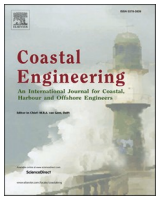
Coastal Engineering 203 (2026) 104856

Contents lists available at ScienceDirect

Coastal Engineering

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



Predicting shoreline changes using deep learning techniques with Bayesian optimisation

Tharindu Manamperi [a,*], Alma Rahat [a], Doug Pender [b], Demetra Cristaudo [b], Rob Lamb [c,d], Harshinie Karunarathna [a]

[a] Faculty of Science and Engineering, Swansea University, United Kingdom
[b] JBA Consulting, United Kingdom
[c] JBA Trust, United Kingdom
[d] Lancaster Environment Centre, Lancaster University, United Kingdom

ARTICLE INFO

Keywords:
Shoreline prediction
Deep learning
LSTM
Bayesian optimisation

ABSTRACT

Accurate prediction of shoreline change is vital for effective coastal planning and management, especially under increasing climate variabilities. This study explores the applicability of deep learning (DL) techniques, particularly Long Short-Term Memory (LSTM) and Convolutional Neural Network-LSTM (CNN-LSTM) models, for shoreline forecasting at monthly to inter-annual timescales, under two modelling approaches—direct input (DI) and autoregressive (AR). All models demonstrated the ability to reproduce temporal shoreline variability, while the autoregressive DL models were performing better.

Further, a noise impact assessment revealed that seasonal decomposition and noise filtering significantly enhanced the model performance. In particular, the models using 52-week data decomposition and residual noise reduction improved the model performance. The reduction of data noises also resulted in narrower ensemble prediction envelopes, indicating that ensemble candidate models behave with low diversity. The temporal data resolution analysis showed that lower data resolutions reduce the predictive performance of the model and at least fortnightly data are required to satisfactorily capture the trend of variability of the shoreline position at this beach.

The use of ensemble predictions, derived from a selected subset of model trials based on their collective performance, proved beneficial by capturing diverse temporal behaviours, thereby offering a quasi-probabilistic forecast with minimal computational cost. Overall, the study underscores the potential of DL models, particularly with autoregressive architectures, for reliable and transferable shoreline change prediction. It also emphasizes the importance of data quality, resolution, and preprocessing in improving model robustness, laying the groundwork for future research into use of DL in multi-scale shoreline predictions.

1. Introduction

Beaches are crucial parts of our landscapes, which are highly dynamic and constantly change due to wave action and other hydrodynamic drivers. Climate change will have a significant impact on the beaches, with the possibility of substantial loss of half the sandy beaches in the world due to coastal erosion (Vousdoukas et al., 2020). It is, therefore, essential for policymakers and engineers to identify areas at risk of coastal erosion and flooding in future (Ibaceta and Harley, 2024), and then sustainably manage them, especially focusing on environmentally and economically high-value areas (Turner et al., 2024).

Beach morphodynamic change prediction is one of the key elements of coastal change assessment and sustainable management with climate change adaptation (Brommer and Bochev-van Der Burgh, 2009; Karunarathna et al., 2018). Sustainable coastal zone management approaches and coastal defence schemes can be designed based on an enhanced understanding and the ability to predict how the beaches change in short (episodic to seasonal) to medium (i.e. between 1 and 10 years) term time scales, when responding to the uncertain driving forces.

Beach change assessment and prediction can be done either in two-dimensional planforms considering both cross-shore and alongshore changes or focusing on the dominant direction of change such as cross-

* Corresponding author.
E-mail addresses: 2252415@swansea.ac.uk, tharindu.udasri@gmail.com (T. Manamperi).

https://doi.org/10.1016/j.coastaleng.2025.104856
Received 8 June 2025; Received in revised form 27 July 2025; Accepted 15 August 2025
Available online 16 August 2025

sectional profile change or shoreline planform change. Particularly, in cross-shore sediment transport dominant beaches, predicting the beach change over these different timescales (storm scale, inter-annual change, up to decadal scales) is essential (Karunarathna and Reeve, 2013; Montaño et al., 2021), as it allows to differentiate whether the changes are, periodic or permanent (Barnard et al., 2017; Dodet et al., 2019; Harley et al., 2017; Masselink et al., 2016; Ranasinghe, 2016).

Conventional numerical coastal morphodynamic prediction models like DELFT3D (Lesser et al., 2004) and XBeach (Roelvink et al., 2009) rely on governing equations that simulate physical processes of wave propagation, hydrodynamics and morphodynamics (Roelvink et al., 2016). However, these models may demand high computational resources and may have limited applicability when applied to predict morphological changes in the long-term (Weber de Melo et al., 2023). Particularly, simulations over a long period of time and longer coastal stretches and, the need for simulating a wide range of scenarios for probabilistic predictions (Ranasinghe, 2020), make them computationally expensive and time-consuming. Also, model skills tend to degrade for medium to longer-term forecasting applications due to a range of inherent uncertainties (Weber de Melo et al., 2023) and requirements for simplification of physical processes as in reduced-physics process models. This may cause them less suitable for long term beach change (Ibaceta et al., 2022; Montaño et al., 2021; Simmons et al., 2017). In addition, most of the sediment transport equations are empirically derived and may only partially valid for some types of beaches.

Another group of models that has been developed to reduce the complexity of conventional process-based models are based on equilibrium state of the beach (Davidson et al., 2013). In these models, the driving force disturbance on the equilibrium state of the beach is emphasised and the evolution of the beach in the short-to long-time scales is predicted (Davidson and Turner, 2009; Yates et al., 2009, 2011). The Yates et al., 2009 showed that the beach state changes due to the energy disequilibrium were dependent on the antecedent beach state and therefore successfully presented 'beach memory' in a shoreline prediction model. When substantial data of historic beach change is available, data driven modelling approaches provide an alternative for predicting future beach change Różyński, 2003 combined Empirical Orthogonal Functions and Canonical Correlation Analysis (CCA) to investigate long term bar change at a beach in Poland. Horrillo-Caraballo and Reeve, 2010 used CCA to predict beach profile changes over annual time scales. Karunarathna et al. (2012); Karunarathna and Reeve (2013) used a hybrid approach that combined a reduced-physics model and measured beach profile data to predict future beach profile change. Although mostly site-specific, both the equilibrium, and data-driven models were successful in reducing the computational demand of process-based models and provided meaningful beach change predictions.

The increasing availability of comprehensive, high-resolution beach change data collected on-site or remotely (Vos et al., 2019), pave the way to explore new data-driven approaches for beach change forecasting. As a result, the use of Machine Learning (ML) for beach change forecasting is becoming popular (Ibaceta and Harley, 2024; Vitousek et al., 2023; Vos et al., 2019). ML algorithms learn the correlations between multiple variables directly from data, capturing patterns and relationships that are difficult to model and parameterised using traditional methods such as simple empirical or statistical data driven models (Gomez-de la Peña et al., 2023; Hashemi et al., 2010; Kim and Lee, 2022; Risha and Liu, 2025). With the proven success in similar fields such as hydrological forecasting and other environmental-related fields (Frame et al., 2022; Mai et al., 2022 and several studies), the use of ML approaches for coastal morphodynamic assessments and forecasting, is becoming increasingly popular (Adeli et al., 2025; Simmons and Splinter, 2025) while it is becoming an attractive alternative as long term morphodynamic datasets are increasingly become available.

It is crucial to ensure that training of ML models is focused on relevant correlations between the driving forces and the outputs, at the appropriate timescales. Therefore, the selection of data, their resolution and the possible driver-dependant relationships, and the optimum data requirement (Splinter et al., 2013) to represent the morphodynamic processes adequately, need a careful attention (Kim and Lee, 2022). The optimum amount of data and the optimised model architecture depend on the complexity of the process being involved and the capacity of the selected ML approach to build relationships between the drivers and the dependent variable from the available data (Splinter et al., 2013). Therefore, it is essential to understand and apply a comprehensive model optimisation process to gain the optimum prediction outcomes from the ML models.

An advanced subfield of ML known as Deep learning (DL) techniques have become popular for applications needing identification of nonlinear and complex correlations from the data. For complex coastal morphodynamic processes that are inherently time-dependent and influenced by uncertain, dynamic driving forces, DL offers a strong capability to learn complex temporal dependencies directly from data (Benidis et al., 2023). Traditional neural networks—the foundational architecture of deep learning—use feedforward structures where data flows through stacked layers of interconnected nodes without feedback, limiting their ability to retain memory of past inputs. This makes them less suited for time-dependent data. Recurrent Neural Networks (RNNs), and particularly Long Short-Term Memory (LSTMs) (Hochreiter and Schmidhuber, 1997), a specific form of RNN, address this limitation by incorporating feedback loops, enabling them to learn from temporal contexts (Lindemann et al., 2021). While the LSTMs are widely acknowledged for time series predictions, several other deep-learning techniques such as Convolutional Neural Networks (CNN) (Krizhevsky et al., 2017) have been developed to extract deep features, i.e. short term fluctuations, temporal shifts, multiple time scale trends and periodicities, from data and to be used alongside LSTMs to improve their ability to perform better (Zhou et al., 2023). Following the success of the LSTMs on time-series predictions, a few attempts have been made to apply this approach to coastal science and engineering. Those include Sea level rise impact assessment (Accarino et al., 2021; Balogun and Adebisi, 2021), wave climate forecasting (Lawal et al., 2023), storm prediction (Frifra et al., 2024; Tiggeloven et al., 2021) and shoreline change (Gomez-de la Peña et al., 2023) predictions.

Gomez-de La Peña et al. (2023) also used CNN and CNN-LSTM hybrid approach to predict shoreline change using daily images of shoreline change for seventeen years using wave-forcing and climatic indices as inputs to the model. Their results have shown that the 2-year shoreline prediction with only wave-forcing inputs, out-performed the benchmark statistical shoreline change models, *ShoreFor* (Davidson et al., 2013) and *SPADS* (Montaño et al., 2021; Simmons and Splinter (2025) explored an autoregressive Multi-Layer Perceptron (MLP) neural network based, shoreline prediction and the results showed that it is capable of reproducing the shoreline variabilities at storm dominated beaches for a time scale long up to a few years. Recently, Chen et al., (2025) had applied a modified *ShoreFor* model (Davidson et al., 2013) at Hasaki Beach. Their results show suboptimal performance, indicating that the equilibrium-based modelling approach may have limited success in capturing the complex, non-linear shoreline dynamics at this Hasaki Beach. These empirical equilibrium models may not work for certain sites (as Chen et al., (2025) have shown the *ShoreFor* model and Banno et al., 2020 have shown the Bruun rule, on Hasaki beach). They may only be good for predictions at certain timescales and certain occurrences of wave climates, though they are advantageous in simplicity and reduced reliance on extensive site-specific calibrations.

Taking shoreline change (defined in Section 3) as a proxy for cross-shore beach profile change, in this study, we explore two LSTM-driven techniques, each under two supervised learning approaches for shoreline change forecasting at the timescales of months to a several years. Hasaki Beach in Japan is taken as our case study site, described in Section 2. Our aim here is to demonstrate application of DL for shoreline prediction by evaluating LSTM models through two supervised

approaches: (i) direct prediction using wave parameters (Hs, Tp) and, (ii) an auto-regressive method using previously predicted shorelines combined with wave parameters. The second approach, which combines LSTM Deep Learning model with auto-regressive prediction approach which embed the antecedent shoreline state as a driving force will be a more realistic approach for shoreline change modelling. We will explore the stability and predictive robustness of LSTMs, a gated (unique memory cell incorporated) RNN under an auto-regressive learning approach (ii), compared to the previously studied direct-input learning approach (i). It benchmarks both methods and explores hyperparameter optimisation using Bayesian techniques, presenting both top-performing models and ensembles to assess robustness. The novelty of our approach lies in the integration of LSTM-based gated memory with autoregressive iterative learning, providing new insights into a stable and cost-effective predictive model.

The study also investigates the impact of noisy inputs, assessing whether smoothing or decomposition techniques improve performance—an important consideration given the irregularities common in real-world beach profile data. Additionally, it examines how input data frequency (e.g., weekly, fortnightly, and seasonal) influences model accuracy, offering practical insights for optimal beach monitoring frequencies. By analysing the effects of data quality, preprocessing, and signal decomposition, this research provides guidance on tailoring LSTM models, to be trained with ground-measured data, for specific forecasting needs, ranging from capturing long-term trends to predicting short-term fluctuations, thereby supporting a more informed and realistic deployment of DL techniques in coastal morphodynamic predictions.

The rest of the paper is structured as follows: Section 2 of the paper, describes the study site - Hasaki beach, from which, the data for model building is derived. In section 3, a brief overview of the modelling approaches and the DL techniques used in the study, along with the model development, is presented. Section 4 presents and discusses the results. Section 6 concludes the paper with conclusive remarks.

2. Study site and data

Hasaki Beach, located in the east coastline of Japan, was selected as the case study site. Hasaki is a 16 km long, longshore uniform sandy beach consisting of sediment with median grain size (D50) of 0.18 mm. The beach is subjected to a wave-dominated, micro-tidal environment with a tidal range of circa 1.4 m. The high (HWL), mean (MWL) and low (LWL) water levels are 1.25 m, 0.65 m and −0.20 m, respectively, with respect to the datum level at Hasaki (Tokyo Peil −0.69 m) (Kuriyama, 2002). Deepwater waves have been measured off the cost of Hasaki at a water depth of 24 m, using an ultrasound wave gauge for 20 min every 2 h. The wave data was collected by the Port and Airport Research Institute of Japan for 32 years from 1987 to 2018. Wave directions have not been measured. A time series of 2-Hourly significant wave height Hs and peak period Tp have been derived by the measured wave data. The long-term mean Hs and Tp determined from the measured wave data are found to be is 1.35 m (with a standard deviation of 0.73 m) and 8 s (with a standard deviation of 1.7 s) respectively.

Hasaki has a unique beach profile survey dataset spanning across several decades since 1987. Weekly cross-shore beach profile measurements are available from the Hasaki Observational Research Station (HORS) pier which extends offshore to a length of approx. 400m, facilitating the measurement of profile heights at 5m cross-shore intervals (Banno et al., 2020). Due to the longshore uniformity of the Hasaki Beach, the profiles measured at HORS are representative of the entire length of the beach (Kuriyama, 2002). The longshore uniformity and the high resolution, extensively long beach profile surveys and concurrent wave measurements, make Hasaki an ideal candidate for applying data-driven ML techniques for beach change forecasting.

The mean beach profile at Hasaki, determined from cross-shore profiles measured weekly over 32 years, from 1987 to 2018 (Fig. 1), has a foreshore-nearshore average slope of approximately 1:40 and an average intertidal area slope of 1:100 and can be classified as a gentle sloped beach. The shoreline position data were extracted from the measured profiles. For this study, the shoreline position is defined as the x axis coordinate at which the beach profile intersects with the 0m elevation w.r.t the Hasaki Port Datum, in line with previous studies from (Banno et al., 2020; Banno and Kuriyama, 2014). It is to be noted that data were arranged such that all chainages are positive.

