

Investigating appropriate artificial intelligence approaches to reliably predict coastal wave overtopping and identify process contributions

Michael McGlade^{a,*}, Nieves G. Valiente^a, Jennifer Brown^b, Christopher Stokes^{a,c}, Timothy Poate^a

^a Faculty of Science and Engineering, University of Plymouth, PL4 8AA, Plymouth, United Kingdom

^b National Oceanography Centre, SO14 3ZH, Southampton, United Kingdom

^c Met Office, Fitzroy Rd, EX13PB, Exeter, United Kingdom

ARTICLE INFO

Keywords:

Artificial intelligence
Wave overtopping
Flood forecasting
Climate change
Random forests
Digital twin

ABSTRACT

Predicting coastal wave overtopping is a significant challenge, exacerbated by climate change, increasing the frequency of severe flooding and rising sea levels. Digital twin technologies, which utilise artificial intelligence to mimic coastal processes and dynamics, may offer new opportunities to predict coastal wave overtopping and flooding reliably and computationally efficiently. This study investigates the effectiveness of training various artificial intelligence models using wave buoy, meteorological, and recorded coastal wave overtopping observations to predict the occurrence and frequency of overtopping at 10-minute intervals. These models have the potential for future large-scale global applications in estimating wave overtopping and flood forecasting, particularly in response to climate warming. The model types selected include machine-learning random forests, extreme gradient boosting, support vector machines, and deep-learning neural networks. These models were trained and tested using recorded observational overtopping events, to estimate wave overtopping and flood forecasting in Dawlish and Penzance (Southwest England). The random forests performed exceptionally well by accurately and precisely estimating coastal wave overtopping and non-overtopping 97 % of the time within both locations, outperforming the other models. Moreover, the random forest model outperforms existing process-based and EurOtop-based models. This research has profound implications for increasing preparedness and resilience to future coastal wave overtopping and flooding events by using these random forest models to predict overtopping and flood forecasting on wider global and climate scales. These trained random forests are significantly less computationally demanding than existing process-based models and can incorporate the important effect of wind on overtopping, which was neglected in existing empirical approaches.

1. Introduction

Anthropogenic climate change is increasing the frequency and magnitude of coastal flooding and wave overtopping (Jennath and Paul, 2024). By mid-century, 300 million individuals worldwide may experience frequent coastal flooding (Kulp and Strauss, 2019). Coastal flooding can cause significant fatalities; in 2022 alone, it resulted in 7398 fatalities worldwide (Rae et al., 2023). Currently, coastal flooding causes considerable infrastructural damage across Europe, amounting to €1.4 billion annually (Vousdoulkas et al., 2018). Predicting coastal wave overtopping accurately and precisely is highly desirable given the severe fatalities, economic damage, and growing concerns over climate change (Whittaker et al., 2018).

Both empirical (for example, ‘EurOtop’ and process-based (for example, ‘SWASH’ (Suzuki et al., 2017) modelling approaches are now used to predict wave overtopping discharge at seawalls. However, they often incur errors due to inadequate spatiotemporal resolution, forcing data, bathymetry, and seawall geometry, applicability to non-deep waters, as well as uncertainties in the underpinning statistical relationships in the case of empirical approaches (Buccino et al. 2023). Furthermore, important physical processes, such as wind effects, are often oversimplified or neglected altogether due to insufficient empirical understanding (EurOtop., 2018). These factors all introduce uncertainties regarding the model performance (Lerma et al., 2018). Systems to forecast and forewarn against coastal overtopping have also been developed, including the Coastal Storm Modeling System (CoSMoS;

* Corresponding author.

E-mail address: michael.mcglade@plymouth.ac.uk (M. McGlade).

<https://doi.org/10.1016/j.ocemod.2025.102510>

Received 24 July 2024; Received in revised form 5 February 2025; Accepted 6 February 2025

Available online 6 February 2025

1463-5003/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Barnard et al., 2014), Resilience-Increasing Strategies for Coasts tool Kit (RISCKit; Van Dongeren et al., 2018), Operational Wave and Water Level (OWWL; Stokes et al., 2021) model, and the Clouds-to-Coast modelling framework (Zou et al., 2013). These systems require spatial downscaling of wave, tide, and meteorological conditions from ocean to coastal scales using process-based models (e.g., Stokes et al., 2021). However, these models are also computationally expensive because they include wave transformation processes such as nearshore shoaling, refraction, dissipation effects, and wave-current interactions (Zhang et al., 2011).

Tools such as EurOtop are computationally inexpensive if adequate forcing data is available near the location of interest. EurOtop provides a set of static empirical equations derived from extensive physical model data, which can be performed using basic computational tools (EurOtop., 2018). Nonetheless, the empirical equations in EurOtop exhibit scatter, as some beach-structure profiles are not well represented by the underpinning database that is mainly representative of laboratory experiments (Brown et al. 2020). This scatter can cause EurOtop predictions to sometimes be significantly different (i.e., out of range) from the observed overtopping results (EurOtop., 2018).

The OWWL model (Stokes et al., 2021) is an example of an overtopping early warning system that utilises EurOtop by integrating nearshore wave and water level forecasts, a roller dissipation model, and up to date coastal profiles to predict wave runup and overtopping at > 200 locations, over 100 km coastline. This moves beyond the static equations within EurOtop originally designed for seawall design and assessment, to a real-time forecasting system (Stokes et al., 2021). However, such systems inherit the shortcomings of EurOtop, notably not accounting for onshore and offshore wind speeds and directions, which are fundamental to overtopping, yet are poorly represented within EurOtop (Stokes et al., 2021). Moreover, coastal processes influencing overtopping are highly complex, interactive, and dynamic, presenting challenges for predictions based on empirical physical model data. There is an increasing desire to improve predicting wave overtopping beyond these conventional means using newly deployed technologies, such as artificial intelligence (AI), offering faster and more reliable results (Habib et al., 2023). However, major advancements regarding observational overtopping data, which is accurate, extensive, and high resolution, are needed (Yue et al., 2022).

The recent development of the WireWall instrument by the National Oceanography Centre (NOC) marked a significant advancement in recording and analysing wave overtopping (Brown et al., 2020; Yelland et al., 2023). WireWall is an in-situ system that records field observations of wave overtopping suitable for comparison with numerical predictions (Lashley et al., 2022). The capacitance wire system samples at 400hz, recording the contact signal of each wave that passes through the array of wires (Brown et al., 2020; Yelland et al., 2023). A quality control is performed to ensure at least two wires are hit within a couple of seconds of each other, and that predetermined signal thresholds are exceeded to remove noise. By correlating the corresponding wind and wave characteristics along with the recorded overtopping by WireWall, it provides, for the first time, an opportunity to develop a deep understanding of which variables are influencing overtopping (Brown et al., 2020).

Moreover, with the increased coastal resolution of large-scale numerical weather prediction models and the potential of AI, approaches that could serve as prototypes for real-time forecasts of coastal flooding are yet to be explored (Den Bieman et al., 2021). WireWall data have been used for validation of some overtopping systems (e.g., Lashley et al., 2022), however, their use as a training dataset for AI is also to be explored.

Several studies have examined training AI models to predict wave overtopping (e.g., Alshahri and Elbisy, 2022; Alvarellós et al., 2024; Den Bieman et al., 2020; Elbisy, 2023; Habib et al., 2023). These studies highlight the significant opportunities for AI to predict wave overtopping and determine which variables influence overtopping the most

and least significantly. However, many of these studies listed above rely on multiple datasets used by EurOtop. These datasets do not account for the important variables influencing wave overtopping, such as the onshore and offshore wind speed and directionality. Moreover, EurOtop uses static empirical equations, not high temporal-resolution (wave-by-wave) recorded observational overtopping data (Yelland et al., 2023).

AI can incorporate machine and deep learning approaches (Gupta et al., 2021). Machine learning models include random forests, extreme gradient booster (XGBoost), and support vector machines (SVM) (Habib et al., 2023). Deep learning AI include recurrent, feedforward, and convoluted neural networks, to name a few (Schmidhuber, 2015). Training different AI models using high-resolution recorded overtopping data from WireWall, coupled with the corresponding water levels, and meteorological and wave characteristic data, may significantly enhance the quality and quantity of the training data. Moreover, WireWall overtopping is recorded throughout the year, providing a seasonal comparison of how the different variables influence overtopping. Training AI models on recorded overtopping data collected throughout the year, with relevant variables, such as wind and wave data, could unlock new revelations around understanding the nature of wave overtopping.

This study aims to train and evaluate the testing performance accuracy of machine and deep-learning AI to predict wave overtopping occurrence and frequency (i.e., number of overtopping events per 10-min interval) in Dawlish and Penzance, southwest England. This study will train random forests, XGBoost, SVM, and neural networks using in-situ overtopping data (i.e. WireWall data). These models were selected to compare machine and deep learning performances for predicting wave overtopping. The study will also compare these AI models predictions against EurOtop overtopping estimates generated by OWWL. By using two different observation-rich study locations, this allows the demonstration of the potential to scaling these AI models to be applicable to various flood-prone coastal locations across the UK and elsewhere.

2. Methods and materials

2.1. Study area

Wave overtopping was investigated in Dawlish (50°35'01" N, 3°27'52" W) and Penzance (50°07'10" N, 5°32'15" W), southwest England (Fig. 1). These locations are susceptible to frequent and intense overtopping, especially between September and April (Dawson et al., 2016). On February 7th, 2014, an intense overtopping event breached the Dawlish seawall, causing significant infrastructural damage amounting to £50 million (Dawson et al., 2016). Penzance also suffers from frequent overtopping, whereby the promenade is often closed between October and March, irrespective of the seawall defences (Yelland et al., 2023).

Both study locations are characterised by predominant south-westerly Atlantic swell waves that refract, resulting in many inshore waves being southernly; and less frequent, easterly wind waves (Fig. 1). Dawlish and Penzance wave period averages 4–10 s (s) and 5–12 s, respectively (National Coastal Monitoring, 2024). Dawlish mean tidal range is approximately 4 - 4.4 ms (m) for spring tides and for neap tides is 0.2 - 2.7 m. For Penzance, the mean tidal range for spring tides range are approximately 5.5 m and for neap tides is 2.5 m. Dawlish has a prevailing south-westerly wind direction, averaging 5–7 m/s. Penzance has a prevailing west-southwest wind direction, averaging 5 - 10.3 m/s. The beach geometry for Dawlish and Penzance contains a gentle to moderate slope elevation (Jane et al., 2018).

The Dawlish seawall (Fig 2a) was constructed using precast concrete sections. This seawall reaches 5.64 m above Ordnance Datum (ODN), with the walls toe extending, on average, 2.8 m further out towards the sea (Adams and Heidarzadeh, 2023). This seawall also includes a

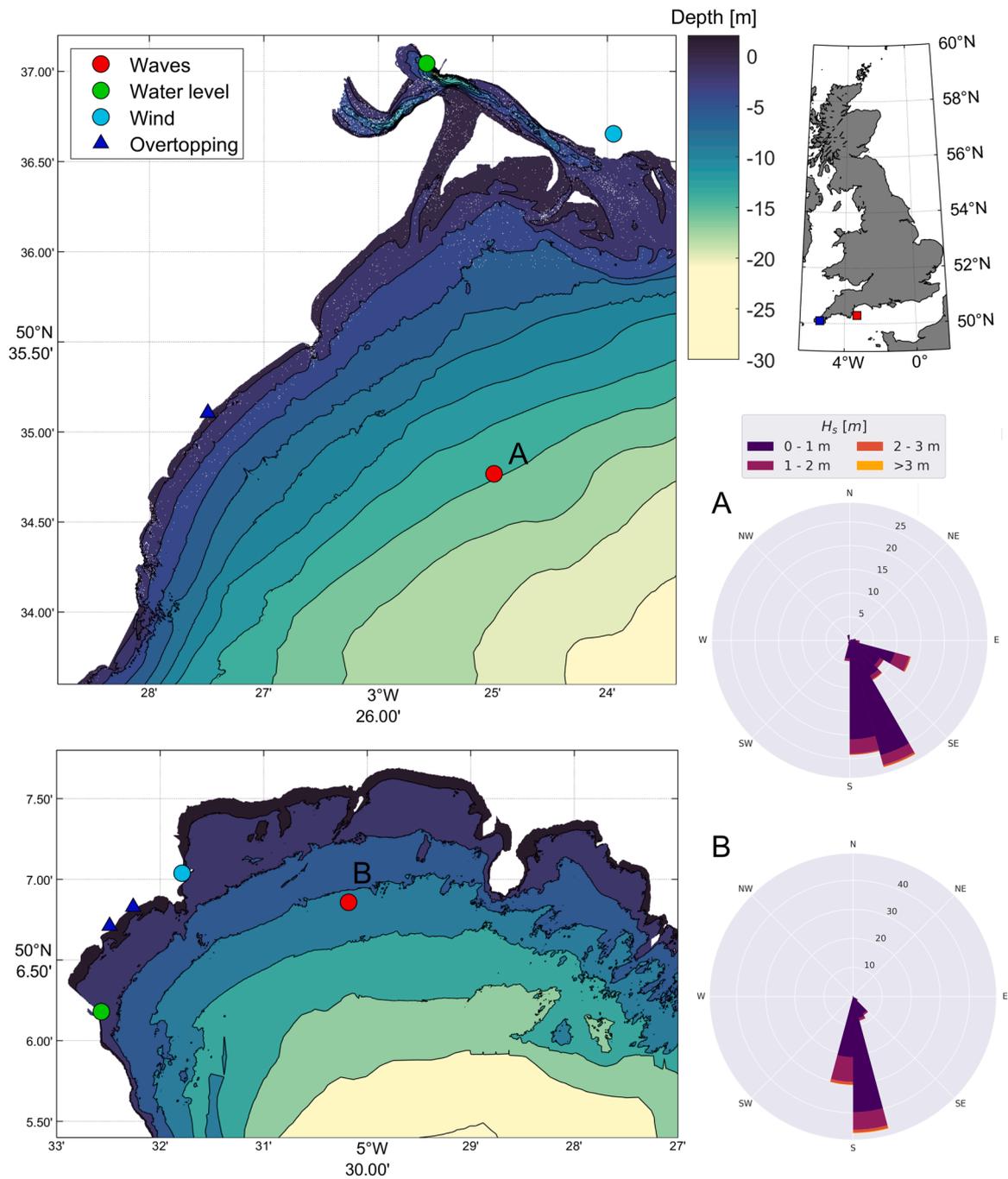


Fig. 1. Bathymetric contoured map with deployment location for Dawlish (top) and Penzance (bottom). Wave roses of Southwest Coastal Monitoring directional waverider buoy for Dawlish for 2010–2023 (A) and Penzance for 2007–2023 (B).

recurred section at the top for deflecting oncoming waves. The Penzance seawall (Fig 2b) reaches approximately 3.7 m above ODN and is slightly curved (35°) for dissipating wave energies.

