave forcing terms. Consequently, time-varying parameters or changes to model structure are suggested as avenues for further research to improve the performance of reduced-complexity models in multi-decadal simulations.

CRediT authorship contribution statement

Oxana Repina: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Rafael C. Carvalho:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Giovanni Coco:** Writing – review & editing, Supervision, Methodology, Conceptualization. **José A.Á. Antolínez:** Writing – review & editing, Software, Methodology, Investigation. **Iñaki de Santiago:** Writing – review & editing, Software, Methodology, Investigation. **Mitchell D. Harley:** Writing – review & editing, Methodology. **Camilo Jaramillo:** Writing – review & editing, Software, Methodology, Investigation. **Kristen D. Splinter:** Writing – review & editing, Methodology. **Sean Vitousek:** Writing – review & editing, Software, Methodology, Investigation. **Colin D. Woodroffe:** Writing – review & editing, Supervision, Resources, Conceptualization.

Data and software availability

Narrabeen Beach profile data are publicly available at http://narrabeen.wrl.unsw.edu.au. ANCHORS water level data and CAWCR wave hindcast data are publicly available at https://dx.doi.org/10.25914/6142dff37250b and http://hdl.handle.net/102.100.100/137152?index=1 respectively. The most recent versions of the five models used in this study are available as follows. CoSMoS-COAST: publicly available on the USGS source code archive at https://doi.org/10.5066/P95T9188. COCOONED: available on request (José A. Á. Antolínez, j.a.a.antolinez@tudelft.nl). ShorelineEvol: available on request (Iñaki de Santiago, idsantiago@azti.es). ShoreFor: publicly available on Github at https://github.com/ShoreShop/ShoreModel_Benchmark. IH-MOOSE: available on request (Camilo Jaramillo, camilo.jaramillo@unican.es).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements and funding sources

Nearshore wave data from the NSW State Wide Nearshore Wave Transformation Tool were kindly provided by Bradley Morris on behalf of the NSW Department of Climate Change, Energy, the Environment and Water. CSIRO is acknowledged for providing the CAWCR offshore wave hindcast dataset, and the Bureau of Meteorology and NCI Australia for the ANCHORS tide gauge data. Narrabeen-Collaroy Beach survey data are collected by the UNSW Water Research Laboratory and previously by the University of Sydney Coastal Studies Unit (Professor Andrew Short). Surveys have been supported by past funding from the ARC DP and LP programs and Warringah/Northern Beaches Council. The R implementation of the model ShoreFor used here was kindly provided by Yen Hai Tran and Eric Barthélemy. We thank two anonymous journal reviewers and Erdinc Sogut, who provided a USGS internal review, for their thoughtful feedback on this manuscript.

Oxana Repina acknowledges the support of an Australian Government Research Training Program Scholarship, as well as the resources and services provided by the University of Wollongong partner share of the NCI, which is supported by the Australian Government. Iñaki de Santiago acknowledges the support of the KOSTARISK joint laboratory (funded by AZTI, UPPA and RPT), Regions4Climate project (Horizon Europe, DOI: 10.3030/101093873) and the Urban Klima 2050–LIFE 18 IPC 000001 project (European Union's LIFE program). Camilo Jaramillo (C.J.) acknowledges the support of the ThinkInAzul programme, supported by MCIN/Ministerio de Ciencia e Innovación with funding from the European Union NextGeneration EU (PRTR-C17.I1) and by Comunidad de Cantabria. C.J. has also been supported by a Margarita Salas post-doctoral fellowship, funded by the European Union-NextGenerationEU, Ministry of Universities, and Recovery and Resilience Facility, through a call from the University of Cantabria. Kristen D. Splinter acknowledges support from an ARC Future Fellowship (FT220100009). This work is a contribution to the IGCP Project 725 'Forecasting Coastal Change'. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.coastaleng.2025.104738.