It should be noted that beach profile data is collected on a weekly basis, while the wave data is collected at 2-h frequency. To ensure the same frequency across different variables, we computed the mean *Hs* and *Tp* over a week, and use that to construct a dataset $(Hs_{[0:T]}, Tp_{[0:T]}, SLP_{[0:T]})$, where 0 is the first data time index and T is the last time index,

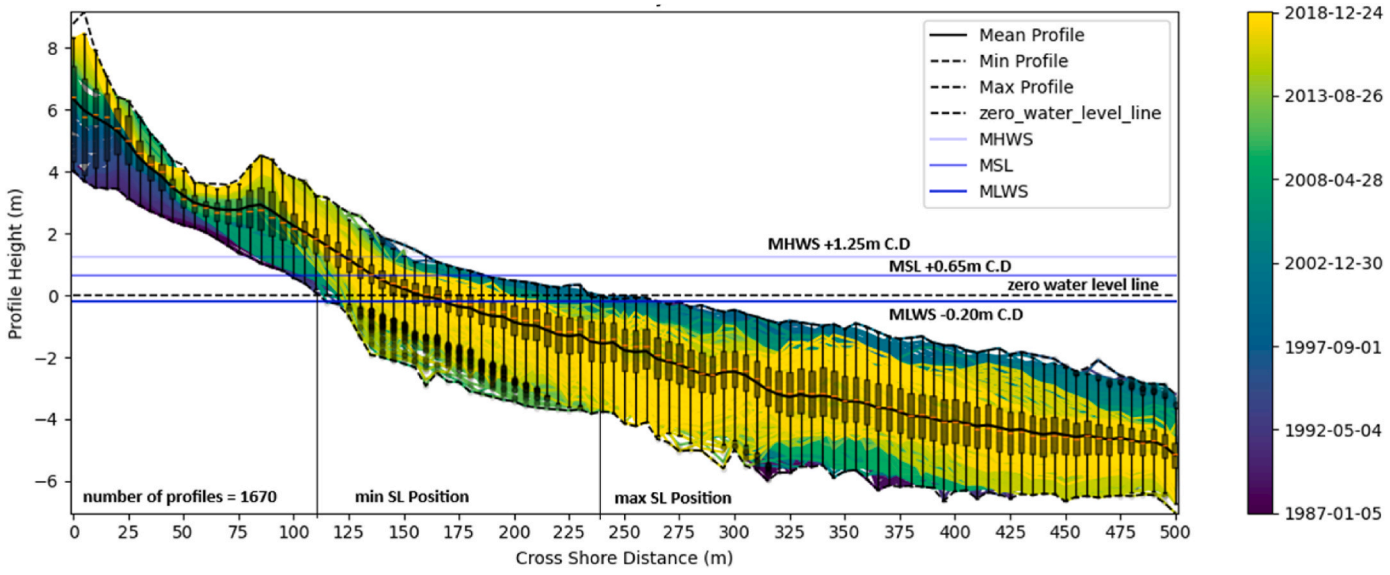


Fig. 1. Observed weekly beach profiles for the period between 1987 and 2018 and the respective water levels for Hasaki beach, Japan. (Mean high water spring (MHWS), Mean Sea level (MSL), Mean low water spring (MLWS) lines respective to Chart Datum (CD) are shown with horizontal blue lines). The box plots represent the profile height variation over the time at each interval step of 5m. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

and the subscript $[0:T]$ represents a sequence of measurements/observations in the time period, considered for training and testing of the models. We visualise the time series data constructed, in Fig. 2.

3. Methods

3.1. Modelling approaches

In this paper, our goal is to predict the future values of shoreline position *SL*P. The cross-shore beach change, and, consequently, the time evolution of the shoreline, is primarily driven by the waves acting on the shoreface where both the significant wave height *Hs* and the peak wave period *Tp* are the dominant drivers. In addition (Pender and Karunarathna, 2013; Eichentopf et al., 2020), and several other studies have found that the antecedent beach state (in this case *SLP*) also have a significant impact on the future shoreline changes following wave action. With this, we decided to explore two modelling approaches here:

1. **Direct Input Model (DI)**: In this approach, we focus only on utilising the wave forcing inputs – *Hs* and *Tp* directly – to make predictions without considering the past values of the target variable. Following setup was used as the model and arrange data inputs accordingly:

$$\overline{SLP_t} = f_{DI}(Hs_{[t-k:t]}, Tp_{[t-k:t]}), \tag{1}$$

Where k is the time window length in weeks, for the historical observations (or the time sequence length) on *Hs* and *Tp* in the time period between $t-k$ and t, and $\overline{SLP_t}$ represents predictions at time t.

2. **Autoregressive Model (AR)**: In this approach, we use the past history of the target *SLP*, alongside the wave forcing inputs, to derive at the future predictions. We express this model as:

$$\overline{SLP_t} = f_{AR}(Hs_{[t-k:t]}, Tp_{[t-k:t]}, SLP_{[t-1]}), \tag{2}$$

following the notations used in (1), but with the addition of measured data $SLP_{[t-1]}$ at the previous time step $t-1$. During the prediction, it will be using the predictions made at the previous time step as an input, demonstrating a more realistic shoreline evolution. Here it is called 'autoregressive' because the antecedent shoreline position has a direct influence on the future predictions.

The rationale behind exploring these approaches is to identify how the deep learning models perform in the presence of prediction feedback as an input, in other words the 'beach memory state', as an input variable. The data structure under above two approaches are given in Annex 01. 3.

It should be noted that in this study, following a preliminary sensitivity analysis through a trial-and-error approach on the sequence length k, we fixed $k = 12$, to enable models to learn the correlations among the wave forcing and the shoreline change using a 12-week past period, to represent an approximate seasonal duration. $k = 12$ was found to provide a suitable balance between capturing relevant temporal dependencies and avoiding overfitting or excessive computational cost.

3.2. Deep learning methods

According to the above setup with a time sequencing nature, we explored two deep learning techniques which are reputed in time series forecasting, in particular LSTM (Long short term Memory) and CNN-LSTM (Convolutional Neural Network incorporated LSTM), to

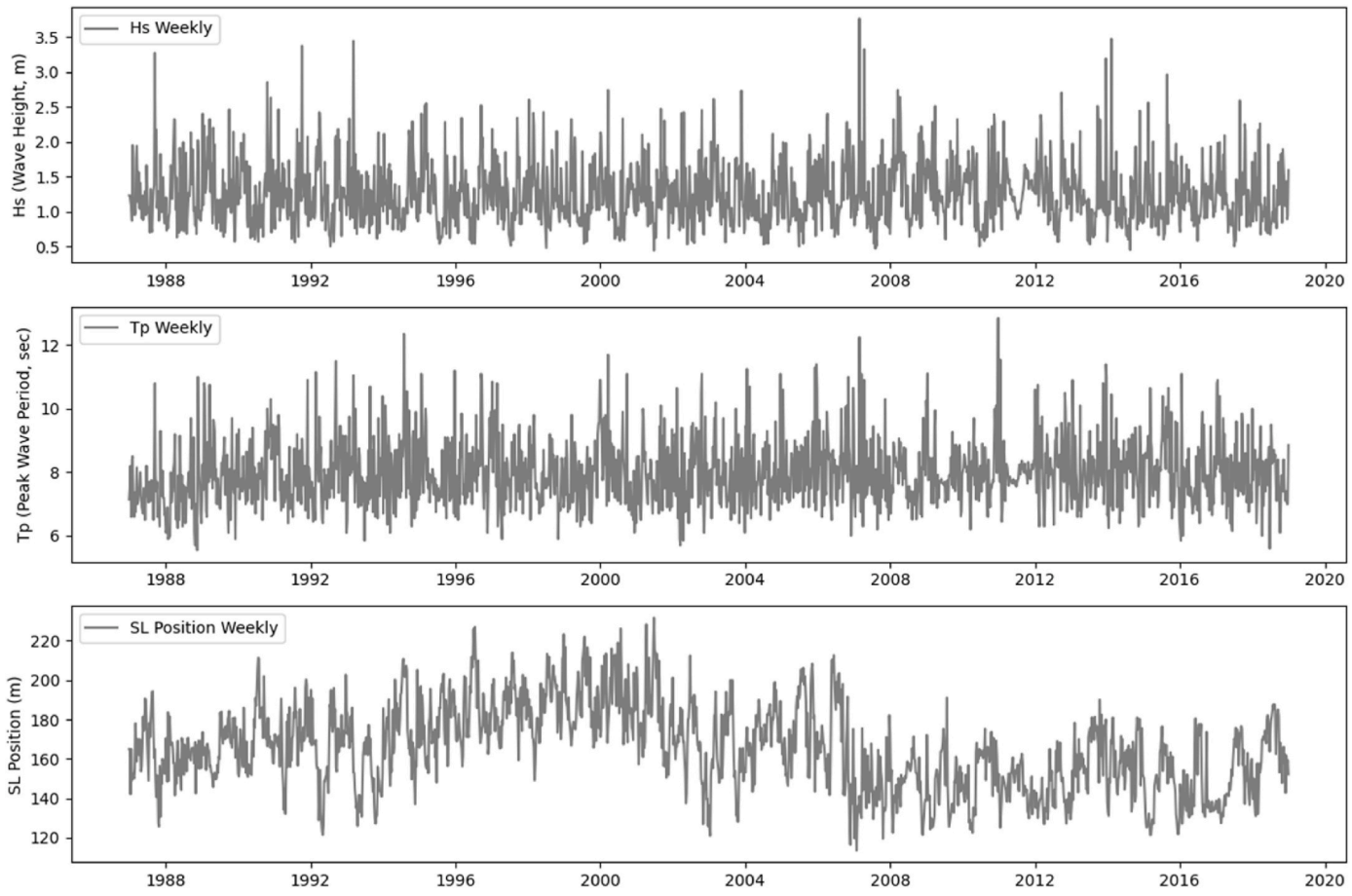


Fig. 2. Time series of (a) wave heights, weekly mean (b) wave periods, weekly mean (c) absolute shoreline positions data at Hasaki from 1987 to 2018.

construct data-driven models for $f_{DI}(Hs_{[t-k:t]}, Tp_{[t-k:t]})$ and $f_{AR}(SLP_{[t-1]}, Hs_{[t-k:t]}, Tp_{[t-k:t]})$.

LSTM is a special version of neural networks which includes a 'memory function' in the neuron cells (Hochreiter and Schmidhuber, 1997), which distinguishes them from the regular neural networks. This 'memory function' provides a unique ability to retain certain data relationship information in the cells for a certain time period and learn the sequential dependencies. This makes LSTMs a favourable candidate for time series predictions. CNNs apply convolutional filters to extract features from input data. Though originally used in computer vision (Lecun et al., 1998), they are now applied to data sequences. CNNs are faster to train and require fewer parameters than fully connected networks, as they rely on localized filtering rather than full connections (Albawi et al., 2017). In time series forecasting, CNNs are used to efficiently capture local patterns and short-term dependencies in the data. Combining CNNs with LSTMs leverages the strength of both architectures, making CNN-LSTM models well-suited to capture both local features and long-term temporal dependencies in time series data (Shi et al., 2015; Zhao et al., 2017). Here, we refrain from providing detailed mathematical and architectural descriptions of these methods as this is readily available from existing literature. Interested readers should refer to the work of Benidis et al., (2022), for a detailed tutorial and survey on the topic of deep learning in the context of time series forecasting.

In this study, two different LSTM types, multi-layer stacked LSTM (named 'LSTM' hereafter) and CNN layers stacked LSTMs (hereafter known as 'CNN-LSTM') will be used to predict time dependent shoreline position at Hasaki Beach using two modelling approaches. The direct input models will be noted as $LSTM_{DI}$ and $CNN\text{-}LSTM_{DI}$, and the autoregressive models will be denoted as $LSTM_{AR}$ and $CNN\text{-}LSTM_{AR}$. (A detailed description of the LSTM cell structure, CNN-LSTM layer structure and auto-regressive predictive approach is provided in Annex 01.)

3.3. Model training

The models were trained using data, prepared according to standard data feature engineering methods as follows. The data available at weekly resolution, was split into training and testing subsets. Following a data distribution analysis, and with the intention of allowing models to be trained and be able to capture as much temporal variations as possible, the first 30 years of data were used for training (Both training and validation) purpose, while the last 2 years of data were used to test the final model performance (as unseen observations).

The data sets were normalised using 'min-max normalisation', based on the minimum and maximum values of the input data at the training set and both sets were transformed (scaled) to a range between 0 and 1. This was done to ensure that all features contribute equally to the learning process and no information from the 'testing period' was used during the 'training'. This scaling transformation is important especially in neural networks where unscaled features can lead to unstable gradients, slower convergence, or biased weight updates. This helps to keep the model weights numerically stable and improves the overall training efficiency, particularly for models like LSTMs that are sensitive to input magnitude over time.

It is important to note that, the structure of any LSTM in general, is defined based on a set of model-defining higher-order sets of parameters called 'hyperparameters' (HP). These HPs decides the depth, memory capacity and the network's ability to learn the features and patterns from the input feature data. Therefore, it is essential to figure out an optimum set of HPs which defines an optimum model (providing the lowest loss function value on a separate validation set, which is not used to train) through a Hyperparameter Optimisation (HPO) process. This can be computationally expensive and yet not yield the optimum model which can generalise the model well for any input dataset, due to the inherent characteristic of the machine learning models and therefore is biased and depend on the data used to train the model. The process used for each of the model and their parameter space is being stated in section 3.4 under 'Hyperparameter Optimisation via cross validation'.

For the model training we optimised a modified Mielke's loss function (Duveiller et al., 2016) through the standard 'Adam' optimiser. This Mielke's loss function combines the error magnitude (in terms of Mean Square Error between model prediction and observed values, MSE) and agreement in trend (in terms of correlation coefficient, r), and defined as below.

$$Mielke's\ loss\ (SLP, \widehat{SLP}) = \frac{\sum_{i=1}^{n} (SLPi - \widehat{SLPi})^2}{{\sigma_{SLP}}^2 + {\sigma_{\widehat{SLP}}}^2 + (\overline{SLP} - \overline{\widehat{SLPi}})^2 + \kappa} \quad (3)$$

where $\kappa = \begin{cases} 2 \cdot |cov(SLP, \widehat{SLP})|, if\ r < 0 \\ 0, otherwise \end{cases}$,

and $SLPi$, $\widehat{SLPi}$, $\overline{\widehat{SLP}}$ represents the observed, model predicted and mean observed shoreline position values, ${\sigma_{SLP}}^2$, ${\sigma_{\widehat{SLP}}}^2$ represents variances of observed and modelled shoreline position values, r represents the correlation coefficient. The conditional kappa term depends on the r and valued at 2 times covariance or zero.

As described in equation (3), Mielke's loss function computes the loss accounting for both the variance structure of observed and model predicted series and the bias between them. This allows to weight larger errors more heavily when they occur at high variance periods such as during storm-induced shoreline erosion events. This results in a more context-aware loss landscape that emphasizes accurate reproduction of high-impact dynamic events. This has also been confirmed by Gomez-de la Peña et al., 2023 where Mielke's loss function has yielded improved performance evaluation over the conventional MSE loss function to train the model.

3.4. Hyperparameter optimisation via cross validation

In order to increase the possibility of reaching at the optimised HP combination within the selected HP space efficiently, a Bayesian Optimisation (BO) method based HPO technique was used in this study. Bayesian optimisation, particularly when combined with Tree Parzen Estimators (TPE) sampling (Watanabe, 2023), is a reputed method for HPO in LSTM networks (Snoek et al., 2012). It efficiently explores hyperparameter space by modelling the objective function (here the error on the validation set) probabilistically and selecting the next HP combination where improvement is most likely, balancing exploration (searching regions of the space that are least tested) and exploitation (focusing on regions known to yield good performance), (Bergstra et al., 2013). Each HP model at all HPO trials, was individually trained and validated on five different subsets of the 'training data set'. In each fold, a different size split of the training data to train the model ('train set') and validation set to validate ('validation set'), is used, preserving the time sequencing. This 'time-series cross-validation' was used to reduce the risk of overfitting and to improve the model generalisation to unseen data (Bergmeir and Benítez, 2012) as it retains the time sequencing order of the data (Gorriz et al., 2024).The average RMSE on five validation sets was considered to rank the trials on the HPO, which provides more generalised model training and validation. With the intention of maintaining a reasonable computational time on optimisation and account for number of hyperparameters in each model type, the number of HP trial runs was limited to 100 for LSTM models ($LSTM_{DI}, LSTM_{AR}$) and 150 for CNN-LSTM models ($CNN\text{-}LSTM_{DI}, CNN\text{-}LSTM_{AR}$). HPO was carried out using the state-of-the-art 'Optuna' package (Akiba et al., 2019). The specific HP ranges used for the HPO of LSTM and CNN-LSTM models are presented in Table 1.