2.2. Wave overtopping and metocean observations

WireWall data recorded overtopping events on Dawlish and Penzance seawall (Yelland et al., 2024). The capacitance wire system detects waves contacting the wires. We use the overtopping data (Brown et al., 2022; Yelland et al., 2024) from wires inland of the crest of the seawall to assess the wave overtopping that has an inland motion and can potentially pose a hazard to people and infrastructure. Penzance WireWall deployment period was November 16th 2021 to March 15th

2022, and March 10th 2021 to March 17th 2022 for Dawlish. The temporal resolution of the WireWall observations was every tide, from three hours before and after high tide, and the data were sampled at 400hz and processed into 10-min statistics, intervals suitable for comparison to operational monitoring of other metocean parameters and for capturing the change in environmental conditions at the coast. Two WireWall systems were deployed at the crest of the seawall in Penzance capturing long-shore variability and two other WireWall systems deployed in Dawlish, one at the seawall crest and a smaller (WireWand) system deployed behind it to measure cross-shore variability (Brown et al., 2022).

For Dawlish, “Rig 1” refers to the WireWall system at the seawall crest and “Rig 2” is closer inland behind Rig 1. Penzance has two rigs at

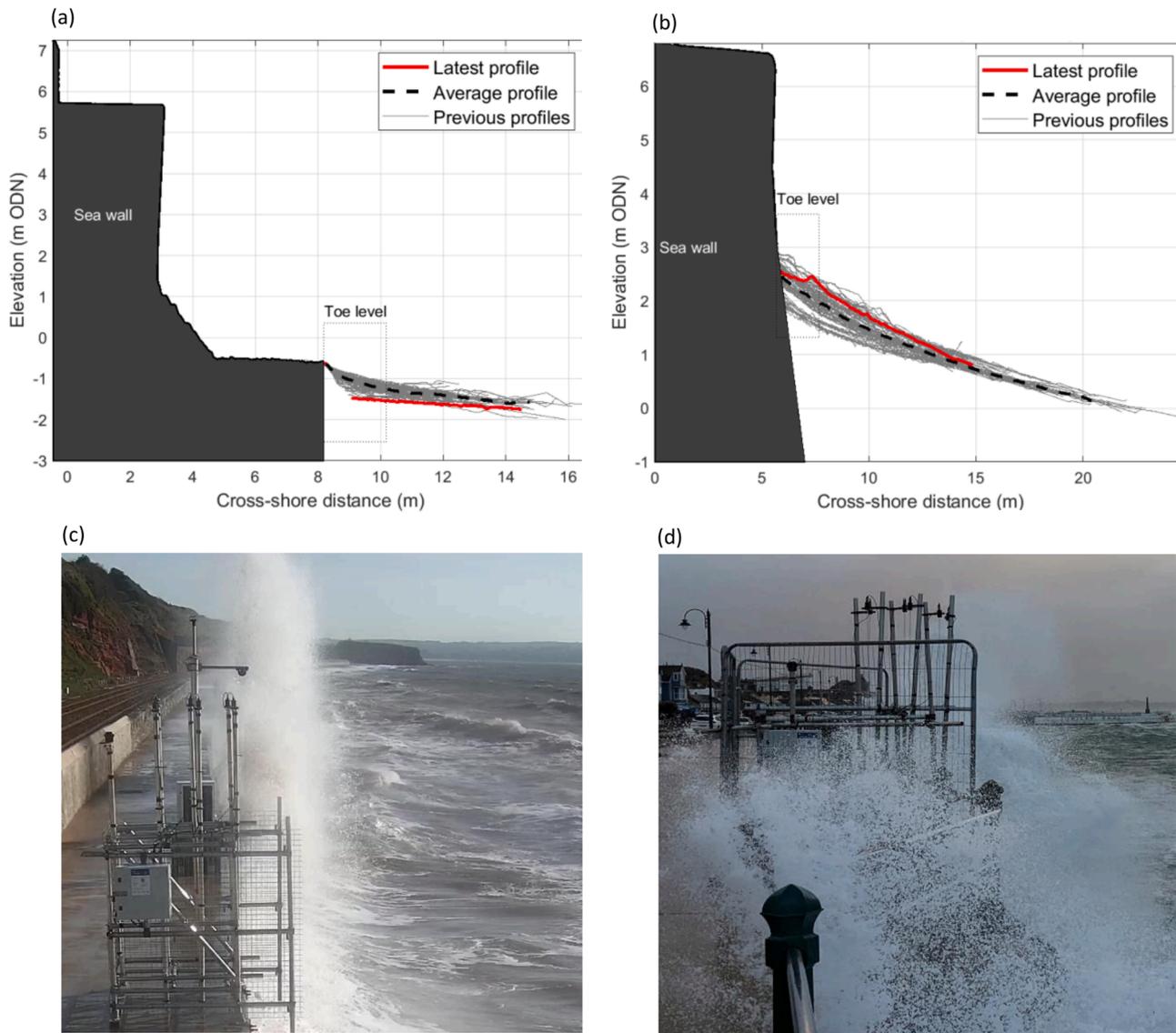


Fig. 2. Seawall cross section for (a) Dawlish and (b) Penzance and WireWall apparatus setup for (c) Dawlish and (d) Penzance.

the seawall crest; “Rig 1” exposed to SE and SW winds (Eastern location near Queens Hotel) and “Rig 2” which is sheltered on SW winds (westerly rig near the harbour). The Dawlish WireWall was configured with two adjacent rows 50 cm apart. Each row consisted of three capacitance sensors measuring the aerated water depth (solid flow and spray) 10 cm, 47 cm, and 84 cm inland from the seawall edge (Fig. 2c). In Penzance, both rigs had similar configurations to Dawlish Rig 1, with sensors placed at 0 cm, 37 cm, and 74 cm inland from the seawall crest (Fig. 2d). This study utilised the middle wires, as the outer and inner wires a couple of times were washed away from the waves. Corresponding meteorological conditions (Fig. 1) like the wind speed (U_{10}) and the mean wind direction (U_{10} Dir) during the recorded overtopping events were measured at 30-min intervals using local weather stations (NNRCMP, 2024). 30-min wave characteristics, such as the significant wave height (H_s), mean period (T_m), peak period (T_p), and the mean wave direction (D_m) were extracted from the directional wave buoy (NNRCMP, 2024). A marine radar system (WaveRadar REX) recorded the water level (WL) every 10-min and 15-min for Dawlish and Penzance, respectively.

2.3. Model data processing

The wind, wave, WL, and WireWall data were processed and interpolated in 10-min intervals. All the feature variables were assessed for multicollinearity using the variance inflation factor. Homoscedasticity was assessed using residual plots and dataset normality was analysed using the Shapiro-Wilk Test. Any relevant feature variables to predict wave overtopping were determined using the random forest variable importance metric (VIM). The VIM uses the random forest out-of-bag data to establish which feature training variables reduce the testing error Gini Impurity (classification) or mean square error (regression) (Janitza and Hornung, 2018). If a particular feature variable trained within the random forest considerably reduces the Gini Impurity or mean squared error, then this variable would be assigned a “high” importance (Janitza and Hornung, 2018). The AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion), which evaluates the goodness of the model fit, while penalising the model complexity, were used to assess the effectiveness of the different AI models training using various feature variables (Güney et al., 2021). See Appendix Table 4 for term definitions.

2.4. AI development

This study trained, tested, and compared supervised machine learning models - specifically random forests, XGBoost, and SVM - with a multilayer deep learning perceptron neural network to predict overtopping and frequency of overtopping during 10-min intervals for the two rigs at the two study locations (i.e., Dawlish and Penzance). These AI models were selected to compare the performance metrics between machine and deep learning models and different architecture designs.

Each model incorporated a binary classifier for estimating overtopping occurrence (yes/no) and a regression predictor for analysing overtopping frequency (i.e., number of overtopping events per 10-min intervals). All the models were computed using the TensorFlow library for neural networks (Version 2.16), XGBoost library for XGBoost algorithms (Version 2.0.3), and scikit-learn library for the random forests and SVM modelling (Version 1.4) (Chen and Guestrin, 2016; Pedregosa et al., 2011).

The input dataset, consisting of wave, wind, and overtopping observations, for Dawlish and Penzance, was randomly partitioned into training (80%), testing (10%), and validation (10%) for all the machine and deep-learning models to predict wave overtopping. By randomly partitioning the data, this avoids the unintentional bias in data splitting. This larger ratio of training data (80%) ensured the models could sufficiently capture intricate relationships and correlations within the dataset, improving the classification accuracy (Ramezan et al., 2021). Additional training data also enhances the AI models understanding of how different feature variables influence overtopping across seasons, improving the predictive robustness and generalisability. Any overfitting exhibited was evaluated using learning curves and 10-fold cross-validation (Nti et al., 2021). We selected 10-fold cross validation over 5-fold, aiming to increase the machine and deep learning model robustness by validating the performance on more data split combinations (e.g., Tsung and Yeh, 2019).

If overfitting was exhibited, the machine or deep learning models were regularised using Lasso regression, i.e., L1 regularisation (Yang et al., 2023). This regularisation type was selected from its ability to handle many feature variables. There was considerable class imbalance

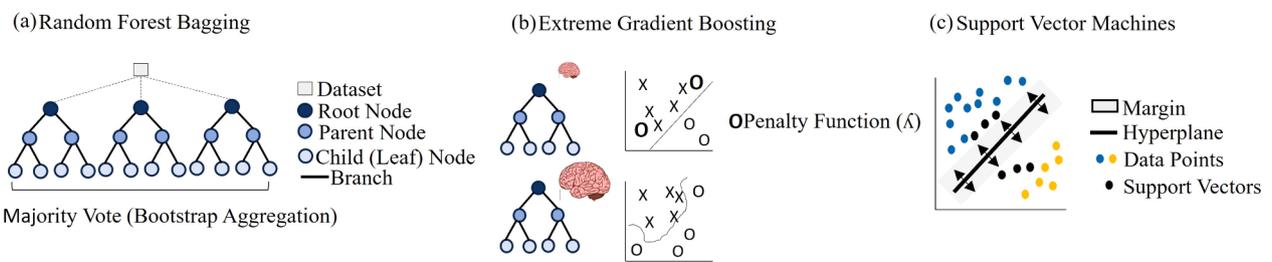
for the overtopping class, which represented approximately 6% of the total dataset, with the remaining 94% representing non-overtopping classes. To address this class imbalance, a synthetic minority over-sampling technique (SMOTE) was deployed for all the models. SMOTE creates synthetic samples for the minority overtopping class to balance the dataset (Wongvorachan et al., 2023). Importantly, some of the majority class “0” non-overtopping values were removed from the initial training dataset to minimise this class imbalance. During low tide periods when the WireWall was not recording, this information was not included in the training, validation or testing as there are no observations to determine if overtopping was or was not occurring. All the AI training datasets, including the code for constructing each model can be found in the supplementary material (McGlade et al., 2024).

2.4.1. Random forests

Random forests are ensemble classifiers containing multiple decision trees that formulate a prediction (Fig. 3a) (Galiano et al., 2012). Each decision tree is uniquely different through sampling the dataset with replacement. Sampling with replacement in a random forest means that each decision tree samples data randomly from the dataset, and after each selection, the data point is placed back into the dataset, allowing it to be potentially selected again with the same statistical probability (Özçift, 2011). By sampling with replacement, each decision tree uniquely trains on different data subsets, learning intricate relationships, correlations, and patterns within each subset (Rigatti, 2017). The subset of data that was not selected during the “sampling with replacement” process were the “out-of-bag” data (Wu et al., 2024). The out-of-bag data determined the variable importance for influencing overtopping and frequency.

A random forest model using bootstrap aggregation was constructed to predict the binary wave overtopping occurrence and wave frequency (i.e., number of overtopping events within a 10-min interval). The Gini Impurity assessed the random forests classification accuracy and the mean squared error determined the accuracy of counting the frequency of overtopping. The random forest model was hyperparameter tuned using systematic grid searches, adjusting variables such as the decision tree depth, number of splits and, decision tree number. The random

Machine Learning



Deep Learning

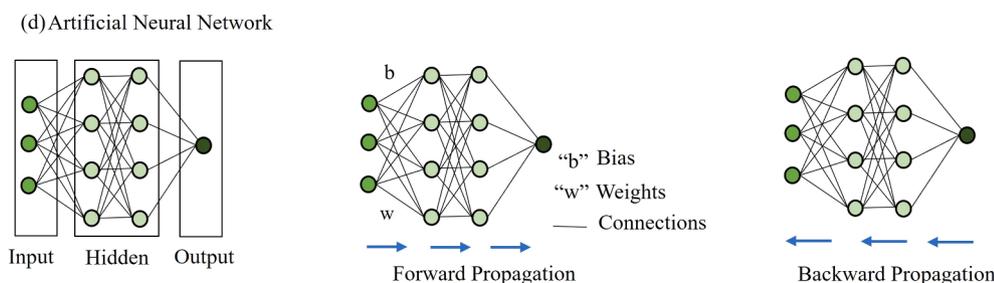


Fig. 3. AI schemes for (a) random forest models (b) XGBoost, (c) SVM, and (d) neural networks.

forest was then regularised (if needed) and the threshold was adjusted to optimise the performance.

2.4.2. XGBoost

An XGBoost is another decision tree classification model (Fig. 3b) (Ghatkar et al., 2019). XGBoost iteratively learns from the misclassifications of previous trees (Velthoen et al., 2023). The XGBoost was constructed using the booster parameter set to 'gbtree.' This XGBoost algorithm was programmed originally for 100 boosting rounds. The number of rounds was selected to view, using learning curves, the optimal number of boosting rounds to maximise performance and minimise overfitting. This XGBoost was then hyperparameter tuned using grid searches, altering the learning rate, booster type, verbosity, eta, decision tree depth, and the maximum number of decision tree nodes. After tuning, the XGBoost was regularised (if necessary), and the threshold was adjusted to yield optimised performance.