References

Aber, J.D., 1997. Why don't we believe the models? Bull. Ecol. Soc. Am. 78 (3), 232–233.
Alvarez-Cuesta, M., Toimil, A., Losada, I.J., 2021. Modelling long-term shoreline evolution in highly anthropized coastal areas. Part 1: model description and validation. Coast. Eng. 169, 103960. https://doi.org/10.1016/j.coastaleng.2021.103960.
Anderson, D., Ruggiero, P., Antolínez, J.A.A., Méndez, F.J., Allan, J., 2018. A climate index optimized for longshore sediment transport reveals interannual and multidecadal littoral cell rotations. J. Geophys. Res.: Earth Surf. 123 (8), 1958–1981. https://doi.org/10.1029/2018JF004689.
Anthony, E.J., Aagaard, T., 2020. The lower shoreface: morphodynamics and sediment connectivity with the upper shoreface and beach. Earth Sci. Rev. 210, 103334. https://doi.org/10.1016/j.earscirev.2020.103334.
Antolínez, J.A.A., Méndez, F.J., Anderson, D., Ruggiero, P., Kaminsky, G.M., 2019. Predicting climate-driven coastlines with a simple and efficient multiscale model. J. Geophys. Res.: Earth Surf. 124 (6), 1596–1624. https://doi.org/10.1029/2018JF004790.
Arsenault, R., Poulin, A., Côté, P., Brissette, F., 2014. Comparison of stochastic optimization algorithms in hydrological model calibration. J. Hydrol. Eng. 19 (7), 1374–1384.
Ashton, A.D., Murray, A.B., 2006. High-angle wave instability and emergent shoreline shapes: 1. Modeling of sand waves, flying spits, and capes. J. Geophys. Res.: Earth Surf. 111 (F4). https://doi.org/10.1029/2005JF000422.
Ashton, A.D., Murray, A.B., Arnoult, O., 2001. Formation of coastline features by large-scale instabilities induced by high-angle waves. Nature 414 (6861), 296–300. https://doi.org/10.1038/35104541.
Atkinson, A.L., Baldock, T.E., Birrien, F., Callaghan, D.P., Nielsen, P., Beuzen, T., Turner, I.L., Blenkinsopp, C.E., Ranasinghe, R., 2018. Laboratory investigation of the Bruun Rule and beach response to sea level rise. Coast. Eng. 136, 183–202. https://doi.org/10.1016/j.coastaleng.2018.03.003.
Baird Australia, 2017. NSW Coastal Wave Model: State Wide Nearshore Wave Transformation Tool, Prepared for the Office of Environment and Heritage (NSW). Sydney, NSW, Australia.
Bamunawala, J., Dastgheib, A., Ranasinghe, R., van der Spek, A., Maskey, S., Murray, A. B., Barnard, P.L., Duong, T.M., Sirisena, T.A.J.G., 2020. Probabilistic application of an integrated catchment-estuary-coastal system model to assess the evolution of inlet-interrupted coasts over the 21st century. Front. Mar. Sci. 7, 1104. https://doi.org/10.3389/fmars.2020.579203.
Bayram, A., Larson, M., Hanson, H., 2007. A new formula for the total longshore sediment transport rate. Coast. Eng. 54 (9), 700–710. https://doi.org/10.1016/j.coastaleng.2007.04.001.
Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Software 40, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011.
Beuzen, T., 2019. pybeach: a Python package for extracting the location of dune toes on beach profile transects. J. Open Source Softw. 4 (44), 1890.
Beuzen, T., Turner, I.L., Blenkinsopp, C.E., Atkinson, A., Flocard, F., Baldock, T.E., 2018. Physical model study of beach profile evolution by sea level rise in the presence of seawalls. Coast. Eng. 136, 172–182. https://doi.org/10.1016/j.coastaleng.2017.12.002.
Bishop-Taylor, R., Nanson, R., Sagar, S., Lymburner, L., 2021. Mapping Australia's dynamic coastline at mean sea level using three decades of Landsat imagery. Rem. Sens. Environ. 267, 112734. https://doi.org/10.1016/j.rse.2021.112734.
Brooks, S.M., Spencer, T., Christie, E.K., 2017. Storm impacts and shoreline recovery: mechanisms and controls in the southern North Sea. Geomorphology 283, 48–60. https://doi.org/10.1016/j.geomorph.2017.01.007.
Bruun, P., 1962. Sea-level rise as a cause of shore erosion. J. Waterw. Harb. Div. 88 (1), 117–132.
Carley, J., Lewis, G., Couriel, E., 2016. Collaroy-Narrabeen Beach Coastal Protection Assessment (No. MHL2491). Manly Hydraulics Laboratory. Prepared for Northern Beaches Council.
Chataigner, T., Yates, M.L., Le Dantec, N., Harley, M.D., Splinter, K.D., Goutal, N., 2022. Sensitivity of a one-line longshore shoreline change model to the mean wave