Based on the ranks from the HPO, the performance of the 'best optimum' trial model (named as best optimised HP model, hereafter) is assessed using the 'testing set', which is the prediction duration. (kept unseen during training and validation process). This results in a single deterministic prediction, based on the unique set of correlations learnt

Table 1
Hyperparameter space selected to optimised for each LSTM model.

Hyperparameters	LSTM Lower/Upper limits	CNN-LSTM Lower/Upper limits
Layers	2, 4	2, 4
Nodes	50, 300	50, 300
Learning rate	0.001, 0.01	0.001, 0.01
Batch size	24, 92 [steps = 8]	24, 92 [steps = 8]
Epochs	50, 200	50, 200
Dropout rate	0.1, 0.4	0.1, 0.4
Kernels	–	24, 64 [steps = 8]
Kernel/filter size	–	2, 4

by the optimised HP model.

Each trial model in the HPO with different HPs, has been trained individually, resulted in a distinct model capable of capturing different temporal dependencies. Therefore, to reflect the model diversity and incorporate a broader range of learned temporal patterns, it was decided to produce an ensemble using the predictions of model variants from the HPO (Mienye and Sun, 2022). The 10 trial models which perform best 'together as an ensemble' among the HPO trial runs are selected to produce the ensemble (Lévesque et al., 2016). The 'mean ensemble' prediction was generated to represent the average prediction, hereafter referred to as the 'mean ensemble' prediction. The prediction envelope was produced from this ensemble using the minimum and maximum prediction at each time step. This ensemble approach captures model diversity rooted in hyperparameter variability and mitigates potential biases tied to selecting a single optimum model—biases that can stem from the HPO process and its defined search space. The 'mean ensemble' prediction is also a way of representing the uncertainty tied into a single model structure (Prince, 2023), though it is not focused here. The workflow structure used in all models is shown in Fig. 3.

3.5. Performance evaluation

A standard performance matrix was used to evaluate the predictions against the measured shorelines, including RMSE, normalised RMSE, Pearson correlation coefficient (r), and Mielke's Index—capturing multi-dimensional performance aspects rather than relying solely on error-based assessment. Normalised RMSE (NRMSE) in the performance matrix is defined as,

$$NRMSE = \frac{RMSE}{(\mathbf{max}(\mathbf{y_t}) - \mathbf{min}(\mathbf{y_t}))} \tag{4}$$

Where, *RMSE* is the Root Mean Square Error between predicted and observed values, $\mathbf{max}(\mathbf{y_t})$ is the maximum of the observed values, $\mathbf{min}(\mathbf{y_t})$ is the minimum of the observed values and $\mathbf{y_t}$ represents the true (observed) values over time. NRMSE was used to evaluate the prediction error, relative to the scale of the data, making it a more comprehensive comparison measure. Mielke's Index of Agreement (MI)

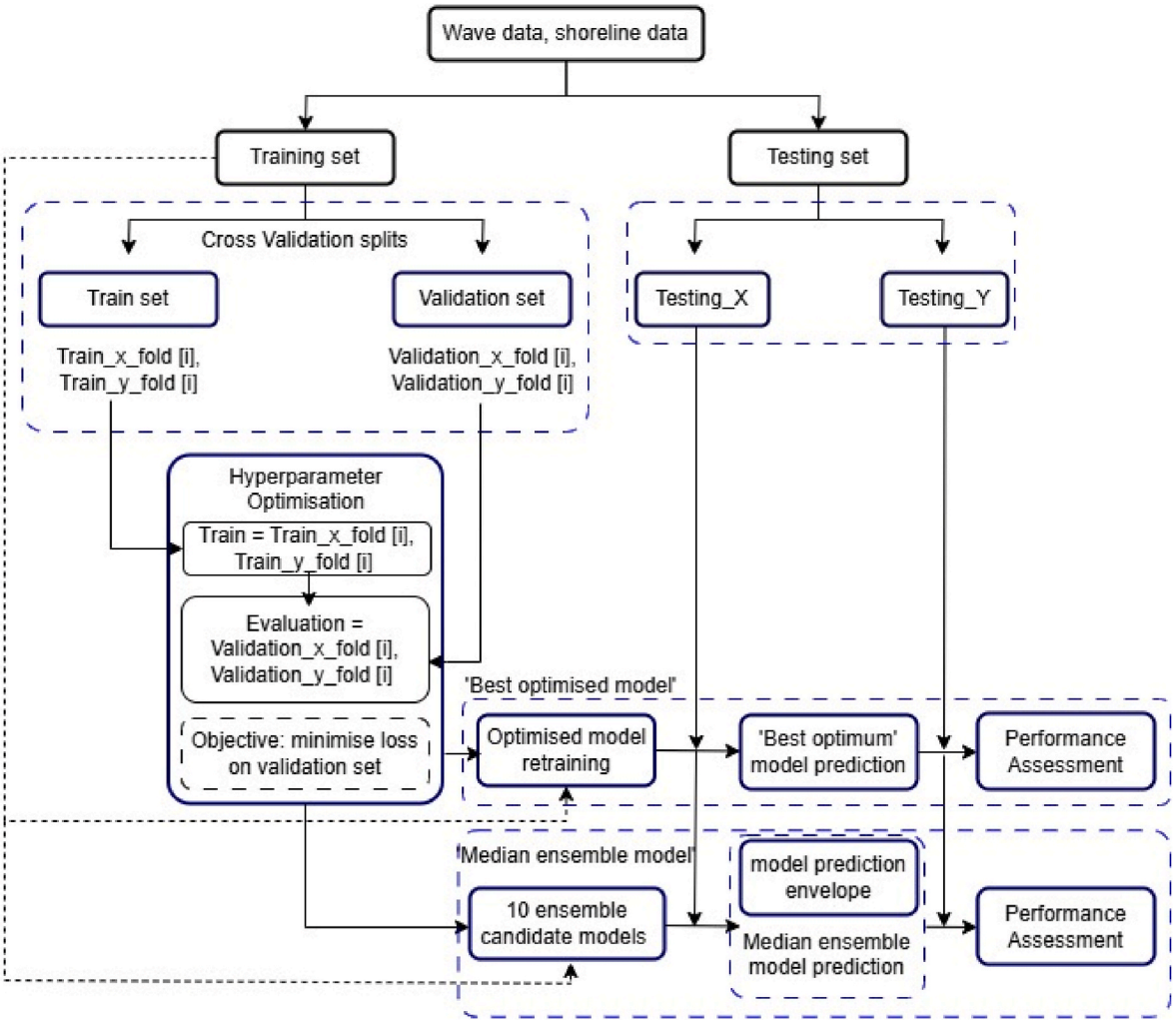


Fig. 3. The workflow structure used in general (irrespective of the model type and modelling approach). Train_x_fold, Train_y_fold, Validation_x_fold, Validation_y_fold, Testing_X, Testing_Y denotes the Inputs and targets at each of the split during the cross-validation process.

is defined as,

$$MI = 1 - \frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{\sum_{t=1}^{n} (|y_t - \bar{y}| + |\hat{y}_t - \bar{y}|)^2} \quad (5)$$

Where, y_t is the observed value at time step t, $\hat{y}_t$ is the predicted value at time step t, $\bar{y}$ is the mean of the observed values and n is the time steps predicted. MI evaluates both the accuracy and precision of the predictions, being sensitive to systematic deviations and magnitude differences. To evaluate the prediction ensemble, apart from the above measures, two additional parameters were considered. The empirical coverage probability (ECP) is defined as,

$$ECP = \left(\frac{1}{N} \sum_{t=1}^{N} I. \{ y_t \epsilon [\hat{y}_t^{\,L}, \hat{y}_t^{\,U}] \} \right) * 100\% \quad (6)$$

Where, $\hat{y}_t^{\,L}$ and $\hat{y}_t^{\,U}$ are the lower and upper bounds of the prediction envelope at time t, $I\{.\}$ is an indicator function that returns 1 if the true value y_t lies within the interval $[\hat{y}_t^{\,L}, \hat{y}_t^{\,U}]$, and 0 otherwise. N is the total number of time steps predicted. Here, ECP measures how often the true shoreline position falls within this envelope and will be an assessment of the reliability of the diverse model envelope. The average bandwidth is calculated to reflects the tightness or spread of the envelope (Schall, 2012) and defined as,

$$Average\ Band\ Width = \left(\frac{1}{N} * \sum_{t=1}^{N} (\hat{y}_t^{\,U} - \hat{y}_t^{\,L}) \right) \quad (7)$$

Where, $(\hat{y}_t^{\,U} - \hat{y}_t^{\,L})$ is the width of the prediction band at time step t.

4. Results and discussion

The four model types described in section 3.2, were trained and validated using the measured shoreline position data allocated for training and, the shoreline position predictions were carried out according to the procedure described in section 3.4. The shoreline positions predicted by the 'best optimised' model and the 'mean ensemble' are compared with the measured shoreline positions for the testing period, for each model type, based on the performance matrix, to explore each model type's predictive capacities. Then the selected model based on performance and complexity, was used to explore the impacts of the quality (noise) and the quantity (resolution) of data on model's predictive capability.

4.1. Model comparison

Fig. 4(a) compares the shoreline position prediction from the 'best optimised' model of the four model types, during the testing period. All four models capture the trend of variability of the measured shoreline position although the predictions from the 'auto-regressive' models ($LSTM_{AR}$; $CNN-LSTM_{AR}$), is seen to out-perform the other two, slightly. A similar comparison of 'mean ensemble' shoreline prediction from the four models is shown in Fig. 4(b).

Similar to the 'best optimised' model prediction, all four 'mean ensemble' predictions follow the trend of the measured shoreline position, though they do not follow short term fluctuations (weekly level), in common. The temporal variability of the 'mean ensemble' prediction of the shoreline positions from all four models are smoother than that from the 'best optimised' predictions. The full performance matrix of all predictions is given in Table 2. The 'best optimised' $CNN-LSTM_{AR}$, prediction has lowest RMSE, NRMSE and MPE, and the highest r, among the best optimised models. Considering the 'mean ensemble'

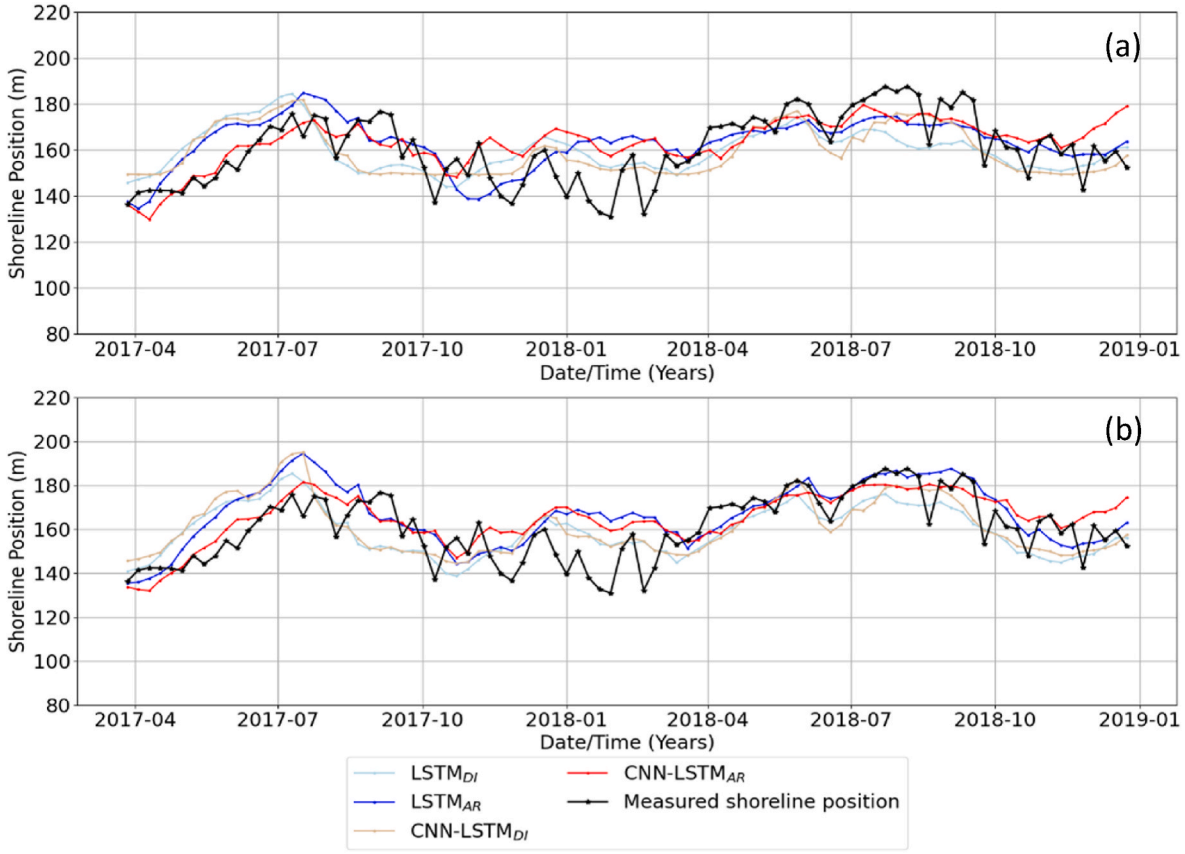


Fig. 4. (a) 'best optimised' model (b) 'mean ensemble' model predicted shoreline position from all models compared with the measured shoreline position.

Table 2
The performance matrix for the models for model application performance assessment.

Model	Optimised HP model					Mean of Ensemble models				
	RMSE (m)	r	MPE %	Mielke's Index	NRMSE (m)	RMSE (m)	r	MPE %	Mielke's Index	NRMSE (m)
LSTM $_{DI}$	13.51	0.45	−0.16	0.40	0.24	12.53	0.57	0.27	0.55	0.22
LSTM $_{AR}$	12.04	0.61	−2.08	0.57	0.21	12.84	0.69	−4.09	0.63	0.23
CNN-LSTM $_{DI}$	12.35	0.57	0.27	0.57	0.22	12.77	0.57	−0.79	0.56	0.23
CNN-LSTM $_{AR}$	11.83	0.63	−2.27	0.57	0.21	11.71	0.69	−3.23	0.63	0.21

predictions, they show similar results with $CNN-LSTM_{AR}$ shows stand out ensemble performance among others. Though it produces lowest RMSE and NRMSE values, the r value is highest and equal to the $LSTM_{AR}$. Considering the MI, the $LSTM_{AR}$ produces the best combined performance with the highest 0.69 for the 'mean ensembles' and equal highest of 0.57 for the 'best optimised' models.

Data density plots comparing the predicted shoreline positions from the 'best optimised' models of the four model types are shown in Fig. 5. The broken white line shows the perfect correlation between the predicted and measured shorelines. It can be seen that the data density surrounding the perfect correlation line is the largest for the 'auto regressive' models while they spread throughout.

indicating that auto regressive models perform well when predicting shoreline change under different levels of severity of forcing conditions where shoreline positions vary around the average. Also, the $LSTM_{AR}$ and $CNN-LSTM_{AR}$ models (Fig. 5(a) and (b)) demonstrate similar predictions at the higher shoreline positions, although they over predict the smaller shoreline positions. The data density plots of mean ensemble predictions (Fig. 5 (e), (f), (g) and (h)) shows better prediction spread and correlation with the observed values, compared to the 'optimised' models.

Fig. 6(a), (b), (c) and (d) show the envelopes of the shoreline positions predicted from the ensembles of $LSTM_{DI}$, $LSTM_{AR}$, $CNN-LSTM_{DI}$, and $CNN-LSTM_{AR}$, respectively. The prediction from the 'best optimised' model, the 'mean ensemble' prediction and the measured shoreline positions are also shown for comparison. The average ECP and the average bandwidth of the envelopes over the prediction period is given in Table 3.