2.4.3. Neural networks

A neural network differs from machine learning models by generating predictions using a design inspired by the human brain (Park et al., 2019). In this design, the neural network receives the training dataset through its input layers (Fig. 3d) (Wright et al., 2022). Like the human brain, the input layer simulates an external sensory stimulus (Prieto et al., 2016). The neural network has hidden layers with multiple neurons that receive and process information from the input layer (Wright et al., 2022). The hidden layer is essentially the centre of the decision-making process, and the output layer represents the final prediction (Shen et al., 2021).

A neural network was computed using a sigmoid activation function for binary classification and a rectified linear unit for regression classification. The neural network was compiled using the Adam optimiser. This optimiser was chosen for its ability to adjust its adaptive learning rate for each feature variable within the dataset, which is critical for parameters with considerably different scales (Nwankpa, 2020). Early stopping was deployed during stochastic gradient descent to prevent overfitting by reducing the number of necessary epochs during the model training. The neural network was hyperparameter-tuned to optimise performance using systematic grid searches, adjusting the number of hidden layers, activation function type, neurons per layer, batch size, and learning rate. After the neural network was constructed, tuned, and regularised (if needed), the threshold was adjusted to harmonise the balance between the recall and precision values (i.e., F1 score, see Section 2.5).

2.4.4. SVM

SVM using a non-linear kernel can map the data into a high dimensional hyperspace (Fig. 3c) (Balraj et al., 2022). The SVM then finds the optimal hyperplane which separates the data within this hyperspace (Balraj et al., 2022). A SVM was computed using a linear kernel to map the dataset onto a higher dimensional hyperplane. This kernel was selected using systematic grid searches to find the optimal kernel type. The SVM was hyperparameter tuned, regularised (if overfitting was exhibited) and evaluated on the testing dataset.

2.5. AI performance metrics

The accuracy of the AI for predicting the binary classification of overtopping was evaluated using the F1 score, described as the harmonic mean between the model precision and recall (Chicco and Jurman, 2020). The F1 score is calculated as:

$$F1 = \frac{TP}{TP + 0.5(FP + FN)} \quad (1)$$

where TP represents the number of true positives, FP is the number of false positives, FN denotes the number of false negatives. The F1-score

can have values between 0 and 1, with 1 representing the best score. This study selected this score to balance consideration between the recall and precision, aiming to lower the total number of false positives and negatives. For machine learning purposes, an F1 score exceeding 0.70 is considered a good and adequate model (Humphrey et al., 2022; Yacouby and Axman, 2020). There were considerable data imbalances for the overtopping class with WireWall recording non-overtopping 94 % of the time and only 6 % overtopping events. The study by Chicco and Jurman (2020) cautions only using the F1 score for evaluating the model accuracy with these data imbalances, recommending using the Matthews Correlation Coefficient (MCC). Unlike the F1 score, MCC addresses the AI performance within all four quadrants of the confusion matrix (true positives, false negatives, true negatives, false positives) (Chicco and Jurman 2020). Therefore, in addition to using the F1 score, this study also evaluated the model performance using MCC. The MCC score is calculated as:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) + (TP + FN) + (TN + FP) + (TN + FN)}} \quad (2)$$

where TP is true positives, TN is true negatives, FP is false positives, and FN is false negatives. Receiver operating characteristics (ROC) evaluated all the machine learning performances by plotting the model's true positive rate (sensitivity) against false positive rate (1-specificity). The total Area Under the ROC curve (AUC) measures the model's ability to categorically distinguish between positive and negative examples. An AUC of 1 indicates perfect classification.

The AI performance predicting overtopping frequency was also evaluated using the correlation coefficient (R^2). However, this study goes beyond the conventional means of only reporting this statistic and analyses a suite of error metrics relating to R^2 . Additional error metrics include the testing dataset root mean square error (RMSE), mean absolute error (MAE), mean square error (MSE), bias and the Brier Skill Score.

2.6. AI framework

The development, training, and evaluation of all the AI models to predict wave overtopping followed a three-step workflow (Fig. 4). The first step involved incorporating the relevant training data, including weather station, wave buoy, tide gauge, and WireWall overtopping data. The second step involves selecting the relevant AI model and following the guidelines to construct the model. The third step involves reviewing the probability of overtopping occurrence along with the overtopping frequency per 10-min window.

2.7. Comparison to EurOtop predictions

To provide a comparison to a traditional method, predictions of overtopping discharge were generated using EurOtop equations. The same wave and tide forcing was used as in the machine learning models, but as EurOtop equations require information about hydrodynamic forcing at the seawall toe, the nearshore wave conditions at the wave buoys were first transformed into the coast using the parametric breaker dissipation model described by Janssen and Battjes (2007). This iteratively solves for wave induced set up and shoaling/broken wave height across a coastal bathymetric profile. The coastal profiles at Penzance and Dawlish were obtained by merging RTK-GPS topographic surveys of the intertidal beach beneath the WireWall with nearshore multibeam bathymetric surveys (nominal vertical accuracies of ± 0.03 m and ± 0.15 – 0.20 m, respectively) collected by the regional coastal monitoring programme (NNRCMP, 2024). More details on the method for computing wave conditions at the seawall are provided in Stokes et al. (2021). The wave overtopping discharge (Eq. (3)) (m^3/s per m crest width) at a given point in time was then estimated using EurOtop's equations for a vertical seawall, given as (EurOtop, 2018)

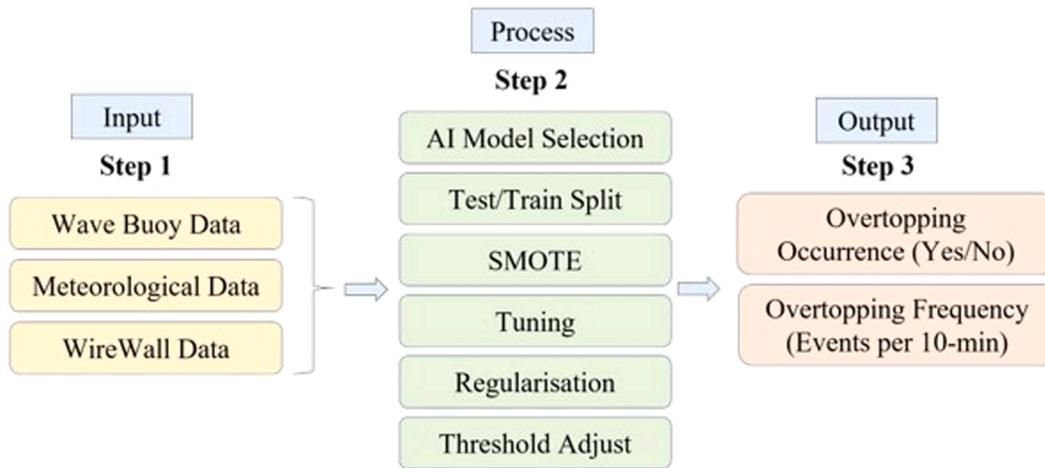


Fig. 4. Workflow process to build the overtopping AI models.

$$\frac{q}{\sqrt{gH_{m0}^3}} a \exp \left[- \left(b \frac{R_c}{H_{m0}} \right)^c \right] \text{ for } R_c \geq 0 \quad (3)$$

which predicts that $Q/\sqrt{gH_{m0}^3}$ increases exponentially as the crest freeboard (R_c) decreases relative to the spectral significant wave height at the seawall toe (H_{m0}). "a, b, and c" are fitted coefficients which vary depending on the structure type and forcing conditions. The left-hand side of Eq. (3) represents the wave overtopping discharge normalised by wave height. The right-hand side represents the crest R_c normalised by wave height. Stokes et al. (2021) found that overtopping enhancement due to onshore winds combined with impulsive breaking using $\gamma_w = 2$ (as suggested in EurOtop., 2018) was found to produce unrealistically high overtopping volumes in many cases and the effect of wind on overtopping was therefore disregarded ($\gamma_w = 1$), highlighting the deficiencies that traditional methods have regarding the influence of wind. The influences of a shallow foreshore at both sites, a toe mound in front of the seawall at Dawlish, and 'impulsive' wave breaking (when breaking occurs directly onto the sea defence) were included in the calculations as per the guidance in EurOtop. (2018). EurOtop generates wave discharge output volumes. Any overtopping volumes exceeding

0.1L/s/m were converted to overtopping (1) classification, and any volumes below 0.1L/s/m were converted to non-overtopping (0). The value of 0.1L/s/m was selected based on the EurOtop Manual (2018) as the tolerable mean wave overtopping discharge to ensure pedestrian safety.

3. Results

3.1. Variables influencing wave overtopping

According to the AIC and BIC estimates, the optimal variable combination to train these machine and deep learning models, which yielded the highest AIC and BIC scores, were the H_s , T_m , U_{10} Dir, D_m , U_{10} , and the freeboard (R_c) (computed as the distance between the seawall crest and WL). In accordance with EurOtop. (2018), the random forest variable importance metric identifies the R_c and H_s as important variables which influence wave overtopping for both rigs in Dawlish and Penzance (Fig. 5). Specifically, H_s influences wave overtopping more noticeably in Dawlish (Fig. 4a) and the R_c in Penzance (Fig. 5). These findings highlight the need to incorporate these variables into AI models to predict wave overtopping accurately. The SVM yielded poor predictive performance results (not shown).

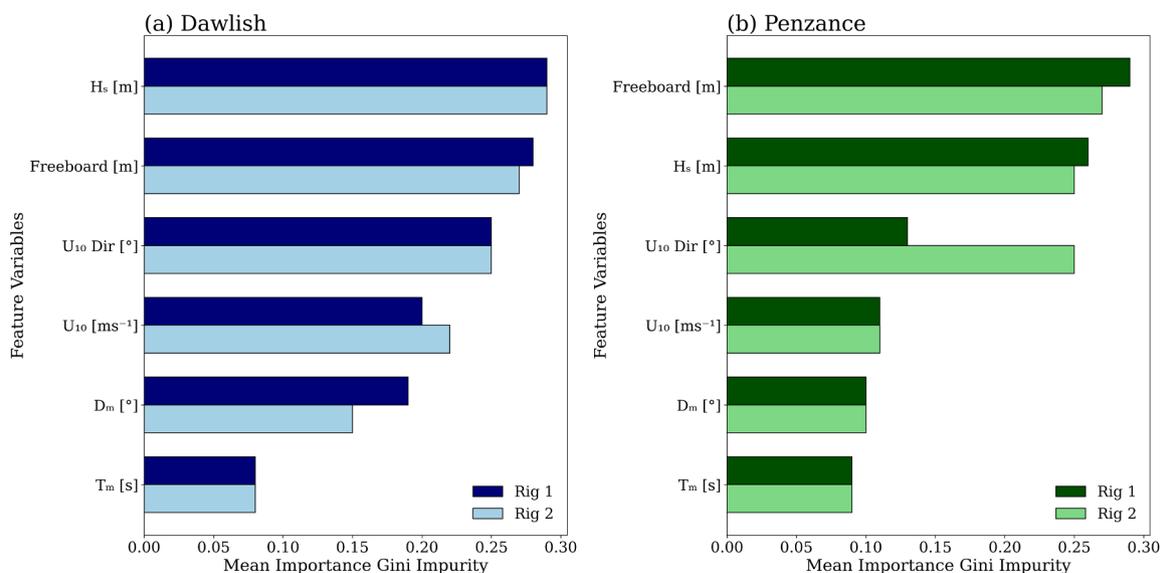


Fig. 5. Variable importance for (a) Dawlish at Rig 1 (dark blue) and at Rig 2 (light blue) and (b) Penzance at Rig 1 (dark green) and at Rig 2 (light green). The relative importance for each feature variable is computed using the Gini Impurity. Variables include H_s , T_m , D_m , U_{10} Dir, U_{10} , and Freeboard (R_c).

In both locations, T_m has the least influence on overtopping for both rigs, whereas U_{10} and U_{10} Dir are very important variables (fourth and third in variable importance, respectively). There is likely significant correlation between the wind and wave variables, potentially leading to interchangeability in feature importance depending on the random forest initialisation. To fully assess U_{10} and U_{10} Dir unique contribution, the random forest model was retrained without these wind features. Including wind variables improved the overtopping occurrence predictions, with F1 scores consistently above 0.75 and $R^2 > 0.70$ (Tables 1 and 2). Excluding the wind features decreased the F1 scores significantly (e.g., Penzance Rig 2; $F1 = 0.27$) and reduced the R^2 , especially at Rig 2 in Dawlish and Penzance ($R^2 = 0.19$, $R^2 = 0.25$, respectively).

The key number feature variables influencing wave overtopping varies between Dawlish and Penzance. In Penzance, H_s and R_c noticeably influence overtopping, suggesting these two variables alone could reliably and accurately predict overtopping (0.3 and 0.28 mean importance, respectively). While R_c and H_s are important in Dawlish, wind-related variables, such as U_{10} , have a more pronounced impact (e.g., 0.18 at Rig 1 in Dawlish). These findings suggest that while all these variables are important for predicting overtopping, locations like Penzance have fewer predominant variables influencing overtopping, whereas, in Dawlish, the influence is more distributed among various variables.

In Dawlish, overtopping frequency increases with increasing H_s once H_s exceeds 2 m (Fig. 6). Overtopping frequency exceeds 100 events per 10-min interval when H_s is around 2 m, and the R_c is approximately 4.5–5 m Chart Datum (CD). The overtopping frequency is significantly influenced by local winds, exceeding 50 overtopping events per 10 min, during moderate wind speeds ($U_{10} = 5–15$ m/s). In Penzance, high overtopping frequency is observed for $H_s \geq 2.5$ m, and R_c is 3.5–5 m CD. Overtopping occurs most frequently during SE winds when the U_{10} is approximately 5–20 m/s.

In both locations overtopping occurrence is less frequent but still happens during low wave conditions ($H_s < 1$ m). Overtopping occurs at both sites during E low SE-E winds. These findings demonstrate the importance of appreciating which feature variables are more relevant than others when predicting overtopping occurrence, potentially reducing the data intensity required for training these AI models, excluding redundant and unnecessary feature data. However, these findings do not suggest or recommend ignoring low scoring variable importance features, such as T_m and D_m , but rather, suggests a careful consideration when inputting such variables into predictive AI modelling like random forests with the aim of reducing performance error metrics.