direction. Coast. Eng. 172, 104025. https://doi.org/10.1016/j.coastaleng.2021.104025.
Coco, G., Senechal, N., Rejas, A., Bryan, K.R., 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.
Codiga, D.L., 2011. Unified Tidal Analysis and Prediction Using the UTide Matlab Functions (Technical Report No. 2011– 01). Graduate School of Oceanography, University of Rhode Island, Narragansett, RI.
Cowell, P.J., Kinsela, M.A., 2018. Shoreface controls on barrier evolution and shoreline change. In: Moore, L.J., Murray, A.B. (Eds.), Barrier Dynamics and Response to Changing Climate. Springer, pp. 243–275.
Cowell, P.J., Roy, P.S., Jones, R.A., 1995. Simulation of large-scale coastal change using a morphological behaviour model. Mar. Geol. 126, 45–61.
Cowell, P.J., Stive, M.J., Niedoroda, A.W., de Vriend, H.J., Swift, D.J., Kaminsky, G.M., Capobianco, M., 2003. The coastal-tract (part 1): a conceptual approach to aggregated modeling of low-order coastal change. J. Coast Res. 19 (4), 812–827.
Cowell, P.J., Thom, B.G., 1994. Morphodynamics of coastal evolution. In: Carter, R.W.G., Woodroffe, C.D. (Eds.), Coastal Evolution: Late Quaternary Shoreline Morphodynamics. Cambridge University Press, Cambridge, UK, pp. 33–86.
Daly, C.J., Bryan, K.R., Winter, C., 2014. Wave energy distribution and morphological development in and around the shadow zone of an embayed beach. Coast. Eng. 93, 40–54. https://doi.org/10.1016/j.coastaleng.2014.08.003.
D'Anna, M., Castelle, B., Idier, D., Rohmer, J., Le Cozannet, G., Thieblemont, R., Bricheno, L., 2021a. Uncertainties in shoreline projections to 2100 at Truc Vert Beach (France): role of sea-level rise and equilibrium model assumptions. J. Geophys. Res.: Earth Surf. 126 (8), e2021JF006160. https://doi.org/10.1029/2021JF006160.
D'Anna, M., Idier, D., Castelle, B., Le Cozannet, G., Rohmer, J., Robinet, A., 2020. Impact of model free parameters and sea-level rise uncertainties on 20-years shoreline hindcast: the case of Truc Vert beach (SW France). Earth Surf. Process. Landf. 45 (8), 1895–1907. https://doi.org/10.1002/esp.4854.
D'Anna, M., Idier, D., Castelle, B., Rohmer, J., Cagigal, L., Mendez, F.J., 2022. Effects of stochastic wave forcing on probabilistic equilibrium shoreline response across the 21st century including sea-level rise. Coast. Eng. 175, 104149. https://doi.org/10.1016/j.coastaleng.2022.104149.
D'Anna, M., Idier, D., Castelle, B., Vitousek, S., Le Cozannet, G., 2021b. Reinterpreting the Bruun Rule in the context of equilibrium shoreline models. J. Mar. Sci. Eng. 9 (9). https://doi.org/10.3390/jmse9090974.
Davidson, M.A., Splinter, K.D., Turner, I.L., 2013. A simple equilibrium model for predicting shoreline change. Coast. Eng. 73, 191–202. https://doi.org/10.1016/j.coastaleng.2012.11.002.
de Santiago, I., Camus, P., González, M., Liria, P., Epelde, I., Chust, G., del Campo, A., Uriarte, A., 2021. Impact of climate change on beach erosion in the Basque Coast (NE Spain). Coast. Eng. 167, 103916. https://doi.org/10.1016/j.coastaleng.2021.103916.
Dean, R.G., Houston, J.R., 2016. Determining shoreline response to sea level rise. Coast. Eng. 114, 1–8. https://doi.org/10.1016/j.coastaleng.2016.03.009.
Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28 (4), 1015–1031. https://doi.org/10.1029/91WR02985.
Duan, Q., Sorooshian, S., Gupta, V.K., 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. J. Hydrol. 158 (3), 265–284. https://doi.org/10.1016/0022-1694(94)90057-4.
Durrant, T., Greenslade, D., Hemer, M., Trenham, C., 2014. A Global Wave Hindcast Focussed on the Central and South Pacific. CAWCR Technical Report No. 070.
Duveiller, G., Fasbender, D., Meroni, M., 2016. Revisiting the concept of a symmetric index of agreement for continuous datasets. Sci. Rep. 6 (1), 19401. https://doi.org/10.1038/srep19401.
Emery, K.O., 1961. A simple method of measuring beach profiles. Limnol. Oceanogr. 6 (1), 90–93.
Gomez-de la Peña, E., Coco, G., Whittaker, C., Montaño, J., 2023. On the use of convolutional deep learning to predict shoreline change. Earth Surf. Dyn. 11 (6), 1145–1160. https://doi.org/10.5194/esurf-11-1145-2023.
González, M., Medina, R., 2001. On the application of static equilibrium bay formulations to natural and man-made beaches. Coast. Eng. 43 (3), 209–225. https://doi.org/10.1016/S0378-3839(01)00014-X.
Goodwin, I.D., Freeman, R., Blackmore, K., 2013. An insight into headland sand bypassing and wave climate variability from shoreface bathymetric change at Byron Bay, New South Wales, Australia. Mar. Geol. 341, 29–45. https://doi.org/10.1016/j.margeo.2013.05.005.
Hague, B.S., Jones, D.A., Trewin, B., Jakob, D., Murphy, B.F., Martin, D.J., Braganza, K., 2021. ANCHORS: a multi-decadal tide gauge dataset to monitor Australian relative sea level changes. Geoscience Data Journal 9 (2), 256–272. https://doi.org/10.1002/gdj3.136.
Hague, B.S., Taylor, A.J., 2021. Tide-only inundation: a metric to quantify the contribution of tides to coastal inundation under sea-level rise. Nat. Hazards 107 (1), 675–695. https://doi.org/10.1007/s11069-021-04600-4.
Hallermeier, R.J., 1981. A profile zonation for seasonal sand beaches from wave climate. Coast. Eng. 4, 253–277. https://doi.org/10.1016/0378-3839(80)90022-8.
Hallin, C., Huisman, B.J.A., Larson, M., Walstra, D.-J.R., Hanson, H., 2019. Impact of sediment supply on decadal-scale dune evolution — analysis and modelling of the Kennemer dunes in The Netherlands. Geomorphology 337, 94–110. https://doi.org/10.1016/j.geomorph.2019.04.003.
Hamon-Kerivel, K., Cooper, A., Jackson, D., Sedrati, M., Guisado Pintado, E., 2020. Shoreface mesoscale morphodynamics: a review. Earth Sci. Rev. 209, 103330. https://doi.org/10.1016/j.earscirev.2020.103330.