The ensemble prediction envelope from the $LSTM_{DI}$ model (Fig. 6 (a)) gives a narrow ensemble envelop, with an average envelop width of 10.50 m. Around 20.4 % of the measured shoreline positions fell within the envelope. The 'mean ensemble' shoreline position prediction closely follows the prediction from the 'best optimised' model, showing consistency and lower diversity in the ensemble prediction. The ensemble envelope of shoreline position predictions from the $LSTM_{AR}$ model is shown in Fig. 6(b). It produces the widest prediction envelop among the four model types with average envelop width of 27.7 m, while 65.3 % of the true shoreline positions fall within this envelop. This indicates the high diversity among the candidate models in the ensemble. The shoreline prediction from the 'best optimised' model closely follows the 'mean ensemble' prediction.

As can be seen in Fig. 6(c), the ensemble shoreline prediction envelope from the $CNN-LSTM_{DI}$ is narrower than that from the $LSTM_{AR}$ and is highly fluctuating. The average ECP is 51.6 %, while the average envelop bandwidth is 19.5 m. The predictions from the 'best optimised' model and the 'mean ensemble' prediction are very similar, though the predictions show a high level of smoothness. While the 'best optimised' model in this case shows a better performance compared to the predictions from the other direct input model, $LSTM_{DI}$, the mean ensemble performances are quite similar, indicating that the CNN inclusion has not had much impacted the model performance is considered as an ensemble mean. However, the wider ensemble bandwidth which captures over 50 % of observed values, compared to 20.4 % in $LSTM_{DI}$ shows large diversity the CNN can integrate, when an ensemble is considered.

Fig. 6(d) shows the ensemble shoreline prediction envelop from the $CNN-LSTM_{AR}$ model. The average bandwidth of the envelop is 23.4 m and higher than both DI models, indicating high diversity among the candidate models in the ensemble. Over 71.4 % of the true shoreline fell within the envelop which is also the highest among all models. The prediction from the 'best optimum' variant and median ensemble are very similar, indicating good ensemble performance in terms of diversity and accuracy.

Overall, all four models have been able to perform satisfactorily at predicting temporal trends of shoreline position variability, particularly considering the past 12-week long memory state. All four models capture the trend of variability of the shoreline although some deviation can be seen when the shoreline position is smaller (indicating a retrieval). The mean ensemble produces better diverse performance, through a prediction envelope, indicating the diversity of the predictions from the ensemble candidates. Also, it reduces the dependency on a model with a single set of HPs. Even though auto-regressive models are expected to propagate the prediction errors forward, they perform better than the direct input models, potentially due to the taking the antecedent shoreline position as an input driver. Considering the performance matrix and the prediction envelops, the $CNN-LSTM_{AR}$ model performs marginally better than $LSTM_{AR}$. However, it was decided to use $LSTM_{AR}$ model in the proceeding analysis and discussions, as it performs closely to the $CNN-LSTM_{AR}$ and shows a better spread of predictions for both 'optimised' and 'mean ensemble' models. It also has a relatively simpler model structure with a smaller number of HPs to be optimised, requiring less HP trial models in the HPO, which will result in less computational resource requirement.

4.2. The effect of data quality and quantity on shoreline position predictions

The performance of ML models on predicting time series data depends on the data quality, noise, frequency and quantity. In this proceeding analysis, the impact of data noise on the time-dependent shoreline position predictions are investigated based on the '$LSTM_{AR}$' model. The shoreline positions predicted based on the weekly wave and shoreline position data (hereafter known as original data) using the $LSTM_{AR}$ model is taken as the baseline scenario for this assessment.

4.2.1. Impact of short term fluctuations of the input driving force data on the shoreline predictability

Preliminary investigations have shown that the performance of LSTM models is sensitive to short term fluctuations of the input driving force time series data that may be caused by physically meaningful events such as storm induced erosion) or irregular variations (data inherent noise). This can adversely impact model training and predictive performance, particularly in autoregressive forecasting setups. Therefore, it may be worthwhile exploring the optimal level of data fluctuations that can provide the desired timescales of trends of shoreline variabilities, through the model outcomes, without compromising the quality of prediction at the timescale of interest, although data inherent noise is out of scope of this study. Here we will investigate the sensitivity of the shoreline predictions to the short-term fluctuations of the measured data.

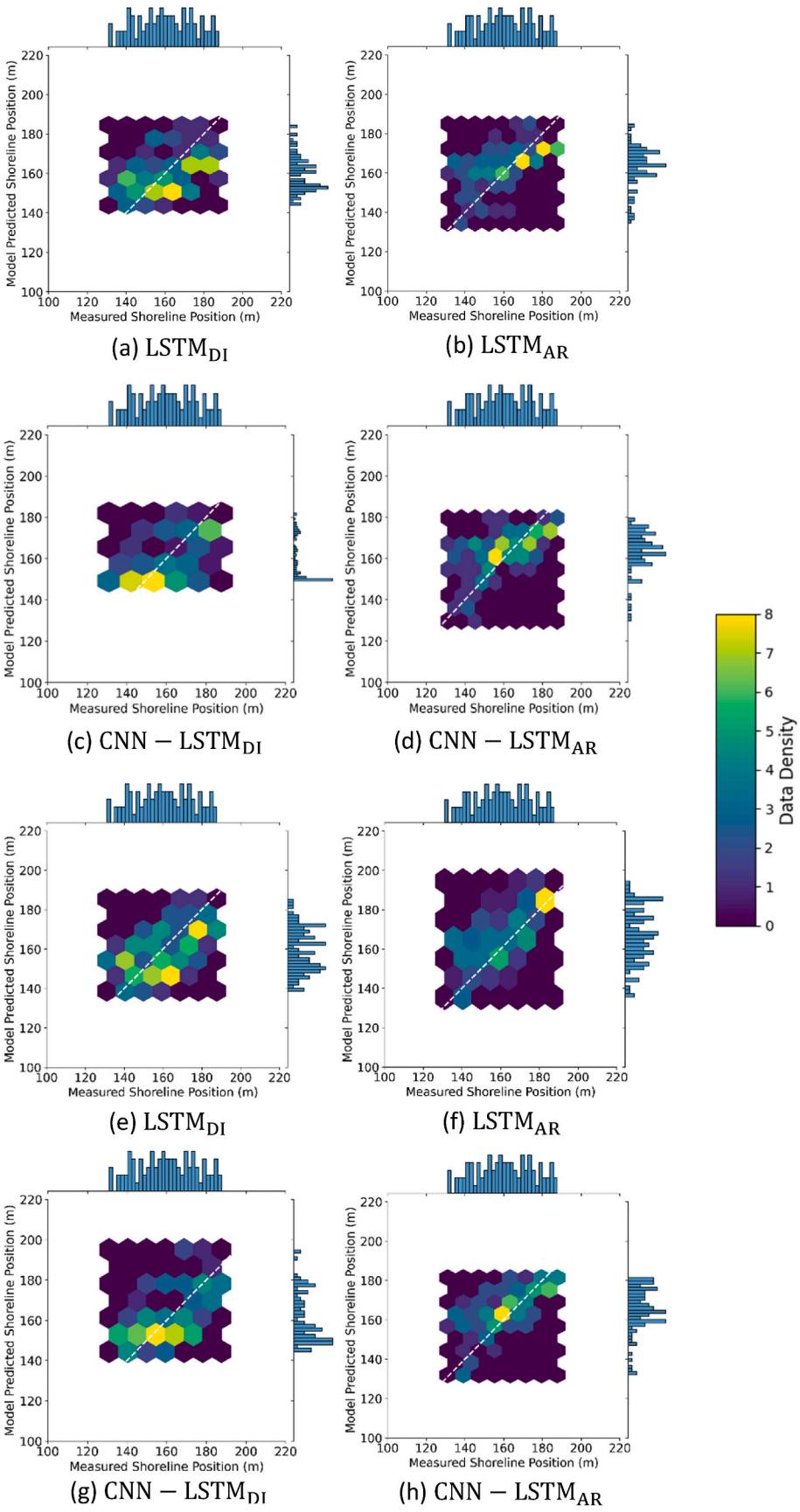


Fig. 5. Data density plots for model scenarios (a) LSTM$_{DI}$ (b) LSTM$_{AR}$ (c) *CNN* − LSTM$_{DI}$ (d) *CNN* − LSTM$_{AR}$ for 'best optimised' models and (e) LSTM$_{DI}$ (f) LSTM$_{AR}$ (g) *CNN* − LSTM$_{DI}$ (h) *CNN* − LSTM$_{AR}$ for 'mean ensemble' models.

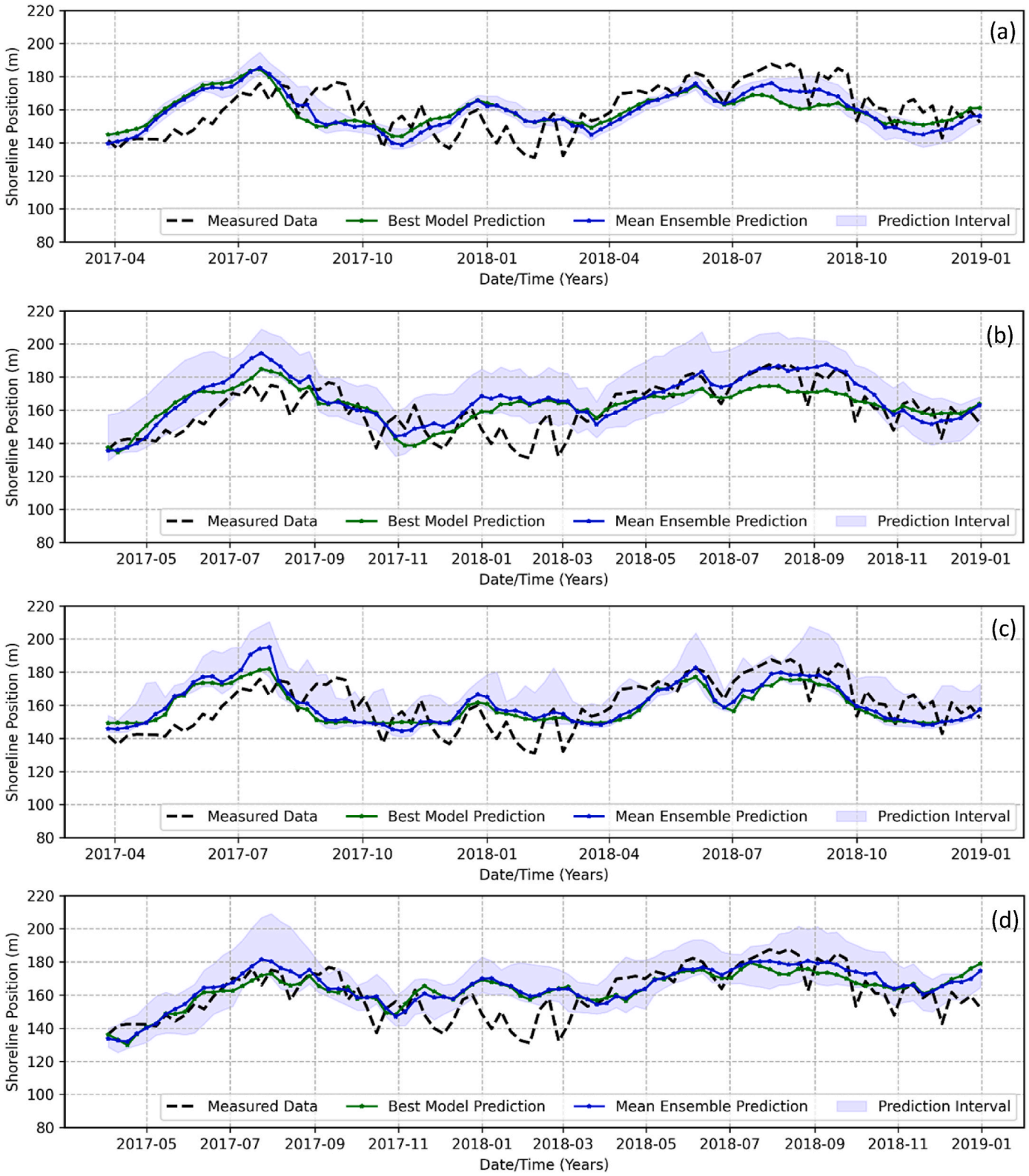


Fig. 6. Shoreline position predictions from 'best optimised' model, 'mean ensemble' prediction and ensemble prediction envelope compared with the measured shoreline position (a) LSTM_{DI} (b) LSTM_{AR} (c) $CNN-\mathrm{LSTM}_{DI}$ (d) $CNN-\mathrm{LSTM}_{AR}$

As the first scenario, the weekly input data (wave forcing and shoreline position) were pre-processed using the commonly known 'moving averaging' technique to eliminate a known short term fluctuation level (described later using the defined 'Noise ratio' and 'variance ratio'). Our focus here is on the monthly to annual scale shoreline change. Therefore, following a preliminary sensitivity analysis carried out involving a number of different moving averaging window sizes, a 4-week window was selected. This is expected to provide a balanced level of smoothing which is sufficient to suppress high-frequency variations (e.g., storm-scale events to weekly scale variations), while preserving

Table 3
Average ECP and average bandwidth of the shoreline position predictions from the selected ensembles of the $LSTM_{DI}$, $LSTM_{AR}$, *CNN* − $LSTM_{DI}$, *and CNN*− $LSTM_{AR}$ *models.*

Model	ECP%	Avg. band width (m)
$LSTM_{DI}$	20.43	10.46
$LSTM_{AR}$	65.22	27.77
CNN-$LSTM_{DI}$	51.61	19.51
CNN-$LSTM_{AR}$	71.47	23.37

the underlying trends relevant to our prediction timescale. Then, Fast Fourier Transform (FFT) was carried out was carried out on the demeaned data to decompose the time series and identify key periodic trends of variation of the shoreline position at timescales of months-to-annual and beyond (Table 4, Fig. 7). From the identified dominant frequencies, two prominent bi-annual (26 weeks), annual (52 weeks) trends of variability of both shoreline position and wave climate within the timescale of our interest were selected as they would represent short-medium term (inter-seasonal and annual) shoreline variabilities. The time series of wave forcing and shoreline position data were decomposed into 26- and 52-week trends, seasonalities and remaining residuals.

Then, the residual component (which we represent here as the short-term fluctuation) at each period was reduced by applying a set of selected noise reduction factors, NF (0, 0.25, 0.5, 0.75) to simulate equally distributed noise level scenarios. NF ranges between 0 and 1, where 0 represents full fluctuation removal and, 1 retains the original short term fluctuation. The chosen values were selected to represent equally spaced increments across this range, allowing for a structured evaluation of model performance under progressively reduced noise conditions. Following that, the time series were reconstructed by combining the noise reduced residuals with the primary trend components, ensuring the trends were retained at the different level noised reduced data. The summary of the model variants built based on all data noise reduction methods are given in Table 5. Two measures were used to compare the noise in each scenario, compared to the original shoreline time series data. The 'variance ratio' was calculated as the ratio between variance of the smoothed time series data and the original weekly time series data. The 'Noise ratio' was calculated as the ratio between variance of the smoothed time series data and the variance of noise, where noise is calculated as the difference between original time series and smoothed time series at each time step. The input data distribution in each variant for all input parameters are given in the box-plots (for training and testing sets separately) in Fig. 8. It can be observed how the data noise is lowered in noise reduced scenarios, when compared with the baseline scenario.

The performance of '$LSTM_{AR}$' model, which is selected for our application, was evaluated against different forcing data noise levels (Table 5) by comparing shoreline position predictions against the testing dataset (Fig. 9). The models were trained and cross-validated, and the 'best optimised' model from the HPO optimisation and the 'mean ensemble' prediction were selected following the same methods described earlier in this section.

Table 4
Results of FFT decomposition of the shoreline time series.