3.2. Predicting wave overtopping occurrences

For Dawlish, the random forest model outperformed the XGBoost and neural network for estimating wave overtopping occurrence (i.e., overtopping versus non-overtopping) at Rig 1 and Rig 2, yielding an F1 score of 0.83 and 0.80, respectively (Table 3). The random forest had a high cross-validation accuracy for estimating overtopping and non-overtopping classes for both rigs at 96 % and 97 %. For both rigs, the random forest model correctly identified wave overtopping 80 % and 81 %, with 85 % and 79 % precision for Rig 1 and Rig 2, respectively. MCC confirms the high F1 scores for both rigs at 0.79 and 0.76. The other

models, particularly the XGBoost, performed very well for estimating overtopping and non-overtopping; however, regarding the predictive performance, the random forest model outperformed the other models. For the specific random forest hyperparameter tuning metrics, as well as the tuning metrics for the XGBoost, neural network, and SVM, see Tables 1–3, Appendix. The ROC curves for the XGBoost, random forest, and neural networks all show low instances of false positives with a high area under the curve (> 0.97) for Dawlish at Rig 1 and Rig 2 (Figs. 7 and 9). The AUC values indicate that each model has an excellent discrimination between overtopping and non-overtopping classes ($AUC = 0.95 – 0.97$). All three models have slightly more challenges detecting overtopping compared to non-overtopping. The Penzance ROC curves for the XGBoost, random forest, and neural networks all show reasonably low instances of false positives, with a high area under the curve for Penzance at Rigs 1 and 2 (Figs. 8 and 10). However, the neural network, particularly at Rig 2, shows high instances of false positives and negatives.

For Penzance, again the random forest model outperformed the other models for estimating wave overtopping within both rigs (Table 4). For Rig 1 and Rig 2, the random forest model yielded an F1 score of 0.86 and 0.75, and an accuracy of 86 % and 85 %, respectively. Moreover, the random forest precision was 86 % and 65 %. Overall, considering overtopping and non-overtopping classes, the random forest achieved a cross-validation accuracy of 97 % and 95 %. MCC confirms the reasonably high F1-scores of 0.81 and 0.71, respectively. Since false negatives are significantly important, minimising such false negatives, specifically for random forests, Gupta et al. (2021) recommends adjusting the model’s recall level. This adjustment involves changing the threshold level, which may negatively impact the precision score. Increasing the recall reduces false negatives at the expense of potentially increasing false positives. For predicting wave overtopping, with the aim of minimising casualties, this approach may indeed be appropriate. The trade-off with reducing false negatives at the expense of more false positives may have economic implications with unnecessary closure of locations for safety purposes.

The random forest model achieved the highest predictive performance for both rigs at Penzance and Dawlish, followed by the XGBoost, with the neural network showing the weakest performance, likely attributing to limited training data. The random forest exhibits variable importance metrics, allowing the model to prioritise on more important variables, potentially giving this model better performances over the others. See Section 4.2 for detailed performance insights.

3.3. Random forest model error performance

The evaluation of the misclassified predictions in the confusion matrices for Dawlish and Penzance shows the random forest outperforming the other AI models (Figs. 7 and 8). This analysis focuses on the misclassified variables within the random forest model testing dataset, particularly examining the relation between the misclassified instances and the different feature variable values (e.g., H_s , T_m). We emphasise analysing false negatives due to their significant implications for safety and damage. In other words, it is important to understand the model classifications and mainly address the random forest limitations when, falsely reporting non-overtopping.

In Dawlish, several false negatives occur when the overtopping

Table 1

Random forest performance (binary model) for the testing dataset comparing the addition and exclusion of U_{10} and U_{10} Dir features.

Rig	Location	Including U_{10}/U_{10} Dir				Excluding U_{10}/U_{10} Dir			
		F1	Recall	Precision	MCC	F1	Recall	Precision	MCC
1	Dawlish	0.83	0.80	0.85	0.79	0.78	0.76	0.81	0.77
2	Dawlish	0.80	0.81	0.79	0.76	0.67	0.58	0.80	0.67
1	Penzance	0.86	0.86	0.86	0.81	0.84	0.82	0.87	0.81
2	Penzance	0.75	0.85	0.85	0.71	0.32	0.27	0.38	0.30

Table 2
Random forest performance (regression) for the testing dataset comparing the addition and exclusion of U_{10} and U_{10} Dir features.

Rig	Location	Including U_{10}/U_{10} Dir				Excluding U_{10}/U_{10} Dir			
		R^2	MAE	RMSE	MSE	R^2	MAE	RMSE	MSE
1	Dawlish	0.81	0.66	3.05	9.28	0.65	0.87	4.07	16.6
2	Dawlish	0.76	0.23	1.90	3.62	0.19	0.40	3.47	12.1
1	Penzance	0.84	2.53	7.87	6.91	0.73	3.41	10.24	104
2	Penzance	0.84	11.8	15.3	2.33	0.25	0.70	5.37	29

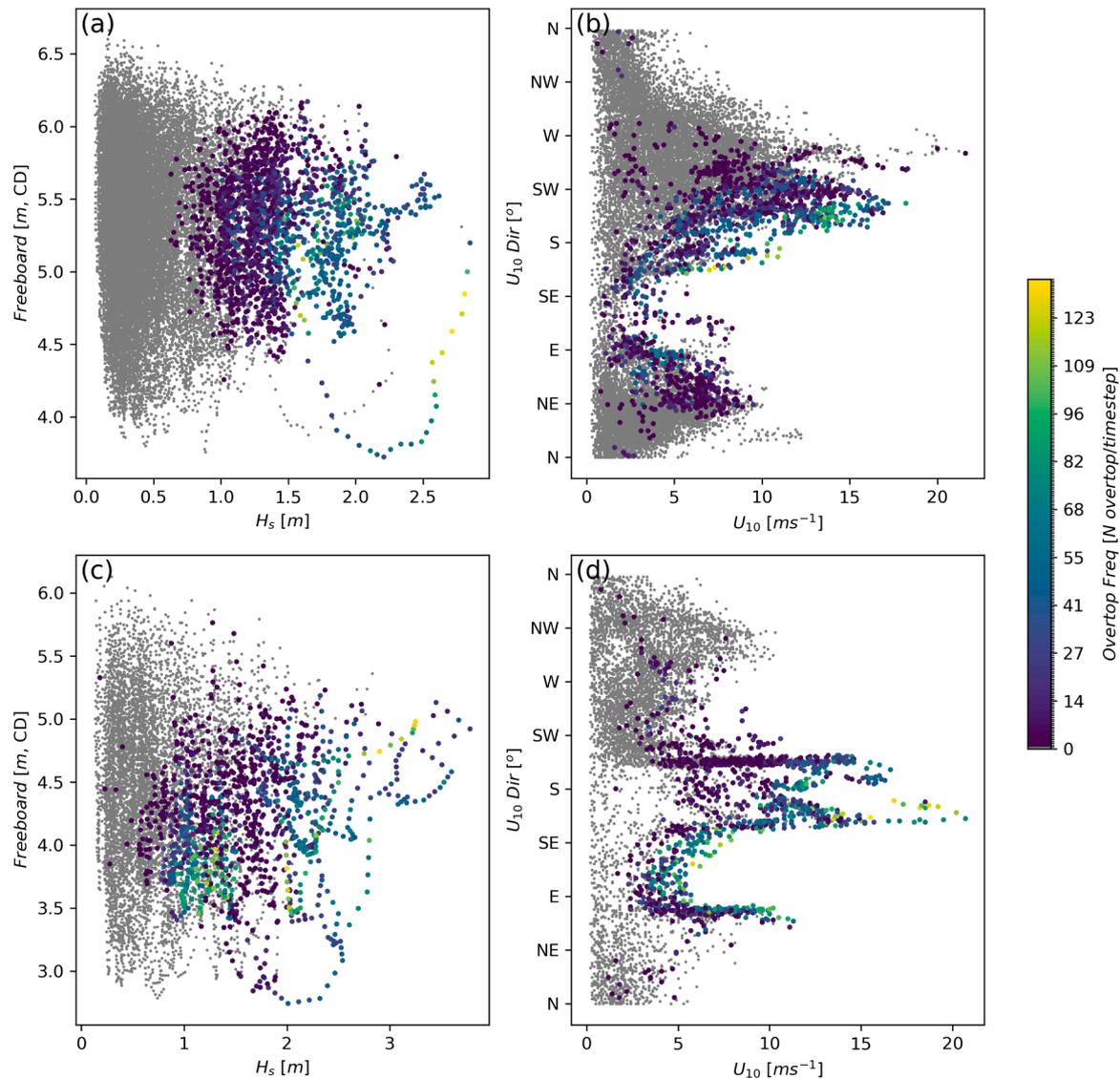


Fig. 6. Wave and wind overtopping relationship for Dawlish (a,b) and Penzance (c,d). (a,c). Scatter plots for freeboard (R_c) and H_s for Dawlish and Penzance, respectively. (b,d). Scatter plots for U_{10} and U_{10} Dir for Dawlish and Penzance, respectively. Scatter colours represent recorded overtopping frequency.

Table 3
AI performance metrics for estimating overtopping occurrence in Dawlish.

Rig	Model	AIC	BIC	F1	Precision	Recall	Accuracy	MCC	Brier Score
1	Random Forest	421	467	0.83	0.85	0.80	96 %	0.79	0.031
	XGBoost	394	345	0.81	0.82	0.79	94 %	0.77	0.032
	Neural Network	400	354	0.70	0.72	0.67	93 %	0.77	0.032
2	Random Forest	345	433	0.80	0.79	0.81	97 %	0.76	0.011
	XGBoost	367	400	0.78	0.74	0.80	95 %	0.71	0.019
	Neural Network	342	361	0.72	0.70	0.74	93 %	0.70	0.034

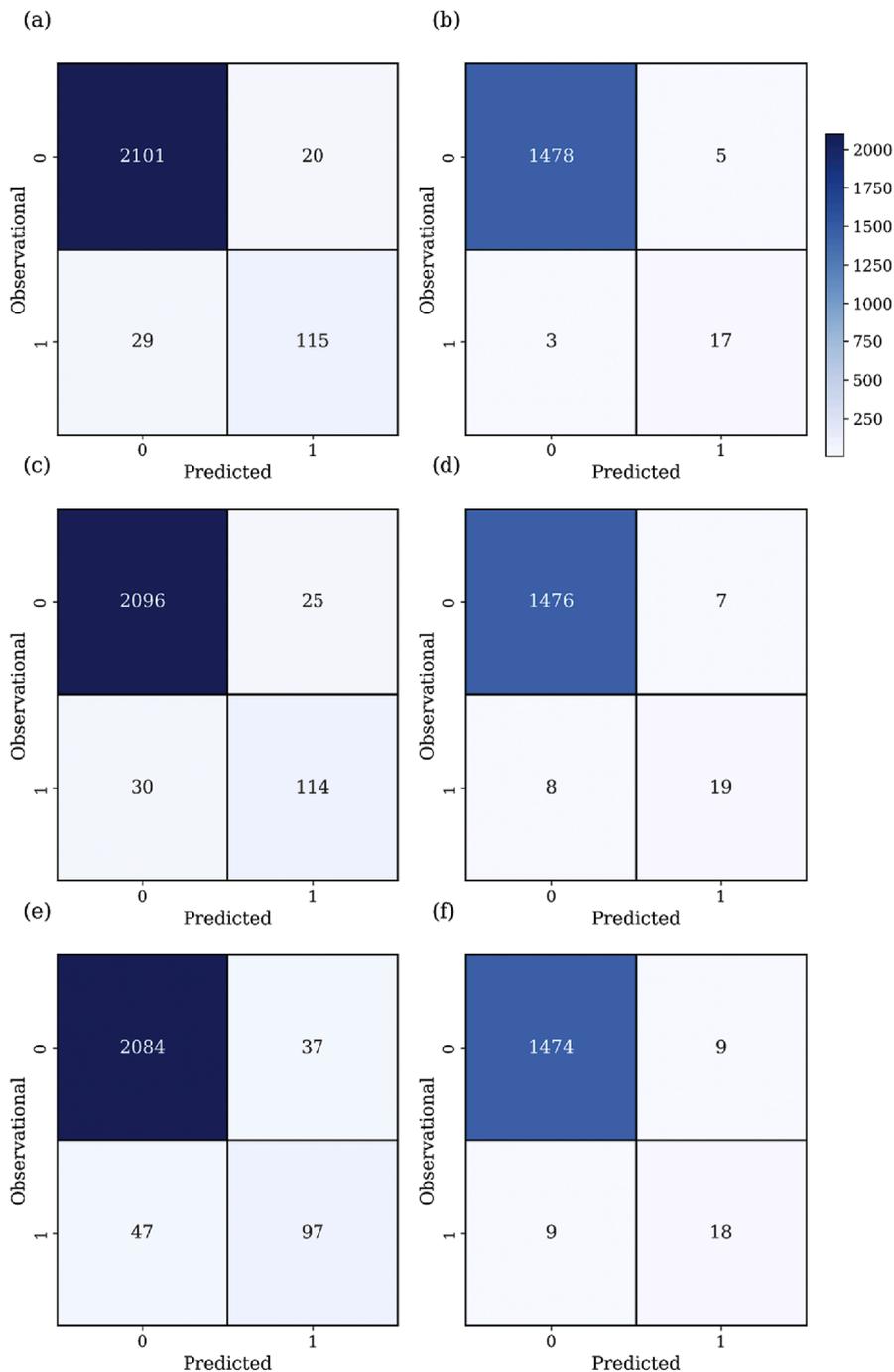


Fig. 7. Confusion matrix for machine and deep-learning models at Dawlish: random forest Rig 1 (a) and rig 2 (b); XGBoost Rig 1 (c) and Rig 2 (d); and neural network Rig 1 (e) and Rig 2 (f). In each panel, top left corresponds to TN, top right to FP, bottom left to FN and bottom right to TP.

frequency is low, typically around 1–2 overtopping events per 10-min window. (Fig. 11). In other words, when the random forest creates false negatives, the overtopping frequency is not severe. There is one recorded instance in April 2021 when the random forest model misclassified no overtopping when the overtopping frequency was four overtopping events per 10-min window. There are several patterns between misclassified false negatives and the feature variable values. For example, when the H_s is between 0.7 and 1.1 m, false negatives considerably increase. Several false negatives derive when the T_m is between 3.8 and 4 s and when the R_c levels range between 2.5 and 3 m.