Hanson, H., 1989. Genesis: a generalized shoreline change numerical model. J. Coast Res. 5 (1), 1–27.
Hanson, H., Kraus, N.C., 1989. GENESIS: Generalized Model for Simulating Shoreline Change (Technical Reference No. 1). US Army Corps of Engineers Waterways Experiment Station. Coastal Engineering Research Centre, Vicksburg, Mississippi.
Harley, M.D., Masselink, G., Ruiz de Alegría-Arzaburu, A., Valiente, N.G., Scott, T., 2022. Single extreme storm sequence can offset decades of shoreline retreat projected to result from sea-level rise. Communications Earth & Environment 3 (1), 112. https://doi.org/10.1038/s43247-022-00437-2.
Harley, M.D., Turner, I.L., 2008. A simple data transformation technique for pre-processing survey data at embayed beaches. Coast. Eng. 55 (1), 63–68. https://doi.org/10.1016/j.coastaleng.2007.07.001.
Harley, M.D., Turner, I.L., Kinsela, M.A., Middleton, J.H., Mumford, P.J., Splinter, K.D., Phillips, M.S., Simmons, J.A., Hanslow, D.J., Short, A.D., 2017. Extreme coastal erosion enhanced by anomalous extratropical storm wave direction. Sci. Rep. 7 (1), 6033. https://doi.org/10.1038/s41598-017-05792-1.
Harley, M.D., Turner, I.L., Short, A.D., 2015. New insights into embayed beach rotation: the importance of wave exposure and cross-shore processes. J. Geophys. Res.: Earth Surf. 120 (8), 1470–1484. https://doi.org/10.1002/2014JF003390.
Harley, M.D., Turner, I.L., Short, A.D., Ranasinghe, R., 2009. An empirical model of beach response to storms–SE Australia. In: Proceedings of the 19th Australasian Coastal and Ocean Engineering Conference, pp. 600–606. Wellington, New Zealand.
Harley, M.D., Turner, I.L., Short, A.D., Ranasinghe, R., 2010. Interannual variability and controls of the Sydney wave climate. Int. J. Climatol. 30 (9), 1322–1335. https://doi.org/10.1002/joc.1962.
Harley, M.D., Turner, I.L., Short, A.D., Ranasinghe, R., 2011a. A reevaluation of coastal embayment rotation: the dominance of cross-shore versus alongshore sediment transport processes, Collaroy-Narrabeen Beach, southeast Australia. J. Geophys. Res.: Earth Surf. 116 (F4). https://doi.org/10.1029/2011JF001989.
Harley, M.D., Turner, I.L., Short, A.D., Ranasinghe, R., 2011b. Assessment and integration of conventional, RTK-GPS and image-derived beach survey methods for daily to decadal coastal monitoring. Coast. Eng. 58 (2), 194–205. https://doi.org/10.1016/j.coastaleng.2010.09.006.
Harley, M.D., Turner, I.L., Splinter, K.D., Phillips, M.S., Simmons, J.A., 2016. Beach response to Australian east coast lows: a comparison between the 2007 and 2015 events, narrabeen-collaroy beach. J. Coast Res. 75 (10075), 388–392. https://doi.org/10.2112/SI75-078.1.
Hsu, J.R.C., Evans, C., 1989. Parabolic bay shapes and applications. Proc. Inst. Civ. Eng. 87 (4), 557–570. https://doi.org/10.1680/iicep.1989.3778.
Hunt, E., Davidson, M., Steele, E.C.C., Amies, J.D., Scott, T., Russell, P., 2023. Shoreline modelling on timescales of days to decades. Cambridge Prisms: Coastal Futures 1–26. https://doi.org/10.1017/cft.2023.5.
Huxley, C., 2009. Shoreline response modelling assessing the impacts of climate change. In: Proceedings of the 18th NSW Coastal Conference. Citeseer, Ballina, NSW, Australia.
Ibaceta, R., Harley, M.D., 2024. Data-driven modelling of coastal storm erosion for real-time forecasting at a wave-dominated embayed beach. Coast. Eng. 193, 104596. https://doi.org/10.1016/j.coastaleng.2024.104596.
Ibaceta, R., Harley, M.D., Turner, I.L., Splinter, K.D., 2023. Interannual variability in dominant shoreline behaviour at an embayed beach. Geomorphology 433, 108706. https://doi.org/10.1016/j.geomorph.2023.108706.
Ibaceta, R., Splinter, K.D., Harley, M.D., Turner, I.L., 2020. Enhanced coastal shoreline modeling using an ensemble Kalman filter to include nonstationarity in future wave climates. Geophys. Res. Lett. 47 (22), e2020GL090724. https://doi.org/10.1029/2020GL090724.
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.
Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. Environ. Model. Software 21 (5), 602–614. https://doi.org/10.1016/j.envsoft.2006.01.004.
Jaramillo, C., González, M., Medina, R., Turki, I., 2021a. An equilibrium-based shoreline rotation model. Coast. Eng. 163, 103789. https://doi.org/10.1016/j.coastaleng.2020.103789.
Jaramillo, C., Jara, M.S., González, M., Medina, R., 2020. A shoreline evolution model considering the temporal variability of the beach profile sediment volume (sediment gain/loss). Coast. Eng. 156, 103612. https://doi.org/10.1016/j.coastaleng.2019.103612.
Jaramillo, C., Jara, M.S., González, M., Medina, R., 2021b. A shoreline evolution model for embayed beaches based on cross-shore, planform and rotation equilibrium models. Coast. Eng. 169, 103983. https://doi.org/10.1016/j.coastaleng.2021.103983.
Kamphuis, J.W., 1991. Alongshore sediment transport rate. J. Waterw. Port, Coast. Ocean Eng. 117 (6), 624–640. https://doi.org/10.1061/(ASCE)0733-950X(1991)117:6(624.
Komar, P.D., 1971. The mechanics of sand transport on beaches. J. Geophys. Res. 76 (3), 713–721. https://doi.org/10.1029/JC076i003p00713.
Kriebel, D.L., Dean, R.G., 1993. Convolution method for time-dependent beach-profile response. J. Waterw. Port, Coast. Ocean Eng. 119 (2), 204–226.
Le Cozannet, G., Bulteau, T., Castelle, B., Ranasinghe, R., Wöppelmann, G., Rohmer, J., Bernon, N., Idier, D., Louisor, J., Salas-y-Mélia, D., 2019. Quantifying uncertainties of sandy shoreline change projections as sea level rises. Sci. Rep. 9 (42).
Leatherman, S.P., Zhang, K., Douglas, B.C., 2000. Sea level rise shown to drive coastal erosion. Eos, Transactions American Geophysical Union 81 (6), 55–57. https://doi.org/10.1029/00EO00034.