Periods correspond to dominant frequencies (weeks)	Approx. in Time	Interpretation
417.75	8 years	Long-term trend
238.71	4.5 years	Multi-year periodicity, possibly related to climatic variabilities in the North Pacific Ocean
52.22	1 year	Annual cycle (most dominant)
26.11	6 months	Biannual/semi-annual cycle

The trend of variability of the predicted shoreline positions from all noise-reduced model variants are in general in good agreement (overall average of 'best optimised' model, NRMSE = 0.21 m and 'mean ensemble', NRMSE = 0.215 m) with the 'measured' (before smoothing) shoreline position trends, as observed in Fig. 9. This indicates that reasonable predictions of trends of variability of the shoreline position can be obtained from smoothed data, if smoothing is carefully done to retain the trends of variability of the shoreline position at the required timescale. The performance matrix of all models compared with smoothed shoreline positions is shown in Table 6, and in visual form with radar plots in Fig. 10. The shoreline predictions from both the 'best optimised' model and the 'mean ensemble' of all noise-reduced model variants give lower RMSE values (less than 9.74 m and 11.02 m respectively) when compared to the baseline model prediction with original data ('best optimised' models RMSE of 12.04 m and, 'mean ensemble' RMSE of 12.84 m). They tend to improve in all MI, r, and NRMSE values as the noise level (Noise ratio) reduces, except for few scenarios.

The model with 52-week seasonal decomposition and 0.25 NF (SD52_NF = 0.25) gives the best overall performance for both 'best optimised' model prediction and the 'mean ensemble' prediction, with the lowest RMSEs of 3.41 m (NRMSE = 0.09 m) and 3.57 m (NRMSE = 0.09 m), respectively among all scenarios. In general, the model variants with lower NF values (lower residual noise level included), gave better predictions, with only notable exception being the SD52_NF = 0 model which performs worse than SD52_NF = 0_25 scenario. This can also be visually observed in Fig. 10 in which the SD52_NF = 0.25 model scenario falls entirely within the 3rd quarter of the radar plot and the majority of the model scenarios with lower NF values also falls within the 3rd quarter. The $LSTM_{AR}$ MMA model (grey) which was used as the benchmark for noise reduction (which has a noise ratio similar to the noise ratios of $LSTM_{AR}$ SD26_NF = 0.50 and $LSTM_{AR}$ SD52_NF = 0.5) shows poor performance, compared to the similar noise ratio input models and mostly falls in the 1st and 4th quarters. All model scenarios have better r and MI values (expect for the $LSTM_{AR}$ SD52_NF = 0.75, 'mean ensemble' which, produces a marginally lower MI value) than the baseline $LSTM_{AR}$ model, indicating that agreement between model predictions and measured data, has improved with noise reduction. (Fig. 10(a) and (b)). Another notable fact is that the best optimised models perform better than the 'mean ensembles' when smoothed data were used. This may be due to the increased ability of optimisation to find an 'optimum model' with less noisy and trend-revelling data. Mean ensembles which capture the diverse trends may have produced slightly diverse trend variabilities.

A quantitative analysis on the improvement of the model performance as a result of the level of short term fluctuations smoothed in the shoreline position data series used is given in the Annex 02. The performance in terms of the normalised RMSE and Mielke's Index against the level of short term fluctuations (given in Table 5), is used in this analysis. The percentage improvement (compared to the baseline $LSTM_{AR}$) of the selected performance indicators (Normalised RMSE and Mielke's Index) from the baseline $LSTM_{AR}$ model is given in the Annex 02, Table A5.1. It can be observed that $LSTM_{AR}$ *MMA* slightly deviates from the gradual performance improvement trends, even with a significantly low 'Noise ratios', but with a high 'variance ratios. The moving averaging could not retain the underlying trends or remove the desired short-term fluctuations appropriate for the focused timescales of the predictions. Therefore, the model training and prediction may still be impacted either by a degree of 'retaining level of fluctuations' or by 'the loss of important trends', compared to other scenarios. This result shows that certain desirable predictability improvement can be achieved with some degree of input data series smoothing, without compromising the underlying trends of variabilities at the desired timescales.

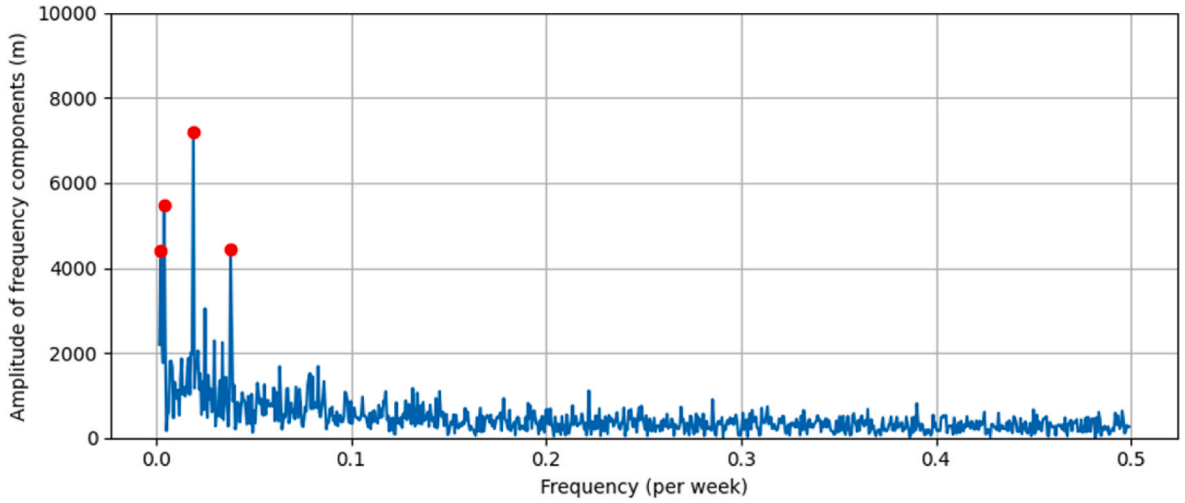


Fig. 7. Frequency spectrum of shoreline position variability after FFT (dominant frequencies are highlighted in red). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 5
Model scenarios developed for model performance assessment on different levels of data noise.

Model	Dataset	Noise reduction	Seasonality (for SD)	Variance ratio	Noise ratio
LSTM $_{AR}$ (baseline)	Weekly	–	–	1	–
LSTM $_{AR}$ MMA	Weekly	average	–	0.84	4.26
LSTM $_{AR}$ SD26_NF = 0.75	Weekly	Residuals 0.25	26 weeks	0.84	31.06
LSTM $_{AR}$ SD26_NF = 0.50	Weekly	Residuals 0.5	26 weeks	0.72	8.14
LSTM $_{AR}$ SD26_NF = 0.25	Weekly	Residuals 0.75	26 weeks	0.64	3.35
LSTM $_{AR}$ SD26_NF = 0	Weekly	Residuals 1.0	26 weeks	0.6	1.8
LSTM $_{AR}$ SD52_NF = 0.75	Weekly	Residuals 0.25	52 weeks	0.82	23.67
LSTM $_{AR}$ SD52_NF = 0.50	Weekly	Residuals 0.5	52 weeks	0.7	6.62
LSTM $_{AR}$ SD52_NF = 0.25	Weekly	Residuals 0.75	52 weeks	0.63	2.79
LSTM $_{AR}$ SD52_NF = 0	Weekly	Residuals 1.0	52 weeks	0.6	1.54

4.2.2. Impact of data resolution and data availability on the shoreline prediction

The availability of high-frequency coastal change data over long time periods is scarce. Only a few sites around the world currently possess such datasets (Banno et al., 2020). Although some remotely sensed data, such as the data collected from satellites, may address the issue of time resolution, often their spatial resolution may not be adequate to accurately resolve important coastal morphodynamic features (Vitousek et al., 2023). Coastal monitoring is an expensive and time-consuming task. Therefore, it will be useful to explore the desirable measuring frequencies and data lengths that will allow meaningful and decision-supportive predictions from data-driven techniques so that monitoring efforts can be optimised.

The impact of data resolution on predicting shoreline change is investigated here by utilising a simple data reduction technique by selecting data at pre-selected intervals. The three data reduction scenarios compared to the original weekly data resolution scenario (baseline model, $LSTM_{AR}$) were summarised below in Table 8. The scenarios were derived by training and testing the '$LSTM_{AR}$ ' model using fortnightly (every two weeks), monthly (every 4 weeks), and inter-seasonal (twice in every season, every 6 weeks), wave forcing and shoreline position datasets. To represent real time monitoring, the data were temporally downsampled with measured values at each resolution, rather than averaging. Their predictive performance was compared with that of the model built on weekly shoreline data ($LSTM_{AR}$ – baseline model). It is to be noted that the lowering of the data resolution in-turn reduces the data length available for model training and prediction, and this is given in Table 7.

A comparison of shoreline predictions from all four original and reduced data scenarios (given in Table 7) for both 'best optimised' model and the 'mean ensemble' prediction, along with a comparison with measure data is shown in Fig. 11. The predictions from the 'best optimised' models of all scenarios show worse performance compared to baseline '$LSTM_{AR}$ ' model at weekly resolution. The predictions from the 'mean ensemble' predictions based on the fortnightly data closely follow the trend of variability of the shoreline position and show marginally better performance (NRMSE of 0.218 m) compared to the baseline model (NRMSE of 0.227 m). The 'mean ensemble' model predictions with both monthly and bi-seasonal data do not follow the trend of variability of the measured shoreline position and worse than the baseline model prediction.

The 'mean ensemble' produces slightly better performance in terms of capturing trend variability, though it unable to capture shoreline accretions accurately. It appears that ensemble balances the lower data resolution limitation by several candidates capturing and aggregating multiple learnt trends, contributing to producing a better mean prediction. This result reflects the fact that, the predictive capacity of LSTM model (even with an auto-regressive modelling approach) reduces with reduction in input data frequency. Furthermore, having a coarser data resolution may have omitted important shoreline variations such as storm induced erosion, which make model unable to capture and learn these variations through the beach memory state. The results reveal that fortnightly shoreline measurements may be adequate to predict the overall trend of variability of the shoreline change at timescales of seasons and above up to few years, at this beach, though it is highly a site-specific result outcome at this Hasaki beach.

Banno et al. (2020) highlighted that the Hasaki beach morphology demonstrates both long-term and short-term variations, but these

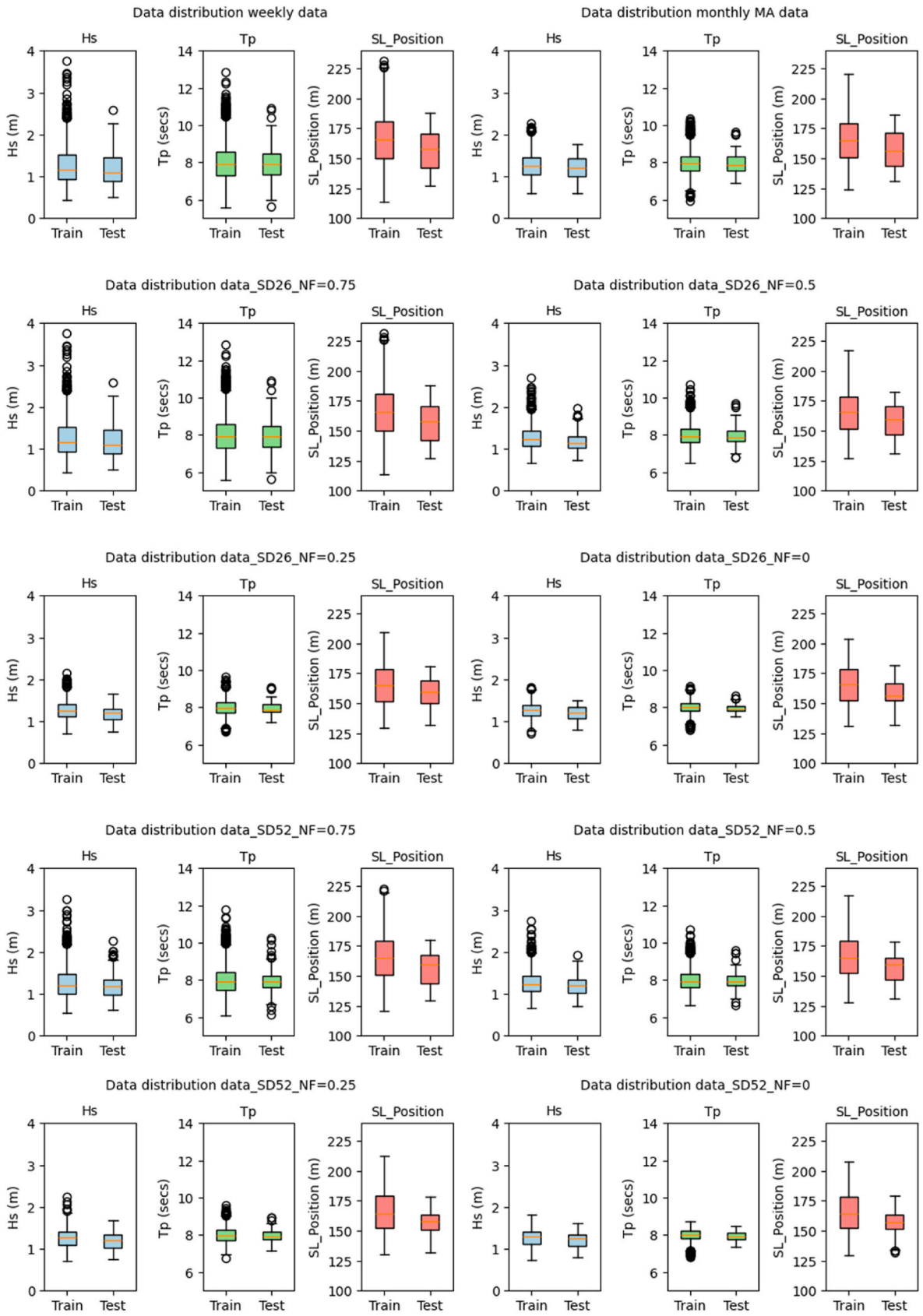


Fig. 8. Data distribution box plots for all scenarios (SD – Seasonal Decomposition, NF – Noise factor considered). Red lines = median, box edges top and bottom = Q3 and Q1 quartiles, whiskers at top and bottom = Q3 + 1.5 * IQR and Q1 – 1.5 * IQR, where IQR is the interquartile range (Q3-Q1), black circles represent outliers, beyond the whiskers, as per standard notation. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