In Penzance, like Dawlish, misclassified false negatives only occur when the overtopping frequency is low (Fig. 12). In late January and

mid-February, the random forest model misclassified non-overtopping in two instances during an overtopping event with a frequency of four. Similar patterns exist between the feature variable values and the recording of false negatives. Many false negatives occur when the R_c is between 4 and 5 m. Moreover, when the H_s is approximately 1 m, the random forest sometimes misclassifies. Overall, false negatives occur when overtopping frequency is low with these false negatives occurring between $H_s = 0.7$ to 1 m and with $R_c = 2.5$ to 3 m for Dawlish, and $R_c = 4.5$ to 5 m in Penzance.

The random forest model misclassification attribute to small ranges in R_c , H_s , and T_m levels (wave steepness). This misclassification within these ranges is unsurprising, given that several of these key variables,

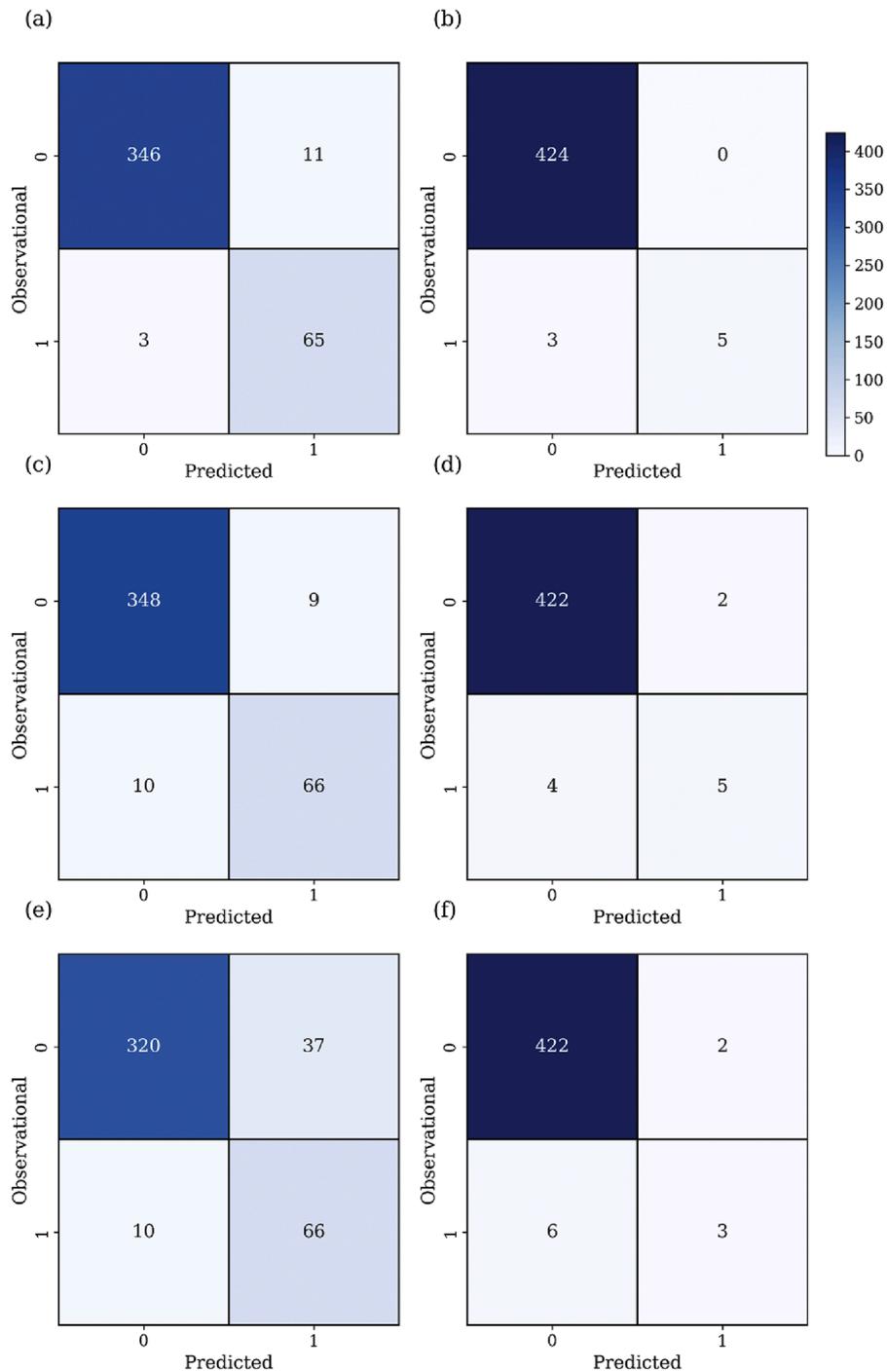


Fig. 8. Confusion matrix for AI models at Penzance: random forest Rig 1 (a) and Rig 2 (b); XGBoost Rig 1 (c) and Rig 2 (d); and neural network Rig 1 (e) and Rig 2 (f). In each panel, top left corresponds to TN, top right to FP, bottom left to FN and bottom right to TP.

particularly the H_s and R_c , were regarded as the most important variables influencing overtopping for both locations (Fig. 5). These misclassifications are similar for Dawlish and Penzance, irrespective of local wave and wave differences. The random forest model is highly robust above and below this specified R_c of 2.5 - 3 m, H_s , and T_m ranges; however, the random forest is sensitive to uncertainty within these ranges. These feature variables have complex non-linear interactions. For example, a combination of high T_m levels and low R_c may typically indicate overtopping; however, if the H_s is also low, overtopping may not occur.

This non-linear relationship may become stochastic during these specified ranges. By “stochastic” in this context, there is noise within

these feature variable ranges, which “confuses” the random forest model to report non-overtopping during overtopping events. Moreover, many of these false negatives, particularly for Dawlish, occur during the summer. These findings are unsurprising given that there are fewer overtopping events during summer, meaning the random forest has fewer training data to learn, suggesting that seasonality is not fully addressed. The random forest needs further improvement for predicting overtopping occurrences during summer periods, indicating that there are potential challenges for the model to discern between overtopping and non-overtopping classes. Wave overtopping is considerably less problematic during the summer compared to spring and winter, with many businesses and local authorities easing restrictions to prevent

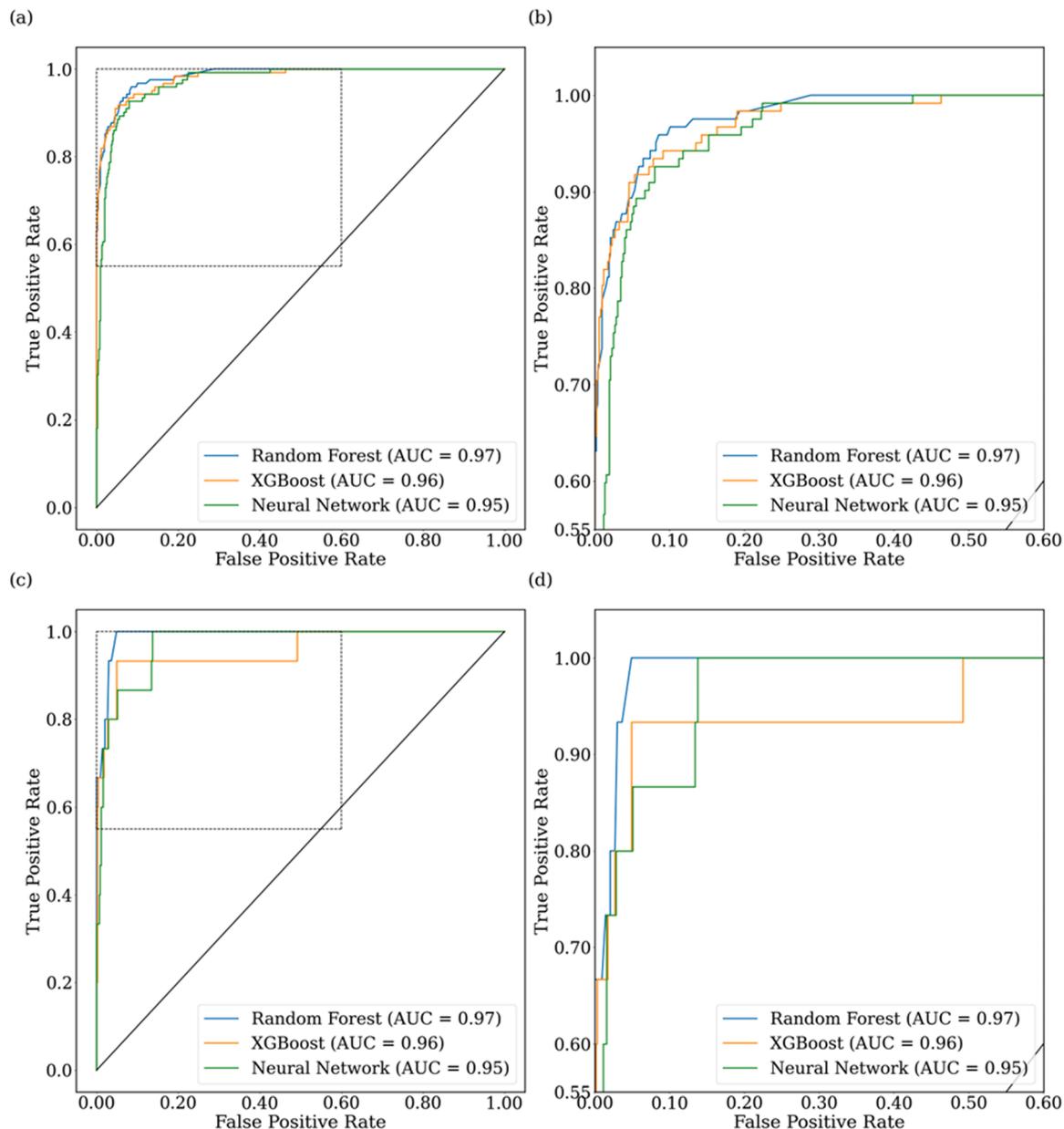


Fig. 9. Receiver operating characteristics (ROC) curves for Dawlish at (a,b) Rig 1 and (c,d) Rig 2. Right panels (b,d) represent the zoomed area in left panels (dashed rectangle).

overtopping hazards. These easing restrictions, coupled with the probability of misclassifying false negatives during the summer periods, further emphasise the importance of addressing such misclassifications during the summer. These findings demonstrate many instances of the random forest correctly classifying overtopping and non-overtopping during the summer; however, the results emphasise the importance of addressing these few instances of misclassification.

3.4. Predicting wave overtopping frequency

Many overtopping predictive tools, such as EurOtop, have been developed for coastal scheme design which focuses on extreme events to test design criteria. These tools have since been adopted for hazard forecasting services. However, nuisance overtopping, under less extreme conditions, is not well predicted by existing tools because the datasets underpinning the development of numerical tools is typically generated using physical models also designed to simulate extreme conditions. The WireWall dataset used here, contains data for all overtopping conditions

(severities) over several seasons, and thus, has enabled the development and testing of three different AI approaches to identify a reliable approach suitable for predicting typical, windy spring tide, conditions, which are likely to be experienced by coastal locations at any time of the year alongside the low probability of extreme conditions.

In Dawlish, the random forest model for both rigs exhibited the strongest predictive performance for estimating the frequency of wave overtopping (Fig. 13; Table 5). There were no statistically significant differences between the random forest predictions and the overtopping frequency estimations in Dawlish for Rig 1 and Rig 2 ($t = 1.37, p = 0.17$; $t = 0.55, p = 0.58$, respectively). The random forest also had very low testing error metrics (Table 5). However, as the overtopping frequency increases the random forest error performance increases (Fig. 13). The XGBoost, particularly for Rig 1, exhibited high predictive performance; however, it had difficulties estimating overtopping frequency, unlike the random forest in Rig 2. Both the random forest and XGBoost exhibited lower mean bias rates than the neural network. In Penzance, the random forest model exhibited the highest predictive performance for estimating

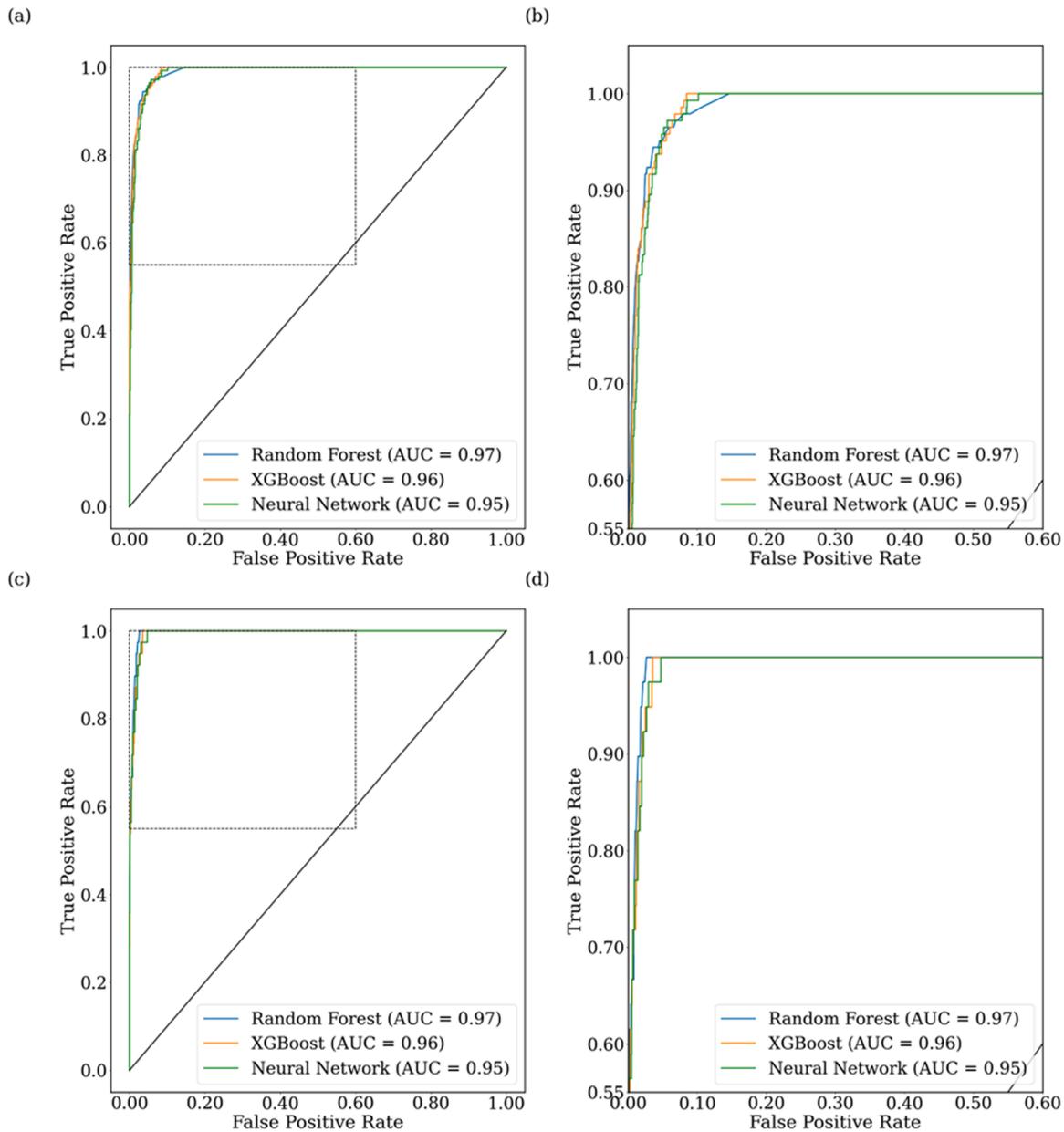


Fig. 10. Receiver operating characteristics (ROC) curves for Penzance at (a,b) rig 1 and (c,d) rig 2. Right panels (b,d) represent the zoomed area in left panels (dashed rectangle).