Liemohn, M.W., Shane, A.D., Azari, A.R., Petersen, A.K., Swiger, B.M., Mukhopadhyay, A., 2021. RMSE is not enough: guidelines to robust data-model comparisons for magnetospheric physics. J. Atmos. Sol. Terr. Phys. 218, 105624. https://doi.org/10.1016/j.jastp.2021.105624.

List, J.H., Sallenger, A.H., Hansen, M.E., Jaffe, B.E., 1997. Accelerated relative sea-level rise and rapid coastal erosion: testing a causal relationship for the Louisiana barrier islands. Mar. Geol. 140 (3), 347–365. https://doi.org/10.1016/S0025-3227(97)00035-2.

Luijendijk, A., Hagenaars, G., Ranasinghe, R., Baart, F., Donchyts, G., Aarninkhof, S., 2018. The state of the world's beaches. Sci. Rep. 8 (1), 6641. https://doi.org/10.1038/s41598-018-24630-6.

Matott, L.S., Hymiak, B., Reslink, C., Baxter, C., Aziz, S., 2013. Telescoping strategies for improved parameter estimation of environmental simulation models. Comput. Geosci. 60, 156–167. https://doi.org/10.1016/j.cageo.2013.07.023.

McLean, R., Thom, B., Shen, J., Oliver, T., 2023. 50 years of beach–foredune change on the southeastern coast of Australia: bengello Beach, Moruya, NSW, 1972–2022. Geomorphology 439, 108850. https://doi.org/10.1016/j.geomorph.2023.108850.

Mil-Homens, J., Ranasinghe, R., van Thiel de Vries, J.S.M., Stive, M.J.F., 2013. Re-evaluation and improvement of three commonly used bulk longshore sediment transport formulas. Coast. Eng. 75, 29–39. https://doi.org/10.1016/j.coastaleng.2013.01.004.

Miller, J.K., Dean, R.G., 2004. A simple new shoreline change model. Coast. Eng. 51 (7), 531–556. https://doi.org/10.1016/j.coastaleng.2004.05.006.

Montaño, J., Coco, G., Antolínez, J.A.A., Beuzen, T., Bryan, K.R., Cagigal, L., Castelle, B., Davidson, M.A., Goldstein, E.B., Ibaceta, R., Idier, D., Ludka, B.C., Masoud-Ansari, S., Méndez, F.J., Murray, A.B., Plant, N.G., Ratliff, K.M., Robinet, A., Rueda, A., Sénéchal, N., Simmons, J.A., Splinter, K.D., Stephens, S., Townend, I., Vitousek, S., Vos, K., 2020. Blind testing of shoreline evolution models. Sci. Rep. 10 (1), 2137. https://doi.org/10.1038/s41598-020-59018-y.

Montaño, J., Coco, G., Cagigal, L., Mendez, F., Rueda, A., Bryan, K.R., Harley, M.D., 2021. A multiscale approach to shoreline prediction. Geophys. Res. Lett. 48 (1), e2020GL090587. https://doi.org/10.1029/2020GL090587.

Morim, J., Hemer, M., Wang, X.L., Cartwright, N., Trenham, C., Semedo, A., Young, I., Bricheno, L., Camus, P., Casas-Prat, M., Erikson, L., Mentaschi, L., Mori, N., Shimura, T., Timmermans, B., Aarnes, O., Breivik, Ø., Behrens, A., Dobrynin, M., Menendez, M., Staneva, J., Wehner, M., Wolf, J., Kamranzad, B., Webb, A., Stopa, J., Andutta, F., 2019. Robustness and uncertainties in global multivariate wind-wave climate projections. Nat. Clim. Change 9 (9), 711–718. https://doi.org/10.1038/s41558-019-0542-5.

Mortlock, T.R., Goodwin, I.D., 2016. Impacts of enhanced central Pacific ENSO on wave climate and headland-bay beach morphology. Cont. Shelf Res. 120, 14–25. https://doi.org/10.1016/j.csr.2016.03.007.

Mortlock, T.R., Goodwin, I.D., McAneney, J.K., Roche, K., 2017. The June 2016 Australian East Coast Low: importance of wave direction for coastal erosion assessment. Water 9 (2). https://doi.org/10.3390/w9020121.

Mull, J., Ruggiero, P., 2014. Estimating storm-induced dune erosion and overtopping along U.S. West Coast beaches. J. Coast Res. 30 (6), 1173–1187. https://doi.org/10.2112/JCOASTRES-D-13-00178.1.

New South Wales Department of Climate Change, Energy, the Environment and Water, 2019. NSW Marine LiDAR Topo-Bathy 2018 Geotif.

Palalane, J., Larson, M., 2020. A long-term coastal evolution model with longshore and cross-shore transport. J. Coast Res. 36 (2), 411–423. https://doi.org/10.2112/JCOASTRES-D-17-00020.1.