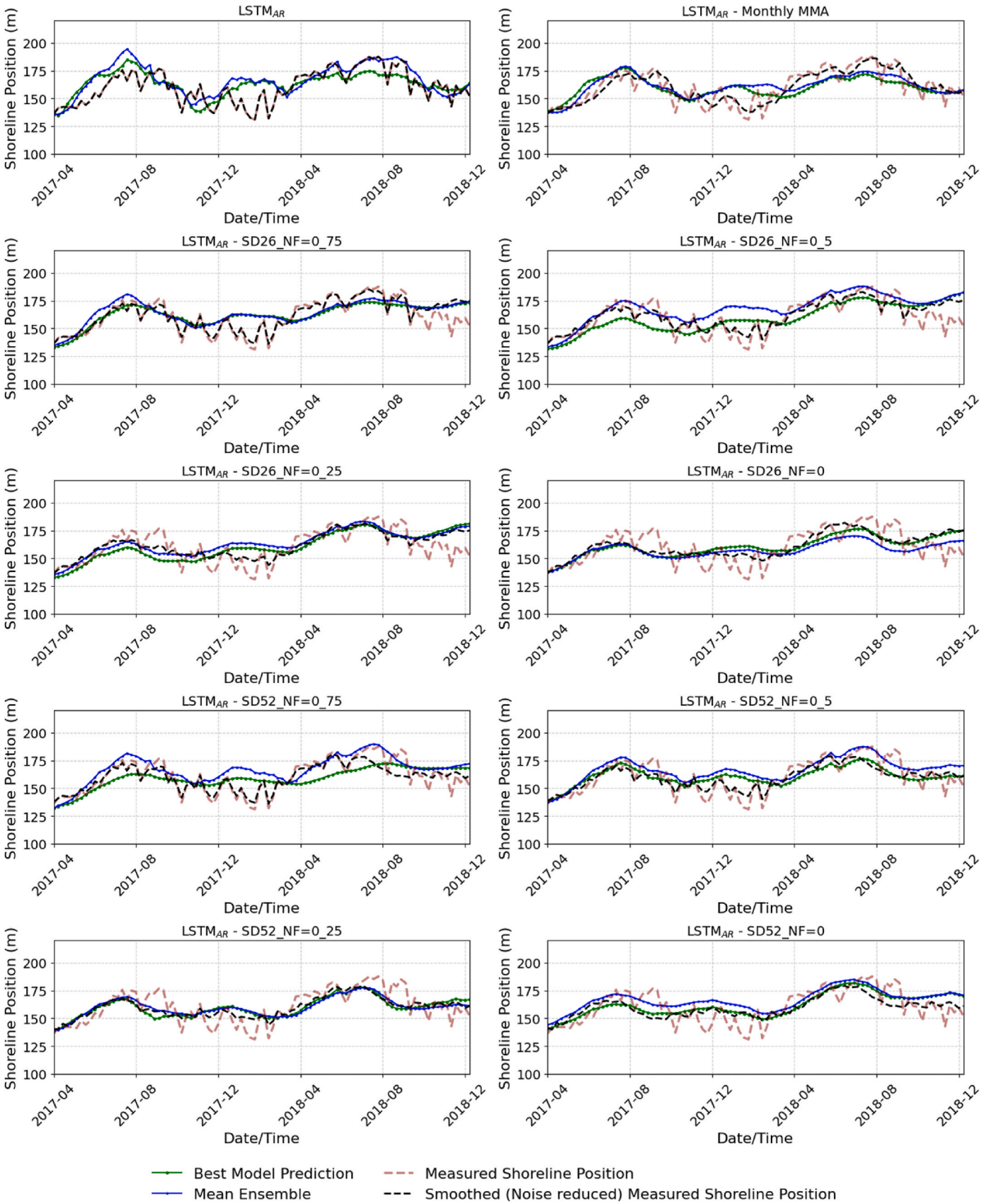


Fig. 9. (a) 'Best' optimised model and (b) 'mean ensemble' prediction forecasted shoreline position for all model scenarios with different noise levels.

Table 6
The performance matrix for the models for performance assessment on data noise assessment scenarios.

Model	Optimised HP model				Mean Ensemble model			
	RMSE (m)	r	MI	NRMSE (m)	RMSE (m)	r	MI	NRMSE (m)
$LSTM_{AR}$ (baseline)	12.04	0.61	0.57	0.21	12.84	0.69	0.63	0.23
$LSTM_{AR}$ MMA	9.74	0.75	0.68	0.20	9.08	0.75	0.70	0.18
$LSTM_{AR}$ SD26_NF = 0.75	8.74	0.76	0.72	0.18	8.87	0.76	0.73	0.18
$LSTM_{AR}$ SD26_NF = 0.50	7.60	0.83	0.82	0.16	9.81	0.82	0.72	0.21
$LSTM_{AR}$ SD26_NF = 0.25	5.98	0.88	0.87	0.13	5.62	0.89	0.87	0.13
$LSTM_{AR}$ SD26_NF = 0	4.22	0.92	0.91	0.09	6.31	0.91	0.77	0.14
$LSTM_{AR}$ SD52_NF = 0.75	9.14	0.61	0.60	0.21	11.02	0.74	0.62	0.25
$LSTM_{AR}$ SD52_NF = 0.50	5.77	0.82	0.79	0.15	9.06	0.84	0.68	0.23
$LSTM_{AR}$ SD52_NF = 0.25	3.41	0.93	0.92	0.09	3.57	0.92	0.91	0.09
$LSTM_{AR}$ SD52_NF = 0	4.66	0.92	0.88	0.12	7.60	0.94	0.71	0.19

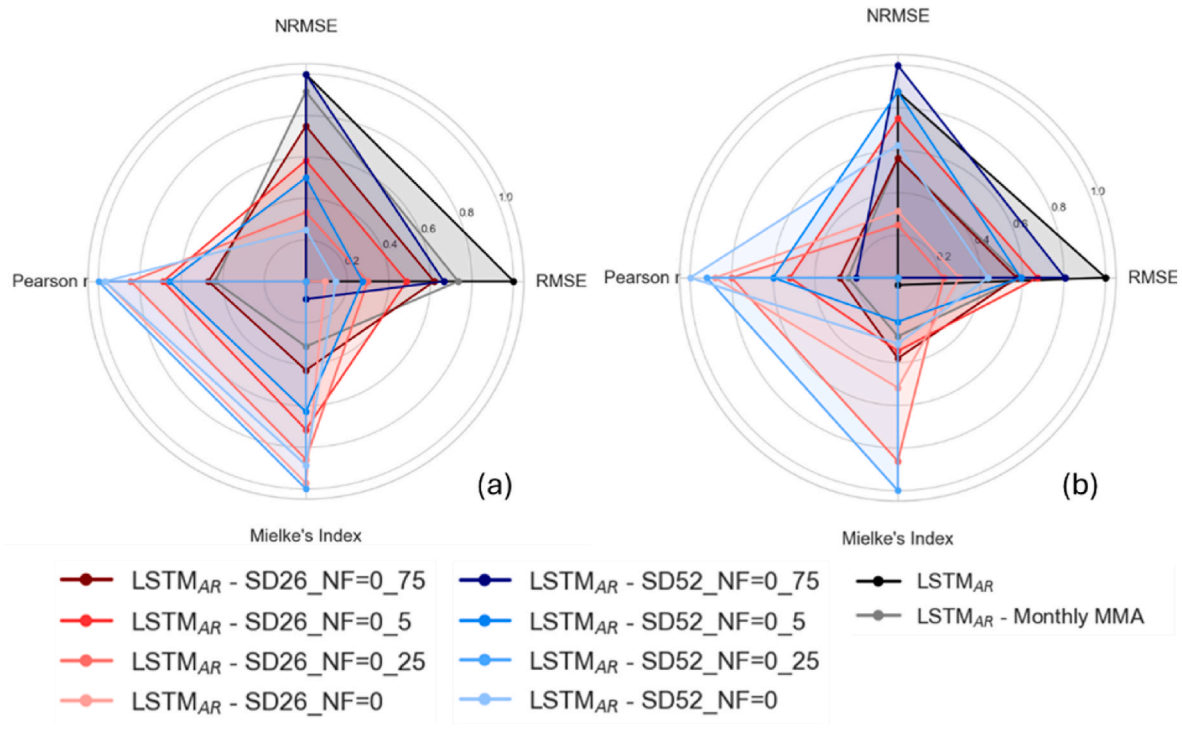


Fig. 10. Radar plots of performance matrix (normalised) (a) for Optimised models, (b) 'mean ensemble' predictions for all model scenarios for performance assessment on data noise reduction level.

changes are mostly associated with physical processes of bar behaviour and berm evolution, which are cyclical and seasonal rather than week-to-week stochastic shifts. This study further shows that while monthly variability exists, it is not abrupt or random in nature. The bar movements, for instance, were shown to evolve gradually over multi-month to multi-year cycles, consistent with previous conceptual models of bar dynamics (Ruessink and Kroon, 1994; Kuriyama et al., 2002). Using the monthly standard deviation of profile elevations, mostly in the bar regions, it has shown substantial morphological adjustments occur within monthly windows. To capture these monthly to yearly trends of the driving processes, weekly resolution input data may not necessarily yield additional critical insight, especially under non-extreme conditions. Therefore, as the results revealed above, a two-week temporal resolution provides a sufficient and fair compromise between having coarser resolution (fortnightly, compared to weekly) and still capturing the important physical processes accurately.

5. Concluding remarks

In this study, we demonstrate the applicability of LSTM-based deep

Table 7
Model scenarios for data resolution impact assessment.

Model	Inputs	Predictive output	Dataset	Amount of training data	Amount of data for prediction
$LSTM_{AR}$ (baseline model)	Hs(t-k: t), Tp(t-k: t), SLP(t-1)	SLP(t)	weekly	1 555	92
$LSTM_{AR}$ fortnightly	same	same	fortnightly	778	46
$LSTM_{AR}$ monthly	same	same	monthly	357	20
$LSTM_{AR}$ inter seasonally	same	same	inter-seasonally	259	16

Table 8
The performance matrix for the models for performance on input forcing data resolution assessment scenarios.

Model	Optimised HP model				Mean of Ensemble models			
	RMSE (m)	r	Mielke's Index	NRMSE (m)	RMSE (m)	r	Mielke's Index	NRMSE (m)
LSTM $_{AR}$	12.04	0.61	0.57	0.21	12.84	0.69	0.63	0.23
LSTM $_{AR}$ Fortnightly	15.05	0.68	0.47	0.27	12.12	0.72	0.63	0.22
LSTM $_{AR}$ Monthly	15.52	0.32	0.25	0.29	15.19	0.41	0.29	0.28
LSTM $_{AR}$ Semi Seasonally	15.30	0.70	0.42	0.28	13.98	0.69	0.47	0.25

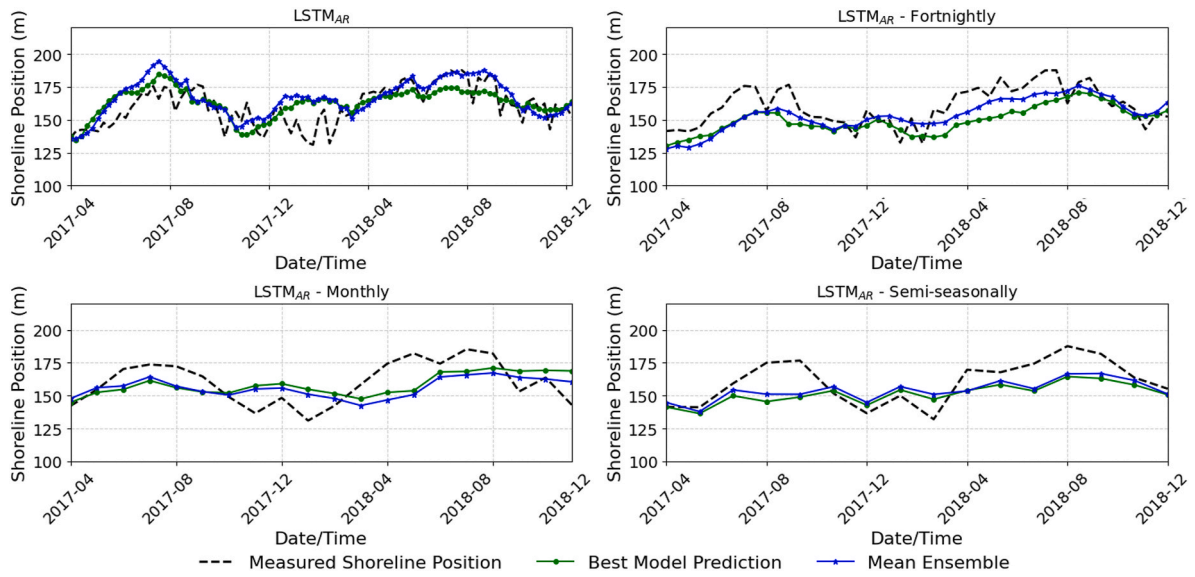


Fig. 11. 'Best optimised' model prediction, 'mean ensemble' prediction and 'ensemble' prediction envelope for all model scenarios in Table 8 for data resolution impact assessment. (a) baseline (weekly resolution) (b) fortnightly resolution (c) monthly resolution (d) semi-seasonal (6 weeks) resolution, using the LSTM$_{AR}$ model.

learning techniques to predict time-dependent shoreline position change of a wave dominated, micro-tidal sandy beach. This exploratory application of two LSTM neural networks (LSTM and CNN-LSTM) under two modelling approaches (direct driving force inputs (DI) and auto-regressive (AR) shoreline input in addition to driving force inputs) as predictive models show that LSTMs can be a promising tool for predicting time-dependent trends of variability of the shoreline position at this beach at the timescales of months to a few years. However, their success as a predictive tool is governed by the availability of an adequate amount of high-quality input data. Also we note that considering the extreme reliance of the modelling approach on data, testing the method on different sites may be necessary to reinforce its efficacy under varying site conditions and beach types.

Although all four model types were able to capture the trends of time variability of the shoreline position at the Hasaki beach, the LSTM$_{AR}$ and CNN − LSTM$_{AR}$ models, which incorporate the effect of 'antecedent beach state' in predicting the shoreline evolution are found to be better than the direct input models. This signifies the importance of the antecedent beach state on future change of the shoreline position, despite the fact that, the auto-regressive prediction approach accumulates error, when progressing ahead with time.

The mean ensemble shoreline position prediction from a selected ensemble of LSTM trial models proved beneficial across certain application scenarios. In certain scenarios among the assessments carried out, the mean ensemble prediction performed better than the 'optimised model' according to the performance matrix involving measures of RMSE, NRMSE, r, MPE, and MI. Unlike a single optimised model, the ensemble captures a range of learned temporal patterns through the ensemble members, offering more diverse, robust and generalised predictions. This has the potential to eliminate overfitting and bias to the selected validation dataset despite the techniques used to minimise them. It was observed that the observed shoreline positions largely fall within the ensemble prediction envelope giving a higher ECP values. Since the ensemble prediction uses existing models from HPO, it requires a minimal extra computational effort while providing additional insights into model variability and uncertainty.

The collection of beach change data at high frequency is expensive and time consuming. Therefore, the impact of input data resolution on the LSTM$_{AR}$ model's prediction was investigated to explore if measuring shoreline positions at lower resolutions would be adequate to determine reasonable predictions. It was found that fortnightly data can provide acceptable shoreline position predictions, and the prediction performance are comparable to those using weekly data. However, data resolutions lower than fortnightly led to poor model performance. Therefore, it can be argued that beach surveys at least at fortnightly intervals are required to drive LSTM modelling approach to perform with reasonably acceptable and reliable future shoreline position predictions at this site.

The model training is found to be sensitive to high short-term fluctuations, which we defined here as data noise. If the focus is on trends of variability of the shoreline position beyond monthly timescales, data smoothing and residual noise reduction are found be effective and useful to overcome this issue. Fourier decomposition of the time dependent shoreline position change data indicated trend of variability at two primary time periods: 26 weeks (bi-annual) and 52 weeks (annual), and two longer time cycles (4.5 years and 8 years). The longer-term cycles (4.5–8 years) likely reflect multi-year trends rather than true seasonal variation and incorporating them into the decomposition would lead to excessive smoothing of the time series, potentially suppressing the short- to medium-term dynamics of interest in this study. Filtering shoreline position fluctuation residuals around the trends at those two seasonal to annual timescales by applying a user defined noise reduction factor,

allowed the model to learn the trend and seasonality better, resulting better model performance. Selecting these two cycles we aimed to isolate the recurrent short term (seasonal to annual) behaviour while preserving the fidelity of the shorter-term trend variations which is critical for model learning and performance assessment. When used moving averaged input forcing data (which is another commonly used approach for data noise reduction), the model performed worse than that with residual data reduction although the results are better comparted to that from the original model. However, the applicability of those techniques depends on the purpose of the predictions and the timescale of interest and may lead to compromising certain short term (storm scale) fluctuations.

Our results on this application show promising evidence that LSTM DL approaches are a successful tool for predicting shoreline change at monthly to inter-annual timescale with reasonable accuracy and accurately captures shoreline position trend variabilities, which is important for coastal management decision making. The method, including pre-processing of the input data are directly transferable to any other site with suitable datasets although pre-processing needs may vary with the available data and the site-specific characteristics. Our analysis also shows that very high frequency field measurements may not always be necessary for making reliable predictions of beach change beyond monthly timescales.