Table 4
AI performance metrics for estimating overtopping occurrence in Penzance.

Rig	Model	AIC	BIC	F1	Precision	Recall	Accuracy	MCC	Brier Score
1	Random Forest	489	476	0.86	0.86	0.86	97 %	0.81	0.006
	XGBoost	467	460	0.85	0.88	0.82	95 %	0.80	0.025
	Neural Network	386	311	0.74	0.64	0.87	93 %	0.64	0.089
2	Random Forest	345	433	0.75	0.65	0.85	95 %	0.71	0.032
	XGBoost	411	423	0.71	0.77	0.66	92 %	0.71	0.023
	Neural Network	364	334	0.59	0.70	0.74	84 %	0.70	0.034

overtopping frequency within Rig 1 (Fig. 14; Table 6). However, the neural network has the highest predictive performance to predict overtopping frequency in Rig 2 (Table 6). These findings demonstrate the ability of random forests predicting overtopping frequency; however, they indicate the opportunities of using different AI model types to estimate overtopping frequency. There were no statistically significant differences between the random forest overtopping predictions and the

overtopping observational data for Rig 1 and Rig 2 ($t = 2.33$; $p = 0.17$; $t = 0.55$, $p = 0.58$, respectively). The XGBoost exhibited a strong predictive performance with low error metrics for both locations. However, the neural network indicated a low predictive performance with a high error (RMSE = 18.9) for estimating overtopping in Rig 1.

The random forest best predicted overtopping frequency for both rigs at Dawlish and Rig 1 for Penzance. The XGBoost performed the second

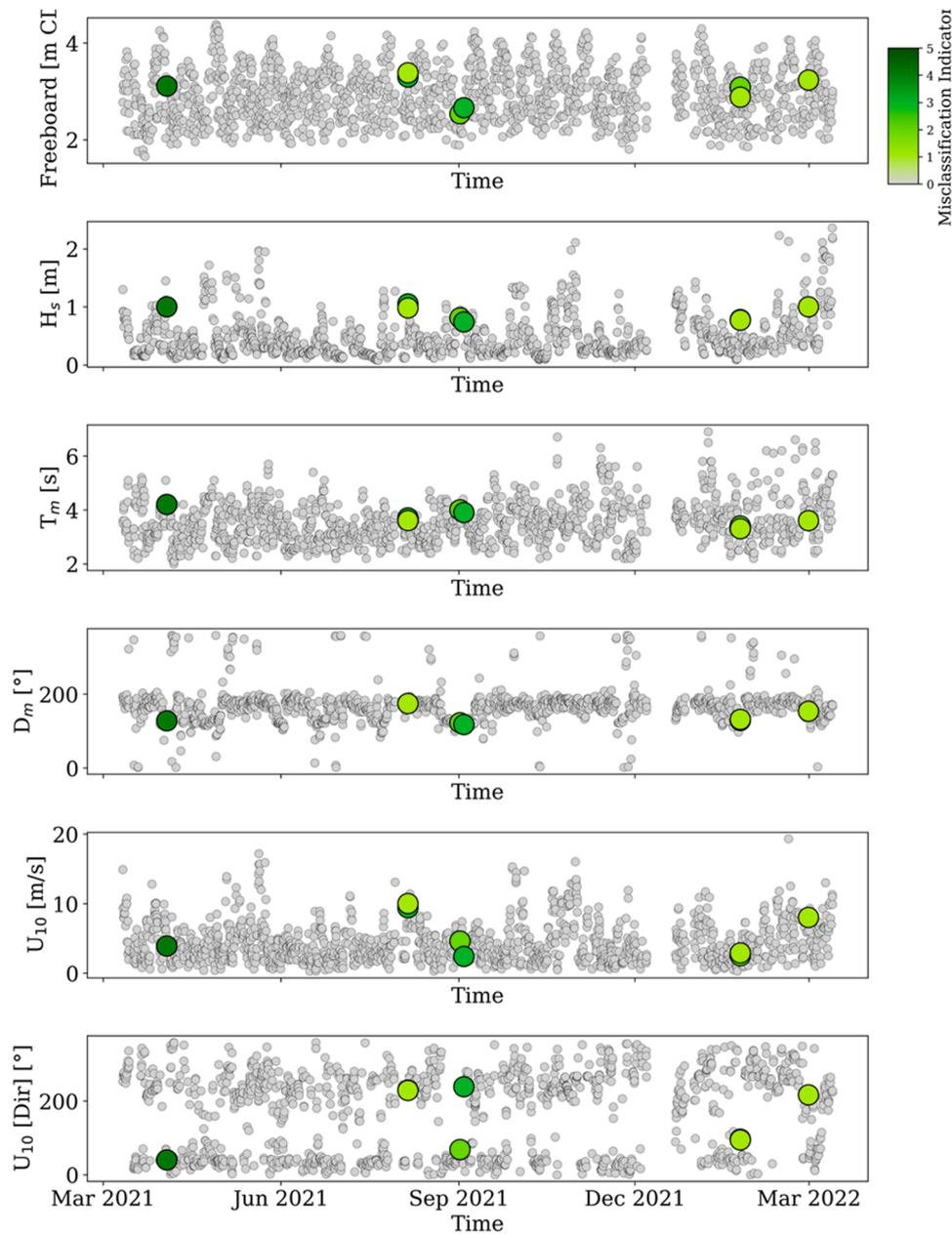


Fig. 11. Random forest false negatives predictions for estimating wave overtopping in Dawlish \pm 3 h either side of high tide. Scatter colours in green represent the number of misclassified events.

best, followed by the neural network. However, the neural network performed the best at Rig 1 in Penzance. These findings demonstrate the neural network exhibits unstable behaviour, with its predictive performance varying significantly (i.e., $R^2 = 0.92$ at Penzance Rig 2 and $R^2 = 0.45$ at both rigs in Dawlish). These varying performances likely result from the imbalances in the training data. Approximately 97 % of overtopping within the training dataset was categorised as non-overtopping, while the remaining 3 % represented overtopping. These data imbalances may cause the neural network to struggle to capture patterns in the minority class (overtopping). At Rig 2, data imbalances were still exhibited (94 % non-overtopping; 4 % overtopping); however, such imbalances were less extreme than Dawlish at both rigs (98 % non-overtopping; 2 % overtopping). The more extreme data imbalances at Dawlish may explain this significant drop in model performance. Random forests ensemble learning captures patterns from different data, reducing the overreliance in the majority class. XGBoost model iteratively learns from mistakes from inaccurately predicting the minority

class, essentially improving over and handling class imbalance. These findings suggest using random forests and XGBoost over neural networks when datasets are imbalanced.

3.5. Comparing AI with EurOtop guidance

The random forest model, which demonstrated having the highest predictive performance for estimating overtopping, was compared against OWWL. To reiterate, OWWL is currently one of the most developed operational suite of models for predicting overtopping (Stokes et al., 2021). OWWL, like many other models rely on static equations derived from EurOtop and omit key parameters like U_{10} for estimating overtopping. The focus of this study is to evaluate the highest performing AI model (i.e., random forest) against OWWL for predicting overtopping. The results confirm the random forest significantly outperforming OWWL for estimating overtopping in Dawlish ($F1 = 0.81$, $F1 = 0.65$, respectively) (Fig. 15). The random forest outperforms OWWL

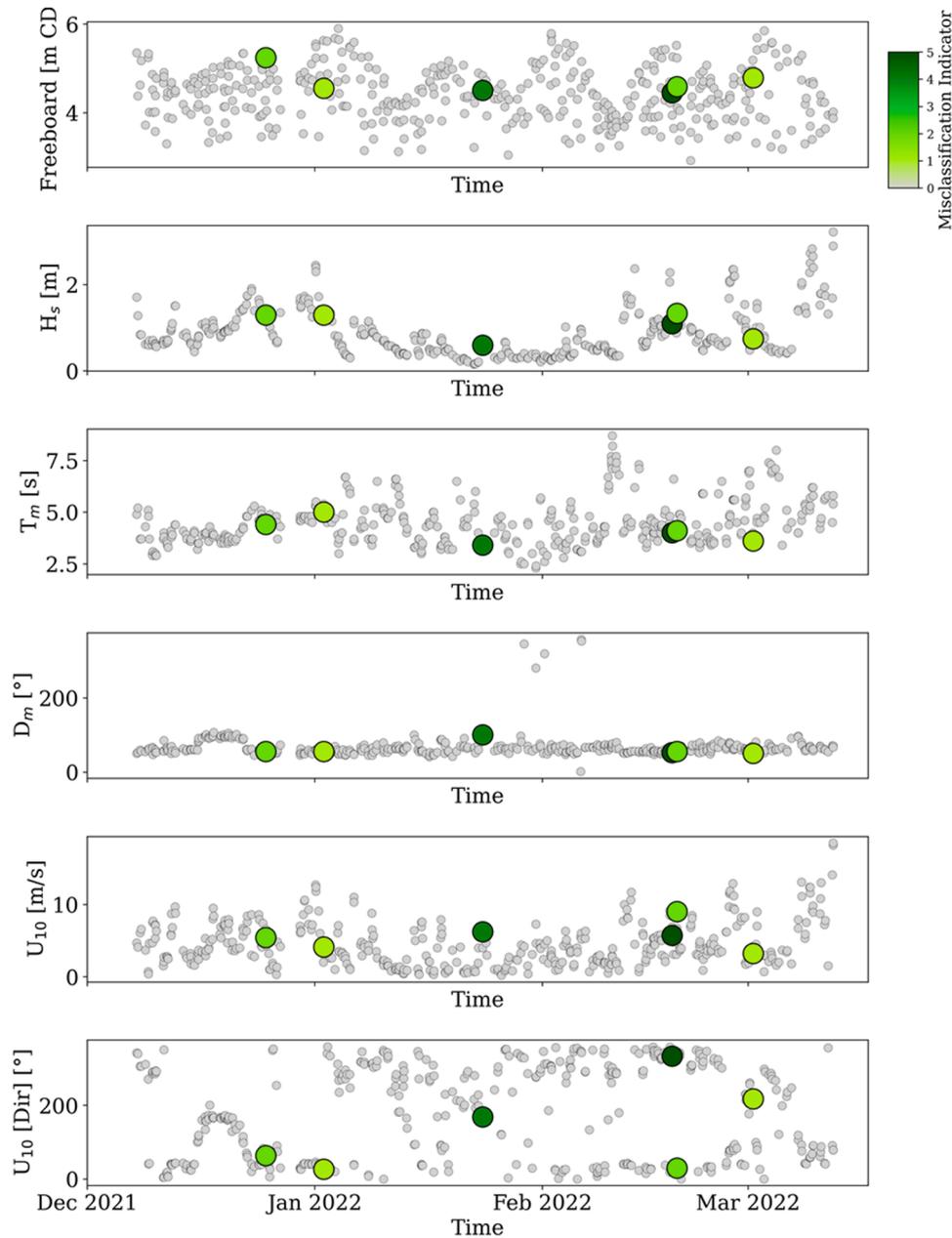


Fig. 12. Random forest false negative predictions for estimating wave overtopping in Penzance ± 3 h either side of high tide. Scatter colours in green represent the number of misclassified events.

slightly for estimating non-overtopping instances ($F1 = 0.98$, $F1 = 0.97$, respectively). These findings indicate the OWWL model having a strong performance for estimating non-overtopping; however, a considerably weaker performance estimating overtopping, which the random forest can reliably estimate. The results indicate the potential importance of incorporating U_{10} and U_{10} Dir when predicting the overtopping class. These are significantly important variables influencing overtopping, that the OWWL model (as well as most current empirical approaches) disregard.

For Penzance, the random forest model significantly outperforms OWWL for estimating overtopping instances ($F1 = 0.86$, $F1 = 0.07$, respectively) (Fig. 15). Moreover, the random outperforms OWWL slightly for estimating the non-overtopping class ($F1 = 0.90$, $F1 = 0.97$, respectively). However, indicated by the standard error bars, there is much improvement needed for the random forest, particularly for Penzance. These error bars likely attribute to the random forest model having challenges addressing false negatives outlined in Section 3.3. The

error bars may be attributed to the smaller training dataset within Penzance, meaning the random forest has fewer data points to learn intricate relationships.