Pathiraja, S., Anghileri, D., Burlando, P., Sharma, A., Marshall, L., Moradkhani, H., 2018. Time-varying parameter models for catchments with land use change: the importance of model structure. Hydrol. Earth Syst. Sci. 22 (5), 2903–2919. https://doi.org/10.5194/hess-22-2903-2018.

Pathiraja, S., Marshall, L., Sharma, A., Moradkhani, H., 2016. Hydrologic modeling in dynamic catchments: a data assimilation approach. Water Resour. Res. 52 (5), 3350–3372. https://doi.org/10.1002/2015WR017192.

Pelnard-Considère, R., 1956. Essai de théorie de l'évolution des formes de rivage en plages de sable et de galets [Model tests confirm the theory of the changing shapes of sand and stone beaches], in: Compte Rendu Des Quatrièmes Journées de l'hydraulique. In: Presented at the Les énergies de la mer, pp. 289–298. Paris.

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. https://doi.org/10.1016/j.margeo.2017.01.005.

Pilkey, O.H., Cooper, J.A.G., 2002. Longshore transport volumes: a critical view. J. Coast Res. 36 (10036), 572–580. https://doi.org/10.2112/1551-5036-36.sp1.572.

Ranasinghe, R., 2016. Assessing climate change impacts on open sandy coasts: a review. Earth Sci. Rev. 160, 320–332. https://doi.org/10.1016/j.earscirev.2016.07.011.

Ranasinghe, R., 2020. On the need for a new generation of coastal change models for the 21st century. Sci. Rep. 10 (1), 2010. https://doi.org/10.1038/s41598-020-58376-x.

Ranasinghe, R., Duong, T.M., Uhlenbrook, S., Roelvink, D., Stive, M., 2013. Climate-change impact assessment for inlet-interrupted coastlines. Nat. Clim. Change 3 (1), 83.

Ranasinghe, R., McLoughlin, R., Short, A., Symonds, G., 2004. The Southern Oscillation Index, wave climate, and beach rotation. Mar. Geol. 204 (3–4), 273–287.

Reeve, D.E., Horrillo-Caraballo, J., Karunarathna, H., Pan, S., 2019. A new perspective on meso-scale shoreline dynamics through data-driven analysis. Geomorphology 341, 169–191. https://doi.org/10.1016/j.geomorph.2019.04.033.

Reeve, D.E., Karunarathna, H., Pan, S., Horrillo-Caraballo, J.M., Różyński, G., Ranasinghe, R., 2016. Data-driven and hybrid coastal morphological prediction methods for mesoscale forecasting. Geomorphology 256, 49–67. https://doi.org/10.1016/j.geomorph.2015.10.016.

Robinet, A., Castelle, B., Idier, D., Harley, M.D., Splinter, K.D., 2020. Controls of local geology and cross-shore/longshore processes on embayed beach shoreline variability. Mar. Geol. 422, 106118. https://doi.org/10.1016/j.margeo.2020.106118.

Robinet, A., Idier, D., Castelle, B., Marieu, V., 2018. A reduced-complexity shoreline change model combining longshore and cross-shore processes: the LX-Shore model. Environ. Model. Software 109, 1–16. https://doi.org/10.1016/j.envsoft.2018.08.010.

Roelvink, D., Huisman, B., Elghandour, A., Ghonim, M., Reyns, J., 2020. Efficient modeling of complex sandy coastal evolution at monthly to century time scales. Front. Mar. Sci. 7, 535. https://doi.org/10.3389/fmars.2020.00535.

Rosati, J.D., Dean, R.G., Walton, T.L., 2013. The modified Bruun Rule extended for landward transport. Mar. Geol. 340, 71–81.

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 (6). https://doi.org/10.3390/jmse9060582.

Sharples, C., Walford, H., Watson, C., Ellison, J.C., Hua, Q., Bowden, N., Bowman, D., 2020. Ocean Beach, Tasmania: a swell-dominated shoreline reaches climate-induced recessional tipping point? Mar. Geol. 419, 106081. https://doi.org/10.1016/j.margeo.2019.106081.

Short, A., Trenaman, N., 1992. Wave climate of the Sydney region, an energetic and highly variable ocean wave regime. Mar. Freshw. Res. 43 (4), 765–791.

Short, A.D., 1984. Beach and nearshore facies: southeast Australia. Mar. Geol. 60 (1), 261–282. https://doi.org/10.1016/0025-3227(84)90153-1.

Short, A.D., Trembanis, A.C., 2004. Decadal scale patterns in beach oscillation and rotation Narrabeen Beach, Australia—time series, PCA and wavelet analysis. J. Coast Res. 20 (2), 523–532.

Silva, A.P. da, Vieira da Silva, G., Strauss, D., Murray, T., Woortmann, L.G., Taber, J., Cartwright, N., Tomlinson, R., 2021. Headland bypassing timescales: processes and driving forces. Sci. Total Environ. 793, 148591. https://doi.org/10.1016/j.scitotenv.2021.148591.

Simmons, J.A., Splinter, K.D., 2022. A multi-model ensemble approach to coastal storm erosion prediction. Environ. Model. Software 150, 105356. https://doi.org/10.1016/j.envsoft.2022.105356.

Smith, P.J., Beven, K.J., Tawn, J.A., 2008. Detection of structural inadequacy in process-based hydrological models: a particle-filtering approach. Water Resour. Res. 44 (1). https://doi.org/10.1029/2006WR005205.