The significantly less computing resource requirements of this method compared to the traditional process-based modelling offers opportunities to quantify the uncertainties inherent to morphodynamic predictions, which will be very useful for coastal management and engineering design. Also, longer term predictions can be attempted and justified when uncertainty of the predictions can be properly quantified. The method also offers opportunities to investigate the impacts of other potential drivers of coastal morphodynamic change, such as tidal fluctuations and mean sea level change, which can be input as forcing parameters to the model, when those data are available at similar temporal resolutions.

CRediT authorship contribution statement

Tharindu Manamperi: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis. **Alma Rahat:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Doug Pender:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization. **Demetra Cristaudo:** Writing – review & editing, Supervision, Resources. **Rob Lamb:** Writing – review & editing, Supervision, Resources. **Harshinie Karunarathna:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Open research

Data (Including the codes and results) archiving is underway and can be made available upon request.

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.

Acknowledgement

The Authors would like to acknowledge the JBA Trust (project No. W22-1128), UK and the Engineering and Physical Sciences Research Council (EPSRC) - Doctoral Training Partnerships (DTP) (EP/W524694/1) grant for funding the PhD study of the first author. Further, would like to acknowledge the Port and Airport Research Institute (PARI), Japan, for providing the beach monitoring data and site knowledge for Hasaki beach, through the research collaboration with the Swansea University. A special acknowledgement to Dr. Masayuki Banno, the head of coastal sediment research group, PARI and, Dr. Yoshiyaki Kuriyama, Coastal Development Institute of Technology, Japan.

Annexes

Annex 01:

1. Schematic diagram of LSTM gated structure with the mechanism of temporal state updating is presented.

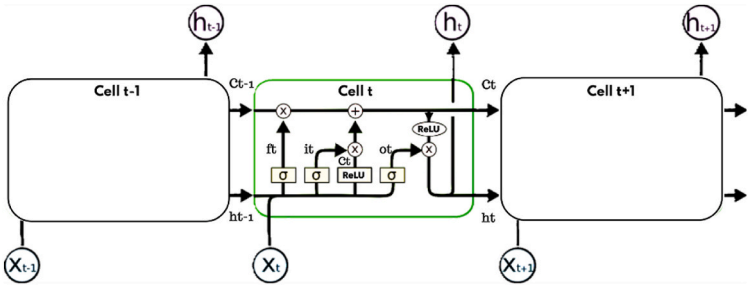


Fig. 1. LSTM gated cell structure. The temporal state update is represented.

At the microstructure level, LSTM cell is a particular gated cell architecture.

They are equipped with a 'forget gate (ft)', an 'input gate (it)', an 'output gate (ot)' and an internal cell state (ct) which represents the memory of the network and carries information across time steps.

The cell state (ct-1) is updated by the forget gate and the input gate, combined with the candidate cell state (ct). The updated cell state is then used as the input cell state in the next time step. The forget gate decides what portion of the cell state from the previous time step should be carried forward to the next time step. It uses a sigmoid activation function to produce values between 0 and 1, for each number in the cell state ct-1.

The standard formulae for each gate operation and candidate cell state is as follows (Hochreiter and Schmidhuber, 1997b).

$$ft = \sigma(Wf \cdot [ht-1, xt] + bf) \quad \text{eq. 1}$$

$$it = \sigma(Wi \cdot [ht-1, xt] + bi) \quad \text{eq. 2}$$

$$\widetilde{Ct} = ReLU(Wc \cdot [ht-1, xt] + bc) \quad \text{eq. 3}$$

$$Ct = ft \cdot Ct - 1 + it \cdot \widetilde{Ct} \quad \text{eq. 4}$$

$$ot = \sigma(Wo \cdot [ht-1, xt] + bo) \quad \text{eq. 5}$$

$$ht = ot \cdot tanh(Ct) \quad \text{eq. 6}$$

where, ft is the forget gate activation vector at time step t, $i\,t$ is the input gate activation vector at time step t, $o\,t$ is the output gate activation vector at time step t, σ is the Sigmoid activation function, ReLu is the Rectified Linear Unit activation function, Wf, Wi, Wc, Wo are weight matrices for the forget gate, input gate, candidate cell state and, output gate, bf, bi, bc, bo are bias vectors for the forget gate, input gate, candidate cell state and, output gate, $ht-1, ht$ are hidden states at previous (t-1) and current (t) time steps, $Ct-1, Ct$ are Cell state vectors at previous (t-1) and current (t) time steps, $\widetilde{Ct}$ is the candidate cell state. The xt are input feature vectors, fed to the network as time-sequential data (at each time step t).

2. Schematic diagram of CNN layers stacked LSTM structure with the mechanism of data flow is presented.

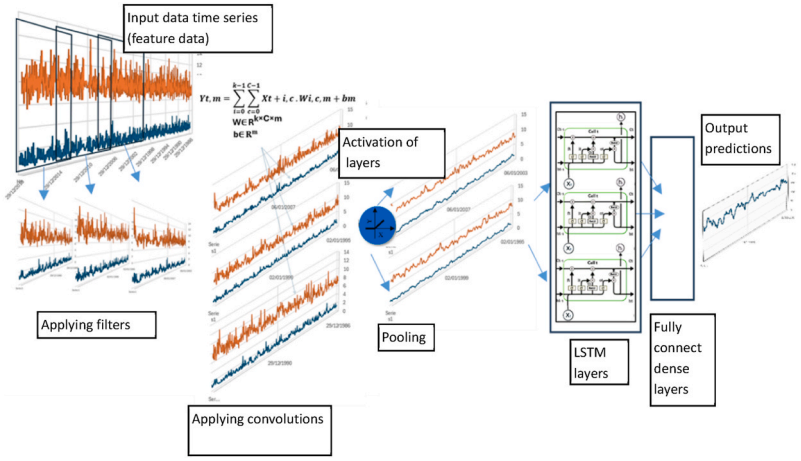


Fig. 2. Stacked CNN and LSTM layers and flow of feature data through layers.

Convolutional Neural Networks (CNN) are a specific structural arrangement of application of convolutional operations over an input data series to capture underlying features.

The time-sequential data input to the CNN layers is considered as 2-dimensional sequences $X \in R^{T \times C}$, T is the number of time steps, and C is the number of input features. The convolutional operation is applied on this 2-dimensional sequences according to the generalised equation as below.

$$Yt, m = \sum_{i=0}^{k-1} \sum_{c=0}^{C-1} Xt + i, c \,.Wi, c, m + bm \quad \text{eq. 7}$$

Where, Yt, m is the output value at time step t for the m th filter. $Xt + i, c$ is the input value at time step $t + i$ for the c th feature. Wi, c, m is the weight of the kernel at position i for the c th input feature and m th filter and bm is the bias term for the m th filter. This summation runs over the kernel size k and the number of input features C.

3. Schematic diagrams illustrating data arrangement for (i) direct input predictive approach, (ii) Auto-regressive predictive approach

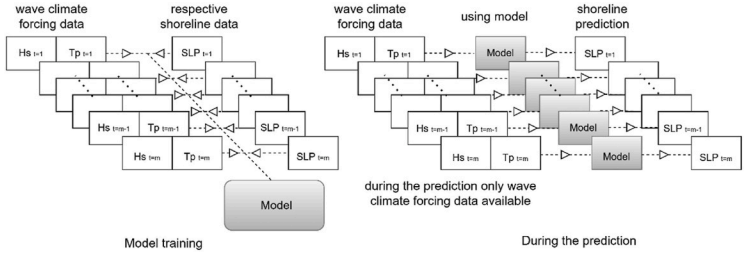


Fig. 3. Data structure during (a) model training and (b) model prediction for (i) direct input prediction approach

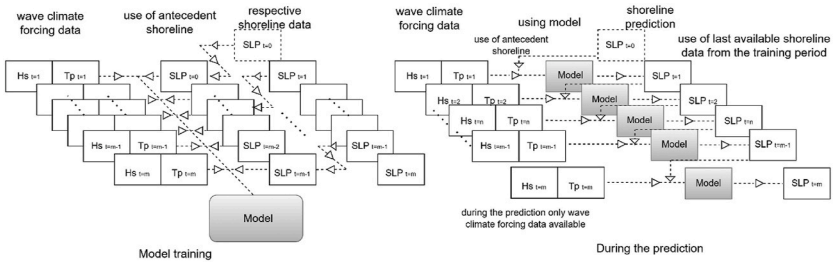


Fig. 4. Data structure during (a) model training and (b) model prediction for (i) Auto-regressive prediction approach

Annex

1. Quantitative improvement of the model performance with input data smoothening

Below plots visually demonstrates the quantitative comparisons of the performance measures vs Noise and variance ratios for the different scenarios used in data noise impact assessment given in section 4.2, Table 5. Here we have selected the Normalised RMSE and Mielke's Index only, as the performance indicators, as this is for demonstration purposes only.

1.1. Normalised RMSE vs Noise ratio for scenarios used in seasonal decomposition-based noise smoothing assessment

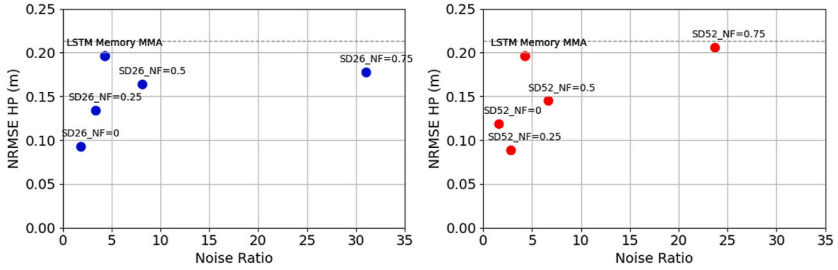


Fig. 5. (a) SD26 scenarios (b) SD52 scenarios in comparison with LSTM AR MMA model. Note: The NRMSE values of the baseline LSTM AR model is marked with the dashed line as the Noise ratio is not defined for this baseline model. All comparisons are done for the Optimised HP models and denoted as HP.

1.2. Normalised RMSE vs Variance ratio for scenarios used in seasonal decomposition-based noise smoothing assessment

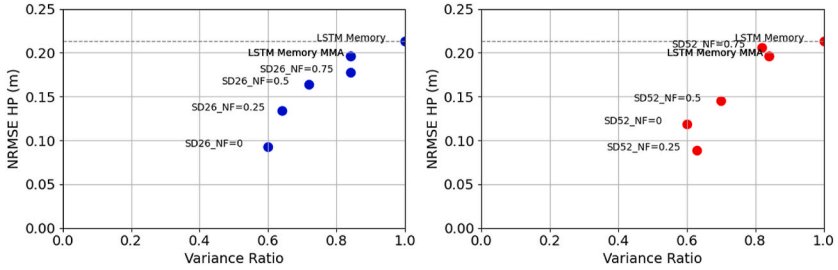


Fig. 6. (a) SD26 scenarios (b) SD52 scenarios in comparison with LSTM AR MMA model and LSTM AR (baseline) model. Note: The NRMSE values of the baseline LSTM AR model is marked normally as the Variance ratio is defined and valued as 1, for this baseline model. All comparisons are done for the Optimised HP models and denoted as HP.

1.3. Mielke's Index vs Noise ratio for scenarios used in seasonal decomposition-based noise smoothing assessment

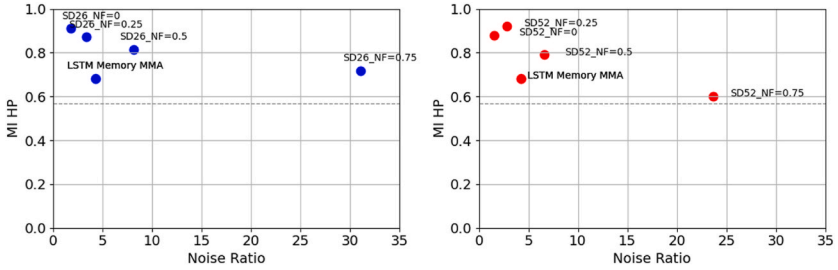


Fig. 7. (a) SD26 scenarios (b) SD52 scenarios in comparison with LSTM AR MMA model. Note: The NRMSE values of the baseline LSTM AR model is marked with the dashed line as the Noise ratio is not defined for this baseline model. All comparisons are done for the Optimised HP models and denoted as HP.

1.4. Mielke's Index vs Variance ratio for scenarios used in seasonal decomposition-based noise smoothing assessment

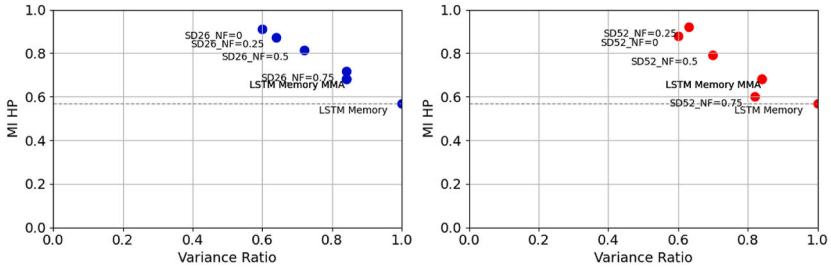


Fig. 8. (a) SD26 scenarios (b) SD52 scenarios in comparison with LSTM AR MMA model and LSTM AR (baseline) model. Note: The NRMSE values of the baseline LSTM AR model is marked normally as the Variance ratio is defined and valued as 1, for this baseline model. All comparisons are done for the Optimised HP models and denoted as HP.

1.5. Percentage improvement of the selected performance indicators with Variance ratios (VR) and Noise ratios (NR) for scenarios used in seasonal decomposition-based noise smoothing assessment

The percentage improvement of the selected performance indicators (Normalised RMSE and Mielke's Index) from the baseline $LSTM_{AR}$ model is given in Table A5.1. The percentage improvement is given compared to the baseline $LSTM_{AR}$ scenario for MI and NRMSE indicators.

Table A5.1
Performance of the noise assessment scenarios along with the Noise and variance ratios of the respective input smoothed shoreline data at noise levels used in Table 5. The percentage improvement is given compared to the baseline $LSTM_{AR}$ scenario for MI and NRMSE indicators

Model	VR	NR	Optimised HP model					Mean Ensemble model				
			RMSE (m)	MI	% improve MI	NRMSE (m)	% improve NRMSE	RMSE (m)	MI	% improve MI	NRMSE (m)	% improve NRMSE
LSTM AR Memory	1.00	–	12.04	0.57	–	0.21	–	12.84	0.63	–	0.23	–
LSTM AR Memory MMA	0.84	4.26	9.74	0.68	17.0	0.20	8.0	9.08	0.70	9.6	0.18	19.4
SD26_NF = 0.75	0.84	31.06	8.74	0.72	21.0	0.18	16.4	8.87	0.73	13.4	0.18	20.7
SD26_NF = 0.5	0.72	8.14	7.60	0.82	30.4	0.16	23.0	9.81	0.72	12.2	0.21	6.6
SD26_NF = 0.25	0.64	3.35	5.98	0.87	34.9	0.13	37.1	5.62	0.87	27.8	0.13	44.5
SD26_NF = 0	0.60	1.80	4.22	0.91	37.7	0.09	56.3	6.31	0.77	18.3	0.14	38.8
SD52_NF = 0.75	0.82	23.67	9.14	0.60	5.5	0.21	3.3	11.02	0.62	−1.6	0.25	−9.7
SD52_NF = 0.5	0.70	6.62	5.77	0.79	28.2	0.15	31.9	9.06	0.68	7.1	0.23	−0.4
SD52_NF = 0.25	0.63	2.79	3.41	0.92	38.4	0.09	58.2	3.57	0.91	31.1	0.09	58.6
SD52_NF = 0	0.60	1.54	4.66	0.88	35.6	0.12	44.1	7.60	0.71	11.8	0.19	14.5

Data availability

Data will be made available on request.