4. Discussion

This study explored the role of AI in predicting wave overtopping occurrence and overtopping frequency (i.e., number of overtopping per 10-min) in Dawlish and Penzance. The results indicated that H_s , R_c , and U_{10} are key variables that influence overtopping. The strong predictive skills presented here demonstrate the significant opportunities of using AI to predict overtopping accurately and precisely. The random forest exhibited the highest predictive performance for estimating occurrences, followed by the XGBoost and the neural network, likely attributing to the ability of the random forest to rank feature importances.

The random forest also exhibited the highest performance for predicting overtopping frequency, followed by the XGBoost and neural

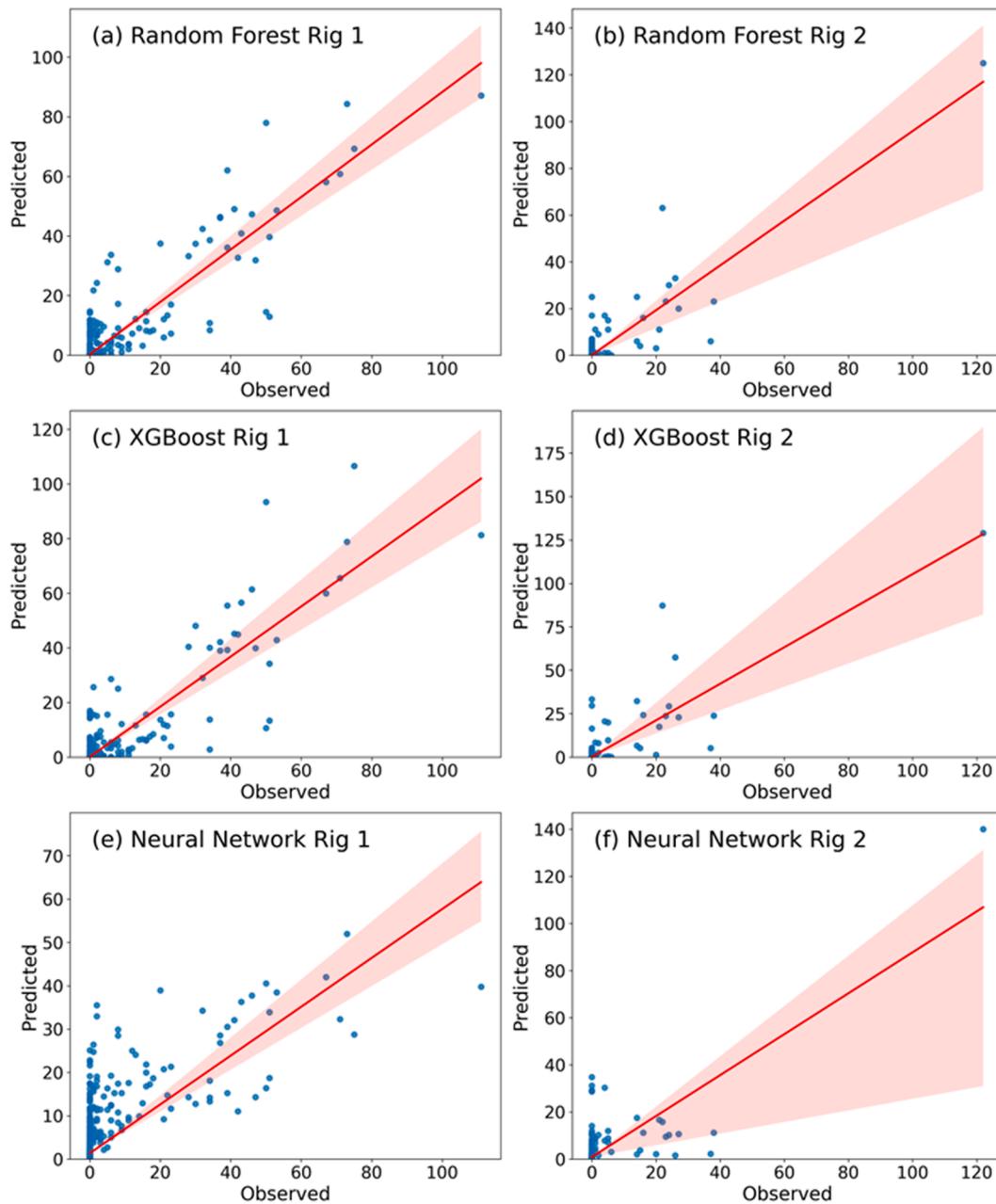


Fig. 13. Regression analysis comparing the random forest predictions against the observational overtopping data in Dawlish: (a) random forest - Rig 1, (b) random forest - Rig 2, (c) XGBoost - Rig 1, (d) XGBoost - Rig 2, (e) neural network - Rig 1, and (f) neural network - Rig 2.

Table 5
AI model performance for estimating wave overtopping frequency in Dawlish.

Rig	Model	R ²	RMSE	MSE	MAE	Mean Bias	T-Test
1	Random Forest	0.81	3.05	9.28	0.66	0.11 %	$t = 1.37; p = 0.17$
	XGBoost	0.77	3.28	10.7	0.69	0.01 %	$t = 0.15; p = 0.88$
	Neural Network	0.45	5.1	26	2.43	0.83 %	$t = 6.42; p < 0.01$
2	Random Forest	0.76	1.90	3.62	0.23	0.08 %	$t = 0.55; p = 0.58$
	XGBoost	0.53	2.65	7	0.30	0.16 %	$t = 1.03; p = 0.30$
	Neural Network	0.45	2.85	8.14	1.08	0.83 %	$t = 5.57; p = 0.675$

network. The neural network presented varying performance results, likely because the neural network was most sensitive to class imbalances present. The central finding of this study is the random forest model outperforms existing models, such as OWWL, which is reliant on EurOtop, for estimating overtopping. This study sheds light around the significant opportunities for AI within the field of coastal engineering and ocean modelling.

4.1. Variables influencing overtopping

H_s, R_c, and U₁₀ were identified as significant variables influencing overtopping. As expected, H_s and R_c are critical variables influencing overtopping (Salaudin and Pearson, 2019; Zheng and Li, 2015). Higher R_c levels increase the wave energy dissipation before waves reach the crest, decreasing the likelihood of waves exceeding the crest height and overtopping. Our study highlights the important effect of U₁₀ on

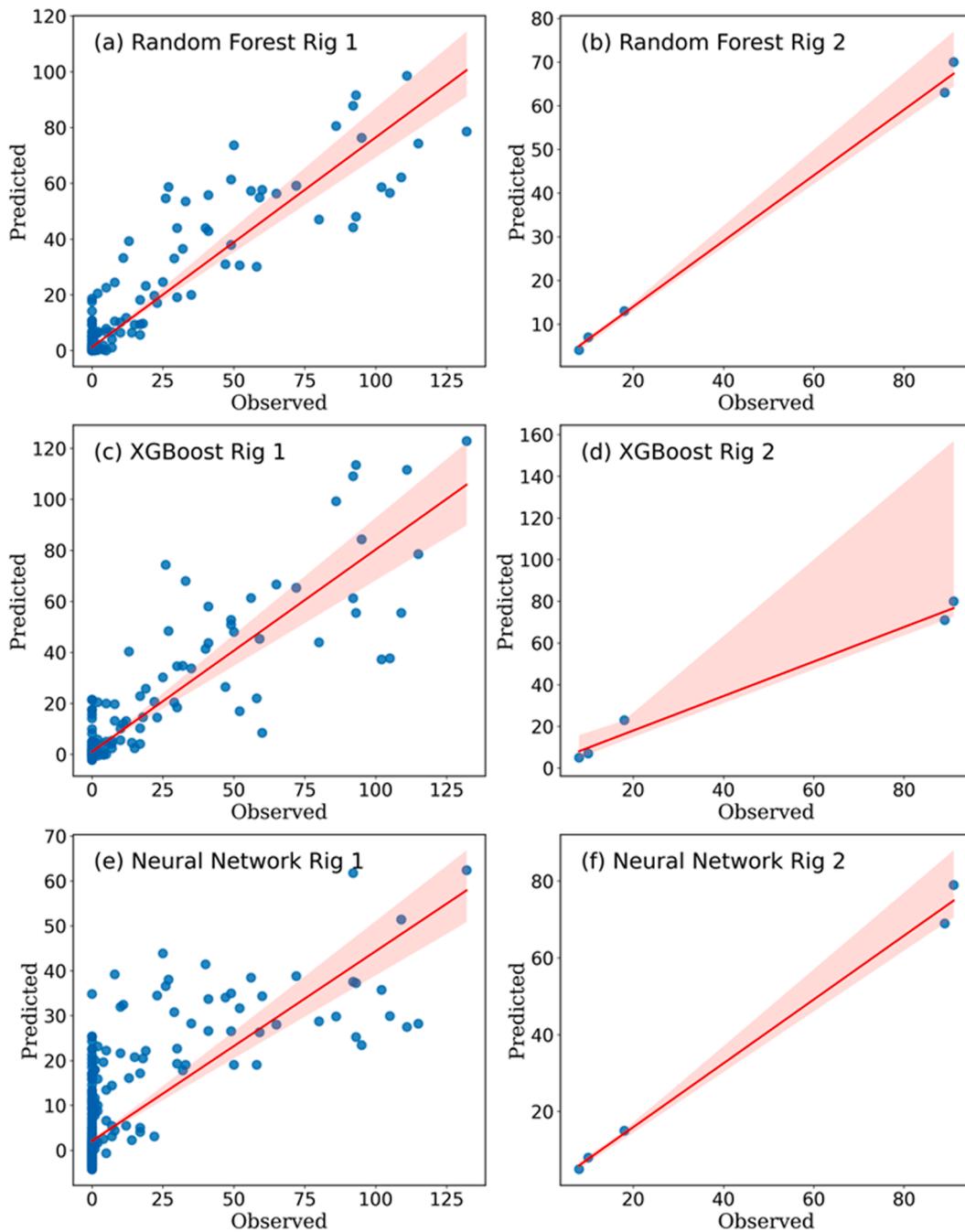


Fig. 14. Regression analysis comparing the random forest predictions against the observational overtopping data in Penzance: (a) random forest - Rig 1, (b) random forest - Rig 2, (c) XGBoost - Rig 1, (d) XGBoost - Rig 2, (e) neural network - Rig 1, and (f) neural network - Rig 2.

overtopping, which is not currently included in existing empirical approaches. Scaling the impact of U_{10} on overtopping is challenging, particularly in laboratory or physical models. Wind effects on wave transformation (from deep water to the structure tow) are highly case-specific, causing significant location-based variability that empirical models struggle to account for. Recent work by [Van Gent et al. \(2024\)](#) addresses the challenge, demonstrating that wind can increase overtopping discharges at seawalls by a factor of up to four and provides an empirical guideline for incorporating this effect. While this random forest does not yet incorporate this guidance directly, acknowledging the influence of U_{10} offers valuable context and may significantly improve the predictive performance of estimating wave overtopping, as demonstrated from our results, which show the exclusion of wind related variables significantly decreasing the random forest predictive

performance.

Several studies have investigated the effects of onshore wind on wave overtopping, with some prediction (static) formula proposed (e.g., [De Chowdhury et al., 2019](#); [Di Leo et al., 2022](#); [Durbridge, 2021](#); [Murakami et al., 2019](#); [Pullen et al., 2009](#)). [Di Leo et al. \(2022\)](#) shows onshore wind significantly enhancing overtopping, particularly with overtopping discharges below 1L/s/m, with the study proposing a wind factor influence on overtopping. [Pullen et al. \(2009\)](#) demonstrated high onshore wind speeds (15–20m/s) increase overtopping discharges by dispersing water farther inland, suggesting an empirical “spray transport factor” that could triple discharge under high wind conditions. U_{10} is very important, particularly for Dawlish Rig 2 (i.e. inland rig), demonstrating the importance of U_{10} dispersing water inland. Additionally, wind direction almost equals the importance of H_s in wave overtopping

Table 6
AI model performance for estimating wave overtopping frequency in Penzance.

Rig	Model	R ²	RMSE	MSE	MAE	Mean Bias	T-Test
1	Random Forest	0.84	7.87	6.91	2.53	0.35 %	$t = 2.33; p = 0.434$
	XGBoost	0.81	8.70	7.54	2.74	0.44 %	$t = 0.12; p = 0.322$
	Neural Network	0.52	13.7	18.9	6.22	1.22 %	$t = 1.33; p < 0.01$
2	Random Forest	0.84	15.3	2.33	11.8	0.07 %	$t = 3.42; p = 0.233$
	XGBoost	0.80	9.88	9.71	8	0.33 %	$t = 5.66; p = 0.122$
	Neural Network	0.92	1.22	1.22	3	0.18 %	$t = 2.22; p = 0.543$

in Rig 2 in Penzance, again demonstrating the significance of including the wind in locations affected by bimodal conditions, such as the ones presented in our study.

Wind shadowing (contribution) can decrease (enhance) H_s and water levels, and ultimately overtopping (Young et al., 2012). The results reveal the ranking of these variables is consistent across the two study locations. These findings reveal the opportunities of using these selected few variables to estimate overtopping across large spatial scales accurately. However, this study only investigated two locations, so the analogy of using these three variables to predict overtopping across multiple locations must require verification and further testing (Yalta and Jenal, 2009). Our study reveals T_m as a minor variable influencing wave overtopping, which agrees with Buccino et al. (2023), showing that in the surf zone, wave energy is linked with T_m , creating a false impression that T_m influences overtopping, where in reality, the study shows overtopping is influenced by wave height and water levels near the seawall. For our study sites, it is thought that the seawalls are more exposed to wind sea than swell waves due to the structure orientation and bay geometry relative to the prevailing swell direction. This may be a factor influencing the importance of T_m .

This study deviates from the EurOtop model, which employs H_s measured directly at the seawall toe. Given the limited availability and operational challenges of direct H_s measurements from the seawall toe, H_s was recorded from nearshore coastal waters using wave buoys. This

modification of obtaining H_s from nearshore coastal waters is supported by Lashley et al. (2023), who advocate using deep-water wave parameters for training predictive models to estimate overtopping. This study obtained wave characteristics from deeper waters, which may explain better how these variables interrelate to influence wave overtopping and strengthen the robustness of the dataset used for training these AI models. While some relationships between H_s , R_c , and U_{10} are well understood and captured in empirical rules, the complexity and interactions between these variables are still poorly understood (Chiapponi et al., 2020). AI offers an opportunity to be able to develop predictive tools where the detailed process parameterisation is not yet available.