Splinter, K.D., Coco, G., 2021. Challenges and opportunities in coastal shoreline prediction. Front. Mar. Sci. 8.

Splinter, K.D., Turner, I.L., Davidson, M.A., 2013. How much data is enough? The importance of morphological sampling interval and duration for calibration of empirical shoreline models. Coast. Eng. 77, 14–27. https://doi.org/10.1016/j.coastaleng.2013.02.009.

Splinter, K.D., Turner, I.L., Davidson, M.A., Barnard, P., Castelle, B., Oltman-Shay, J., 2014. A generalized equilibrium model for predicting daily to interannual shoreline response. J. Geophys. Res.: Earth Surf. 119 (9), 1936–1958. https://doi.org/10.1002/2014JF003106.

Splinter, K.D., Turner, I.L., Reinhardt, M., Ruessink, G., 2017. Rapid adjustment of shoreline behavior to changing seasonality of storms: observations and modelling at an open-coast beach. Earth Surf. Process. Landf. 42 (8), 1186–1194. https://doi.org/10.1002/esp.4088.

Stive, M.J.F., Wang, Z.B., 2003. In: Lakhan, V.C. (Ed.), Morphodynamic Modeling of Tidal Basins and Coastal Inlets. Elsevier Oceanography Series. Elsevier, pp. 367–392. https://doi.org/10.1016/S0422-9894(03)80130-7.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106 (D7), 7183–7192. https://doi.org/10.1029/2000JD900719.

Thiéblemont, R., Le Cozannet, G., Rohmer, J., Toimil, A., Álvarez-Cuesta, M., Losada, I. J., 2021. Deep uncertainties in shoreline change projections: an extra-probabilistic approach applied to sandy beaches. Nat. Hazards Earth Syst. Sci. Discuss 2021, 1–24. https://doi.org/10.5194/nhess-2020-412.

Toimil, A., Camus, P., Losada, I.J., Le Cozannet, G., Nicholls, R.J., Idier, D., Maspataud, A., 2020. Climate change-driven coastal erosion modelling in temperate sandy beaches: methods and uncertainty treatment. Earth Sci. Rev. 202, 103110. https://doi.org/10.1016/j.earscirev.2020.103110.

Toimil, A., Losada, I.J., Camus, P., Díaz-Simal, P., 2017. Managing coastal erosion under climate change at the regional scale. Coast. Eng. 128, 106–122.

Tolson, B.A., Sharma, V., Swayne, D.A., 2014. Parallel implementations of the dynamically dimensioned search (DDS) algorithm. In: Proceedings of the 7th International Symposium on Environmental Software Systems, pp. 573–582. Prague, Czech Republic.

Tolson, B.A., Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resour. Res. 43 (1). https://doi.org/10.1029/2005WR004723.

Tran, Y.H., Barthélemy, E., 2020. Combined longshore and cross-shore shoreline model for closed embayed beaches. Coast. Eng. 158, 103692. https://doi.org/10.1016/j.coastaleng.2020.103692.

Tran, Y.H., Marchesiello, P., Almar, R., Ho, D.T., Nguyen, T., Thuan, D.H., Barthélemy, E., 2021. Combined longshore and cross-shore modeling for low-energy embayed sandy beaches. J. Mar. Sci. Eng. 9 (9), 979. https://doi.org/10.3390/jmse9090979.

Turki, I., Medina, R., Coco, G., Gonzalez, M., 2013. An equilibrium model to predict shoreline rotation of pocket beaches. Mar. Geol. 346, 220–232. https://doi.org/10.1016/j.margeo.2013.08.002.

Turner, I.L., Harley, M.D., Short, A.D., Simmons, J.A., Bracs, M.A., Phillips, M.S., Splinter, K.D., 2016. A multi-decade dataset of monthly beach profile surveys and