References

Accarino, G., Chiarelli, M., Fiore, S., Federico, I., Causio, S., Coppini, G., Aloisio, G., 2021. A multi-model architecture based on Long Short-Term Memory neural networks for multi-step sea level forecasting. Future Gener. Comput. Syst. 124, 1–9. https://doi.org/10.1016/j.future.2021.05.008.

Adeli, A., Dastgheib, A., Roelvink, D., 2025. Shoreline dynamics prediction using machine learning models: from process learning to probabilistic forecasting. Front. Mar. Sci. 12, 1562504. https://doi.org/10.3389/fmars.2025.1562504.

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: a next-generation hyperparameter optimization framework. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631. https://doi.org/10.1145/3292500.3330701.

Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network. 2017 International Conference on Engineering and Technology (ICET), pp. 1–6. https://doi.org/10.1109/ICEngTechnol.2017.8308186.

Balogun, A.-L., Adebisi, N., 2021. Sea level prediction using ARIMA, SVR and LSTM neural network: assessing the impact of ensemble Ocean-Atmospheric processes on models' accuracy. Geomat. Nat. Hazards Risk 12 (1), 653–674. https://doi.org/10.1080/19475705.2021.1887372.

Banno, M., Kuriyama, Y., 2014. Prediction OF future shoreline change with sea-level rise and wave climate change at HASAKI, Japan. Coastal Engineering Proceedings 1 (34), 56. https://doi.org/10.9753/icce.v34.sediment.56.

Banno, M., Nakamura, S., Kosako, T., Nakagawa, Y., Yanagishima, S., Kuriyama, Y., 2020. Long-Term observations of beach variability at hasaki, Japan. J. Mar. Sci. Eng. 8 (11), 871. https://doi.org/10.3390/jmse8110871.

Barnard, P.L., Hoover, D., Hubbard, D.M., Snyder, A., Ludka, B.C., Allan, J., Kaminsky, G. M., Ruggiero, P., Gallien, T.W., Gabel, L., McCandless, D., Weiner, H.M., Cohn, N., Anderson, D.L., Serafin, K.A., 2017. Extreme oceanographic forcing and coastal response due to the 2015–2016 El Niño. Nat. Commun. 8 (1), 14365. https://doi.org/10.1038/ncomms14365.

Benidis, K., Rangapuram, S.S., Flunkert, V., Wang, Y., Maddix, D., Turkmen, C., Gasthaus, J., Bohlke-Schneider, M., Salinas, D., Stella, L., Aubet, F.-X., Callot, L., Januschowski, T., 2022. Deep learning for time series forecasting: Tutorial and literature survey. ACM Comput. Surv. 55 (6), 1–121. https://doi.org/10.1145/3533382.

Benidis, K., Rangapuram, S.S., Flunkert, V., Wang, Y., Maddix, D., Turkmen, C., Gasthaus, J., Bohlke-Schneider, M., Salinas, D., Stella, L., Aubet, F.-X., Callot, L., Januschowski, T., 2023. Deep learning for time series Forecasting: tutorial and Literature Survey. ACM Comput. Surv. 55 (6), 1–36. https://doi.org/10.1145/3533382.

Bergmeir, C., Benítez, J.M., 2012. On the use of cross-validation for time series predictor evaluation. Inf. Sci. 191, 192–213. https://doi.org/10.1016/j.ins.2011.12.028.

Bergstra, J., Yamins, D., Cox, D.D., 2013. Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. Proceedings of the 30th International Conference on International Conference on Machine Learning 28. I-115-I–123.

Brommer, M.B., Bochev-van Der Burgh, L.M., 2009. Sustainable Coastal Zone management: a concept for forecasting long-term and large-scale Coastal evolution. J. Coast Res. 251, 181–188. https://doi.org/10.2112/07-0909.1.

Chen, X., Banno, M., Mori, N., 2025. Equilibrium shoreline change model with non-stationary free parameters at Hasaki Beach, Japan. Coastal Engineering 200, 104761. https://doi.org/10.1016/j.coastaleng.2025.104761.

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.

Davidson, M.A., Turner, I.L., 2009. A behavioral template beach profile model for predicting seasonal to interannual shoreline evolution. J. Geophys. Res.: Earth Surf. 114 (F1). https://doi.org/10.1029/2007JF000888.

Dodet, G., Castelle, B., Masselink, G., Scott, T., Davidson, M., Floc'h, F., Jackson, D., Suanez, S., 2019. Beach recovery from extreme storm activity during the 2013–14

winter along the Atlantic coast of Europe. Earth Surf. Process. Landf. 44 (1), 393–401. https://doi.org/10.1002/esp.4500.
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.
Eichentopf, S., Van Der Zanden, J., Cáceres, I., Baldock, T.E., Alsina, J.M., 2020. Influence of storm sequencing on breaker bar and shoreline evolution in large-scale experiments. Coast. Eng. 157, 103659. https://doi.org/10.1016/j.coastaleng.2020.103659.
Frame, J.M., Kratzert, F., Klotz, D., Gauch, M., Shalev, G., Gilon, O., Qualls, L.M., Gupta, H.V., Nearing, G.S., 2022. Deep learning rainfall–runoff predictions of extreme events. Hydrol. Earth Syst. Sci. 26 (13), 3377–3392. https://doi.org/10.5194/hess-26-3377-2022.
Frifra, A., Maanan, M., Maanan, M., Rhinane, H., 2024. Harnessing LSTM and XGBoost algorithms for storm prediction. Sci. Rep. 14 (1), 11381. https://doi.org/10.1038/s41598-024-62182-0.
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.
Gorriz, J.M., Clemente, R.M., Segovia, F., Ramirez, J., Ortiz, A., Suckling, J., 2024. Is K-fold cross validation the best model selection method for Machine Learning? arXiv: 2401.16407). arXiv. https://doi.org/10.48550/arXiv.2401.16407.
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.
Hashemi, M.R., Ghadampour, Z., Neill, S.P., 2010. Using an artificial neural network to model seasonal changes in beach profiles. Ocean. Eng. 37 (14–15), 1345–1356. https://doi.org/10.1016/j.oceaneng.2010.07.004.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.
Horrillo-Caraballo, J.M., Reeve, D.E., 2010. An investigation of the performance of a data-driven model on sand and shingle beaches. Mar. Geol. 274 (1–4), 120–134. https://doi.org/10.1016/j.margeo.2010.03.010. Scopus.
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., Splinter, K.D., Harley, M.D., Turner, I.L., 2022. Improving multi-decadal coastal shoreline change predictions by including model parameter non-stationarity. Front. Mar. Sci. 9, 1012041. https://doi.org/10.3389/fmars.2022.1012041.
Karunarathna, H., Brown, J., Chatzirodou, A., Dissanayake, P., Wisse, P., 2018. Multi-timescale morphological modelling of a dune-fronted sandy beach. Coast. Eng. 136, 161–171. https://doi.org/10.1016/j.coastaleng.2018.03.005.
Karunarathna, H., Horrillo-Caraballo, J., Reeve, D., 2012. Prediction of cross-shore beach profile evolution using a diffusion type model. Cont. Shelf Res. 48, 157–166. https://doi.org/10.1016/j.csr.2012.08.004.
Karunarathna, H., Reeve, D.E., 2013. A hybrid approach to model shoreline change at multiple timescales. Cont. Shelf Res. 66, 29–35. https://doi.org/10.1016/j.csr.2013.06.019.
Kim, T., Lee, W.-D., 2022. Review on applications of machine learning in coastal and Ocean engineering. Journal of Ocean Engineering and Technology 36 (3), 194–210. https://doi.org/10.26748/KSOE.2022.007.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90. https://doi.org/10.1145/3065386.
Kuriyama, Y., 2002. Medium-term bar behavior and associated sediment transport at Hasaki, Japan. J. Geophys. Res. 107 (C9), 3132. https://doi.org/10.1029/2001JC000899.
Lawal, Z.K., Yassin, H., Lai, D.T.C., Idris, A.C., 2023. Coastal wave modeling and forecasting with LSTM optimization for sustainable energy harvesting (2023040282). Preprints. https://doi.org/10.20944/preprints202304.0282.v2.
Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324. https://doi.org/10.1109/5.726791. Proceedings of the IEEE.
Lesser, G.R., Roelvink, J.A., van Kester, J.A.T.M., Stelling, G.S., 2004. Development and validation of a three-dimensional morphological model. Coast. Eng. 51 (8), 883–915. https://doi.org/10.1016/j.coastaleng.2004.07.014.
Lévesque, J.-C., Gagné, C., Sabourin, R., 2016. Bayesian hyperparameter optimization for ensemble learning. arXiv. https://doi.org/10.48550/arXiv.1605.06394 arXiv: 1605.06394).
Lindemann, B., Müller, T., Vietz, H., Jazdi, N., Weyrich, M., 2021. A survey on long short-term memory networks for time series prediction. Proced. CIRP 99, 650–655. https://doi.org/10.1016/j.procir.2021.03.088.
Mai, J., Shen, H., Tolson, B.A., Gaborit, É., Arsenault, R., Craig, J.R., Fortin, V., Fry, L.M., Gauch, M., Klotz, D., Kratzert, F., O'Brien, N., Princz, D.G., Rasiya Koya, S., Roy, T., Seglenieks, F., Shrestha, N.K., Temgoua, A.G.T., Vionnet, V., Waddell, J.W., 2022. The Great Lakes runoff intercomparison Project phase 4: the Great Lakes (GRIP-GL). Hydrol. Earth Syst. Sci. 26 (13), 3537–3572. https://doi.org/10.5194/hess-26-3537-2022.
Masselink, G., Castelle, B., Scott, T., Dodet, G., Suanez, S., Jackson, D., Floc'h, F., 2016. Extreme wave activity during 2013/2014 winter and morphological impacts along the Atlantic coast of Europe. Geophys. Res. Lett. 43 (5), 2135–2143. https://doi.org/10.1002/2015GL067492.
Mienye, I.D., Sun, Y., 2022. A Survey of ensemble learning: concepts, algorithms, applications, and prospects. IEEE Access 10, 99129–99149. https://doi.org/10.1109/ACCESS.2022.3207287.
Montaño, J., Coco, G., Cagigal, L., Mendez, F., Rueda, A., Bryan, K.R., Harley, M.D., 2021a. A multiscale approach to shoreline prediction. Geophys. Res. Lett. 48 (1). https://doi.org/10.1029/2020GL090587.
Montaño, J., Coco, G., Chataigner, T., Yates, M., Le Dantec, N., Suanez, S., Cagigal, L., Floc'h, F., Townend, I., 2021b. Time-Scales of a dune-beach System and implications for shoreline modeling. J. Geophys. Res.: Earth Surf. 126 (11), e2021JF006169. https://doi.org/10.1029/2021JF006169.
Pender, D., Karunarathna, H., 2013. A statistical-process based approach for modelling beach profile variability. Coast. Eng. 81, 19–29. https://doi.org/10.1016/j.coastaleng.2013.06.006.
Prince, S.J.D., 2023. Understanding Deep Learning. The MIT Press, Cambridge, MA. http://udlbook.com.
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.
Risha, M., Liu, P., 2025. Shoreline prediction models: a review of the evolution from empirical to AI machine learning approaches. https://doi.org/10.22541/essoar.174231482.27711801/v1.
Roelvink, D., Reniers, A., Van Dongeren, A., Van Thiel De Vries, J., McCall, R., Lescinski, J., 2009. Modelling storm impacts on beaches, dunes and barrier islands. Coast. Eng. 56 (11–12), 1133–1152. https://doi.org/10.1016/j.coastaleng.2009.08.006.
Roelvink, D., Walstra, Dirk-Jan R., van der Wegen, Mick, Ranasinghe, R., 2016. Modeling of coastal morphological processes. In: Springer Handbook of Ocean Engineering. Springer, Cham. https://doi.org/10.1007/978-3-319-16649-0_28.
Różyński, G., 2003. Data-driven modeling of multiple longshore bars and their interactions. Coast. Eng. 48 (3), 151–170. https://doi.org/10.1016/S0378-3839(03)00024-3.
Ruessink, B.G., Kroon, A., 1994. The behaviour of a multiple bar system in the nearshore zone of Terschelling, the Netherlands: 1965–1993. Marine Geol 121 (3–4), 187–197. https://doi.org/10.1016/0025-3227(94)90030-2.
Schall, R., 2012. The empirical coverage of confidence intervals: point estimates and confidence intervals for confidence levels. Biom. J. 54 (4), 537–551. https://doi.org/10.1002/bimj.201100134.
Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.K., Woo, W., 2015. Convolutional LSTM Network: a Machine Learning Approach for Precipitation Nowcasting.
Simmons, J.A., Harley, M.D., Marshall, L.A., Turner, I.L., Splinter, K.D., Cox, R.J., 2017. Calibrating and assessing uncertainty in coastal numerical models. Coast. Eng. 125, 28–41. https://doi.org/10.1016/j.coastaleng.2017.04.005.
Simmons, J.A., Splinter, K.D., 2025. Data-driven shoreline modelling at timescales of days to years. Coast. Eng. 197, 104685. https://doi.org/10.1016/j.coastaleng.2024.104685.
Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical bayesian optimization of machine learning algorithms. arXiv. https://doi.org/10.48550/arXiv.1206.2944 arXiv: 1206.2944).
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.
Tiggeloven, T., Couasnon, A., van Straaten, C., Muis, S., Ward, P.J., 2021. Exploring deep learning capabilities for surge predictions in coastal areas. Sci. Rep. 11 (1), 17224. https://doi.org/10.1038/s41598-021-96674-0.
Turner, I.L., Leaman, C.K., Harley, M.D., Thran, M.C., David, D.R., Splinter, K.D., Matheen, N., Hansen, J.E., Cuttler, M.V.W., Greenslade, D.J.M., Zieger, S., Lowe, R. J., 2024. A framework for national-scale coastal storm hazards early warning. Coast. Eng. 192, 104571. https://doi.org/10.1016/j.coastaleng.2024.104571.
Vitousek, S., Buscombe, D., Vos, K., Barnard, P.L., Ritchie, A.C., Warrick, J.A., 2023. The future of coastal monitoring through satellite remote sensing. Cambridge Prisms: Coastal Futures 1, e10. https://doi.org/10.1017/cft.2022.4.
Vos, K., Harley, M.D., Splinter, K.D., Simmons, J.A., Turner, I.L., 2019. Sub-annual to multi-decadal shoreline variability from publicly available satellite imagery. Coast. Eng. 150, 160–174. https://doi.org/10.1016/j.coastaleng.2019.04.004.
Vousdoukas, M., Ranasinghe, R., Mentaschi, L., Plomaritis, T., Athanasiou, P., Luijendijk, A., Feyen, L., 2020. Sandy coastlines under threat of erosion. Nat. Clim. Change 10, 260–263. https://doi.org/10.1038/s41558-020-0697-0.
Watanabe, S., 2023. Tree-Structured parzen estimator: understanding its Algorithm components and their roles for better empirical performance (arXiv:2304.11127). https://doi.org/10.48550/arXiv.2304.11127.
Weber de Melo, W., Pinho, J., Iglesias, I., 2023. Coastal morphodynamic emulator for early warning short-term forecasts. Environ. Model. Software 165, 105729. https://doi.org/10.1016/j.envsoft.2023.105729.
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.
Zhao, Z., Chen, W., Wu, X., Chen, P.C.Y., Liu, J., 2017. LSTM network: a deep learning approach for short-term traffic forecast. IET Intell. Transp. Syst. 11 (2), 68–75. https://doi.org/10.1049/iet-its.2016.0208.
Zhou, F., Chen, Y., Liu, J., 2023. Application of a new hybrid deep learning model that considers temporal and feature dependencies in rainfall–runoff simulation. Remote Sens. 15 (5). https://doi.org/10.3390/rs15051395. Article 5.