inshore wave forcing at Narrabeen, Australia. Sci. Data 3 (1), 160024. https://doi.org/10.1038/sdata.2016.24.
US Army Corps of Engineers, 1984. Department of the Army, waterways experiment station, Corps of Engineers. In: Shore Protection Manual, fourth ed. Vicksburg, Mississippi.
van Ijzendoorn, C.O., de Vries, S., Hallin, C., Hesp, P.A., 2021. Sea level rise outpaced by vertical dune toe translation on prograding coasts. Sci. Rep. 11 (1), 12792. https://doi.org/10.1038/s41598-021-92150-x.
van Maanen, B., Nicholls, R.J., French, J.R., Barkwith, A., Bonaldo, D., Burningham, H., Brad Murray, A., Payo, A., Sutherland, J., Thornhill, G., Townend, I.H., van der Wegen, M., Walkden, M.J.A., 2016. Simulating mesoscale coastal evolution for decadal coastal management: a new framework integrating multiple, complementary modelling approaches. Geomorphology 256, 68–80. https://doi.org/10.1016/j.geomorph.2015.10.026.
Viola, C.N.A., Verdon-Kidd, D.C., Hanslow, D.J., Maddox, S., Power, H.E., 2021. Long-term dataset of tidal residuals in New South Wales, Australia. Data 6 (10). https://doi.org/10.3390/data6100101.
Vitousek, S., Barnard, P.L., 2015. A nonlinear, implicit one-line model to predict long-term shoreline change. In: Proceedings of Coastal Sediments 2015. World Scientific, San Diego, CA.
Vitousek, S., Barnard, P.L., Limber, P., 2017a. Can beaches survive climate change? J. Geophys. Res.: Earth Surf. 122 (4), 1060–1067.
Vitousek, S., Barnard, P.L., Limber, P., Erikson, L., Cole, B., 2017b. A model integrating longshore and cross-shore processes for predicting long-term shoreline response to climate change. J. Geophys. Res.: Earth Surf. 122 (4), 782–806. https://doi.org/10.1002/2016JF004065.
Vitousek, S., Cagigal, L., Montaño, J., Rueda, A., Mendez, F., Coco, G., Barnard, P.L., 2021. The application of ensemble wave forcing to quantify uncertainty of shoreline change predictions. J. Geophys. Res.: Earth Surf., e2019JF005506 https://doi.org/10.1029/2019JF005506.
Vitousek, S., Vos, K., Splinter, K.D., Erikson, L., Barnard, P.L., 2023. A model integrating satellite-derived shoreline observations for predicting fine-scale shoreline response to waves and sea-level rise across large coastal regions. J. Geophys. Res.: Earth Surf. 128 (7), e2022JF006936. https://doi.org/10.1029/2022JF006936.
Vos, K., Harley, M.D., Splinter, K.D., Simmons, J.A., Turner, I.L., 2019a. Sub-annual to multi-decadal shoreline variability from publicly available satellite imagery. Coast. Eng. 150, 160–174. https://doi.org/10.1016/j.coastaleng.2019.04.004.
Vos, K., Harley, M.D., Turner, I.L., Splinter, K.D., 2023. Pacific shoreline erosion and accretion patterns controlled by El Niño/Southern Oscillation. Nat. Geosci. 16 (2), 140–146. https://doi.org/10.1038/s41561-022-01117-8.
Vos, K., Splinter, K.D., Harley, M.D., Simmons, J.A., Turner, I.L., 2019b. CoastSat: a Google Earth Engine-enabled Python toolkit to extract shorelines from publicly available satellite imagery. Environ. Model. Software 122, 104528. https://doi.org/10.1016/j.envsoft.2019.104528.
Vousdoukas, M.I., Ranasinghe, R., Mentaschi, L., Plomaritis, T.A., Athanasiou, P., Luijendijk, A., Feyen, L., 2020. Sandy coastlines under threat of erosion. Nat. Clim. Change 10 (3), 260–263. https://doi.org/10.1038/s41558-020-0697-0.
White, N.J., Haigh, I.D., Church, J.A., Koen, T., Watson, C.S., Pritchard, T.R., Watson, P. J., Burgette, R.J., McInnes, K.L., You, Z.-J., Zhang, X., Tregoning, P., 2014. Australian sea levels—trends, regional variability and influencing factors. Earth Sci. Rev. 136, 155–174. https://doi.org/10.1016/j.earscirev.2014.05.011.
Wishaw, D., Leon, J., Fairweather, H., Crampton, A., 2021. Influence of wave direction sequencing and regional climate drivers on sediment headland bypassing. Geomorphology 383, 107708. https://doi.org/10.1016/j.geomorph.2021.107708.
Woodroffe, C.D., Cowell, P.J., Callaghan, D.P., Ranasinghe, R., Jongejan, R., Wainwright, D.J., Barry, S., Rogers, K., Dougherty, A.J., 2012. Approaches to Risk Assessment on Australian Coasts: A Model Framework for Assessing Risk and Adaptation to Climate Change on Australian Coasts. National Climate Change Adaptation Research Facility. Gold Coast.
Wright, L.D., Short, A.D., 1984. Morphodynamic variability of surf zones and beaches: a synthesis. Mar. Geol. 56 (1), 93–118. https://doi.org/10.1016/0025-3227(84)90008-2.
Wright, L.D., Short, A.D., Green, M.O., 1985. Short-term changes in the morphodynamic states of beaches and surf zones: an empirical predictive model. Mar. Geol. 62 (3), 339–364. https://doi.org/10.1016/0025-3227(85)90123-9.
Yates, M.L., Guza, R.T., O'Reilly, W.C., 2009. Equilibrium shoreline response: observations and modeling. J. Geophys. Res.: Oceans 114 (C9). https://doi.org/10.1029/2009JC005359.
Yates, M.L., Guza, R.T., O'Reilly, W.C., Hansen, J.E., Barnard, P.L., 2011. Equilibrium shoreline response of a high wave energy beach. J. Geophys. Res.: Oceans 116 (C4). https://doi.org/10.1029/2010JC006681.
Zarifsanayei, A.R., Antolínez, J.A.A., Etemad-Shahidi, A., Cartwright, N., Strauss, D., 2022. A multi-model ensemble to investigate uncertainty in the estimation of wave-driven longshore sediment transport patterns along a non-straight coastline. Coast. Eng. 173, 104080. https://doi.org/10.1016/j.coastaleng.2022.104080.
Zeng, L., Xiong, L., Liu, D., Chen, J., Kim, J.-S., 2019. Improving parameter transferability of GR4J model under changing environments considering nonstationarity. Water 11 (10). https://doi.org/10.3390/w11102029.
Zhang, K., Douglas, B.C., Leatherman, S.P., 2004. Global warming and coastal erosion. Clim. Change 64 (1–2), 41–58.
Zhang, X., Liu, P., 2021. A time-varying parameter estimation approach using split-sample calibration based on dynamic programming. Hydrol. Earth Syst. Sci. 25 (2), 711–733.