The results highlight that when considering the interaction effects of many wave and weather conditions, the H_s , R_c , and U_{10} when interacting is quantitatively significant in terms of their influence on overtopping. These findings are significant because, for the first time, this study identifies which variables influence overtopping when considering these complex interaction effects. Moreover, the results show the ranking of these variables is generally consistent across the two study locations, revealing an opportunity of using these select few variables to estimate overtopping across large spatial scales accurately. However, this study only investigated two locations, so the analogy of using these three variables to predict overtopping across multiple locations must require verification and further testing (Yalta and Jenal, 2009). Moreover, consideration of beach level, especially for climate assessment, is another important parameter that should be considered in future development (Stokes et al., 2021).

4.2. Predicting wave overtopping and frequency

Random forests outperformed the other AI models for predicting the frequency of overtopping. This outperformance may be attributed to the random forest’s inbuilt features around “sampling with replacement” and the “out-of-bag” data (Hong et al., 2020; Mitchell, 2011; Özçift, 2011; Salles et al., 2015). Sampling with replacement reduces the likelihood of individual decision trees overspecialising on a particular pattern or noise within the dataset, decreasing the probability of overfitting. Identifying less important features may also simplify the random forest, causing the random forest to prioritise feature variables that lead to the greatest reduction in the leaf node’s Gini Impurity, and ignoring

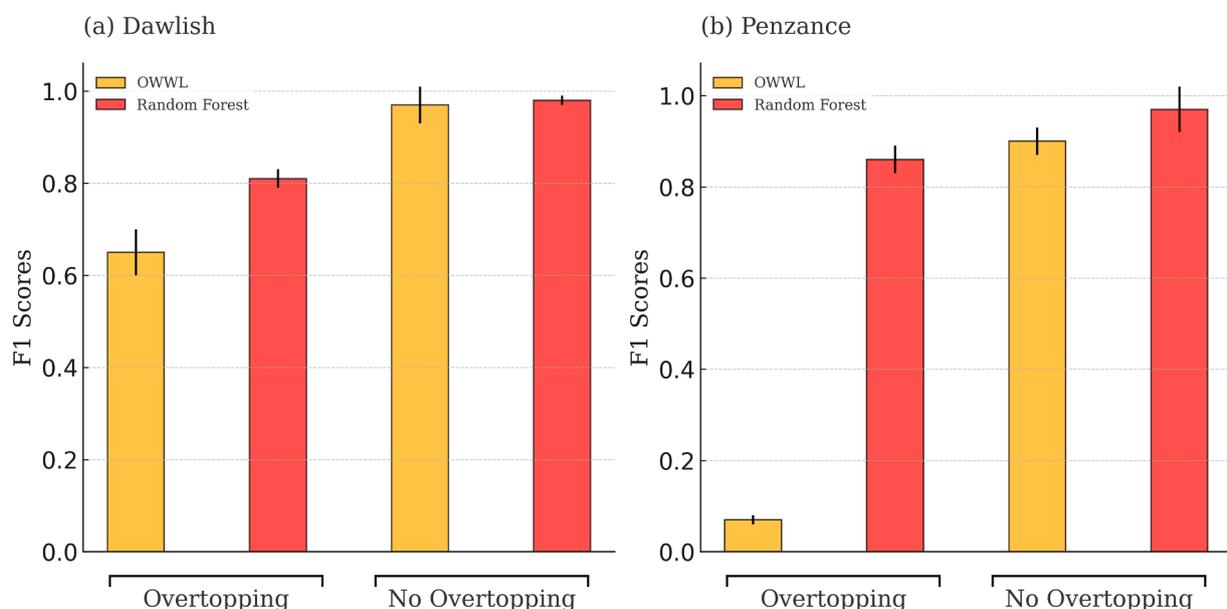


Fig. 15. Comparing the random forest and OWWL predictions for overtopping and non-overtopping classes in (a) Dawlish and (b) Penzance. The bars indicate the standard error for each prediction at the 95 % confidence interval. Orange bars illustrate the OWWL, and red denote the random forest.

redundant and unnecessary features (González et al., 2015). The implications of sampling with replacement and the out-of-bag samples place random forests at a unique advantage, allowing them to specialise in different data subsets and identifying key variables influencing the model predictions potentially better than the neural network and XGBoost. Habib et al. (2023) showed similar findings, which analysed and compared the predictive performance of random forests, gradient-boosted decision trees, and neural networks for estimating overtopping discharge at vertical seawalls, again showing the random forests outperforming the other model types.

There are several reasons why sampling with replacement and this out-of-bag data is relevant for predicting wave overtopping. The variables influencing wave overtopping are highly correlated and interrelated, and models that can distinguish between variables influencing wave overtopping and those creating noise within the data is fundamental. The results revealed many variables, like the T_m and D_m , were not as statistically important regarding their influence on wave overtopping. The ability of the random forest model to place less priority on these variables, unlike the different AI models used, may explain why the random forest outperformed the others.

The neural network demonstrated prediction instability, likely due to limited training data and challenges with class imbalances. The EurOtop manual includes a neural network tool for predicting wave overtopping, raising questions about its predictive stability under similar conditions. Although examining EurOtop's stability is beyond this study's scope, EurOtop has undergone significant improvements, including refining the training dataset and hyperparameter tuning, to improve stability (Zanuttigh et al. 2016).

Meanwhile, Habib et al. (2023) demonstrated that, despite these improvements, decision tree models - particularly random forests - outperformed neural networks in predictive accuracy. Their study did not reveal that the EurOtop neural network produced unstable predictions; instead, it confirmed that the decision tree architecture consistently outperformed tuned neural networks. These results suggest that, despite refinements, the EurOtop neural network may not match the predictive performance of random forests. Future research should investigate the instability of the EurOtop neural network for examining overtopping discharges despite its model refinements.

4.3. The challenges of false negatives

The results revealed that the random forest generated false negative predictions in a few instances. False negatives occur when the random forest reports no overtopping during an observed overtopping event (Bold et al., 2022). The random forest false negatives always corresponded with very low overtopping frequency (< 4 waves overtopping in a 10-min window), and these occur within a narrow threshold range of some of the key feature variables like H_s and R_c .

Gupta et al. (2021) recommends adjusting the model's recall level to minimise false negative predictions. This recall adjustment involves changing the random forest threshold, specifically decreasing the threshold level (Berger and Guda, 2020). This decrease in threshold increases the susceptibility of the random forest classifying more false positives (Arora et al., 2016). Lowering the threshold causes the random forest to classify more instances of overtopping, which inversely increases the probability of incorrectly classifying overtopping during non-overtopping (false positives) (Arora et al., 2016). If identifying wave overtopping is more of a priority of saving human life over the economic implications of incorrectly labelling false positives, then this approach of increasing the model recall by decreasing the threshold would be appropriate. However, in many instances, reporting false positives can result in the unnecessary closure of infrastructure and could be viewed as unfavourable (Arroita et al., 2017).

By avoiding the unnecessary classification of false positives and false negatives without adjusting the model threshold levels, a more practical approach would be to either increase the model training dataset or to

enhance the data sampling to train these random forest models. This study trained these models using an 80 % training to 20 % testing ratio; however, with a more balanced and larger dataset, a slight increase in the training ratio by 5 % could potentially reduce the likelihood of reporting false positives, without necessarily increasing the reporting of false negatives (Uçar et al., 2020). This practical approach to predict hazards and identify key processes leading to hazardous overtopping conditions could have significant implications for the field of coastal engineering and risk assessment.

The random forest was trained on data from only one storm season (2021–2022), primarily capturing high-energy storm events. This lack of training across multiple storm seasons may decrease the robustness of this random forest to capture inter-annual variability in overtopping, particularly during varying storm conditions. Expanding the training dataset in various storm seasons (both high and low-energy storm events) could enhance the random forest generalisation for predicting wave overtopping. Additionally, the water levels at Dawlish were measured within a nearby estuary rather than directly at the overtopping measuring location. This setup could introduce discrepancies in the water level measurements. Like the wave buoy data, the machine learning tools will have learned about the relationship between water level at the tide gauge (i.e. within the estuary) and the occurrence of overtopping at the seawall. This is not necessarily a problem, as long as any predictions made using the machine learning models use an accurate prediction of water level within the estuary (e.g. include tidal and fluvial components) to drive the predictions.

5. Conclusions

This study investigated the role of AI predicting coastal wave overtopping occurrence and frequency (number of overtopping events per 10-min intervals) in two locations in southwest England (Dawlish and Penzance). The random forest models have the highest predictive performance and lowest error metrics for estimating wave overtopping and non-overtopping occurrence with a 97 % accuracy. These models have also successfully identified variables that are statistically significant for influencing wave overtopping and frequency. Moreover, this study indicates the importance of U_{10} and U_{10} Dir for predicting wave overtopping, which are neglected in most current approaches (e.g., EurOtop). The AI models presented here provide a reliable, computationally efficient, predictive tool suitable for hazards forecasting services. These findings go beyond the existing approach of estimating overtopping, which are less accurate and time-consuming, and normally rely on processed-based models. The major implication is that this approach could be implemented anywhere if field observations of overtopping are available.

These AI models can be developed to use nearshore data as an input parameter, which are more readily available from operational monitoring networks and national numerical forecast services. This removes the need to accurately parameterise wave shoaling and breaking, across varying beach-structure profiles to obtain the wave-water level and wave conditions at a structures toe, as required by EurOtop. This makes AI models very suitable for early warning of coastal wave overtopping events in a changing climate where sea levels are rising and existing infrastructure ageing.

Data and code availability

All the training datasets for the models, including the model code alone with any supplementary material can be found using Zenodo: <https://zenodo.org/records/12759748>

CRedit authorship contribution statement

Michael McGlade: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Nieves G. Valiente:** Writing – review &

editing, Supervision, Project administration, Investigation, Funding acquisition, Formal analysis. **Jennifer Brown:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Formal analysis, Data curation. **Christopher Stokes:** Writing – review & editing, Software, Formal analysis, Data curation. **Timothy Poate:** Writing – review & editing, Data curation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Michael McGlade reports financial support was provided by UK Research and Innovation Natural Environment Research Council. Michael McGlade reports financial support was provided by Met Office. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the Natural Environment Research Council (NERC) and the Met Office through their SPLASH project (NE/Z503423/1). The National Oceanography Centre (NOC) (NE/Z503435/1) also provided WireWall information and data. We would like to thank all the reviewers, both internally within this project and externally, for significantly enhancing the quality of this research.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.ocemod.2025.102510](https://doi.org/10.1016/j.ocemod.2025.102510).

Data availability

Data will be made available on request.

References

- Adams, K., Heidarzadeh, M., 2023. Extratropical cyclone damage to the seawall in Dawlish, UK: eyewitness accounts, sea level analysis and numerical modelling. *Natural Hazards* 116, 637–662. <https://doi.org/10.1007/s11069-022-05692-2>.
- Alshahri, A.H., Elbisy, M.S., 2022. Prediction of wave overtopping discharges at coastal structures using artificial neural networks and support vector machine techniques. *Geomate J.* 23, 56–62. <https://geomatejournal.com/geomate/article/view/3582>.
- Alvarez, A., Figuero, A., Rodríguez-Yáñez, S., Sande, J., Peña, E., Rosa-Santos, P., Rabuñal, J., 2024. Deep learning-based wave overtopping prediction. *Appl. Sci.* 14, 2611. <https://doi.org/10.3390/app14062611>.
- Arora, M., Kanjilal, U., Varshney, D., 2016. Evaluation of information retrieval: precision and recall. *Int. J. Indian Culture Bus. Manag.* 12, 224–236. <https://doi.org/10.1504/IJICBM.2016.074482>.
- Balraj, G., Victoire, A.A., S., J., Victoire, A., 2022. Variational mode decomposition combined fuzzy - twin support vector machine model with deep learning for solar photovoltaic power forecasting. *PLoS One* 17, 1–12. <https://doi.org/10.1371/journal.pone.0273632>.
- Barnard, P.L., van Ormondt, M., Erikson, L.H., Eshleman, J., Hapke, C., Ruggiero, P., Adams, P., Foxgrover, A.C., 2014. Development of the Coastal Storm Modeling System (CoSMoS) for predicting the impact of storms on high-energy, active-margin coasts. *Natural Hazards* 74, 1095–1125. [10.1007/s11069-014-0642-8](https://doi.org/10.1007/s11069-014-0642-8).
- Berger, A., Guda, S., 2020. Threshold optimization for F measure of macro-averaged precision and recall. *Pattern. Recognit.* 102, 107250. <https://doi.org/10.1016/j.patcog.2020.107250>.
- Buccino, M., Di Leo, A., Tuozzo, S., Lopez, L.F.C., Calabrese, M., Dentale, F., 2023. Wave overtopping of a vertical seawall in a surf zone: a joint analysis of numerical and laboratory data. *Ocean Eng.* 288, 116144. <https://doi.org/10.1016/j.oceaneng.2023.116144>.
- Den Bieman, J.P., van Gent, M.R., van den Boogaard, H.F., 2021. Wave overtopping predictions using an advanced machine learning technique. *Coast. Eng.* 166, 103830. <https://doi.org/10.1016/j.coastaleng.2020.103830>.
- Bold, R., Al-Khateeb, H., Ersotelos, N., 2022. Reducing false negatives in ransomware detection: a critical evaluation of machine learning algorithms. *Appl. Sci.* 12, 12941. <https://doi.org/10.3390/app122412941>.
- Brown, J., Yelland, M., Pascal, R., Pullen, T., Cardwell, C., Jones, D., Pinnell, R., Silva, E., Balfour, C., Hargreaves, G., Barry, M., Paul, B., Prime, T., Burgess, J., Eastwood, L., Gold, M., Cai, B., Farrington, B., 2020. WireWall - a new approach to measuring coastal wave hazard. *NOC* 115. <https://nora.nerc.ac.uk/id/eprint/528538>.
- Brown, J., Yelland, M., Masselink, G., Poate, T., Stokes, K., Pascal, R., Jones, D., Cardwell, C., Walk, J., Marin, B., Ganderton, P., Darroch, L., Gardner, R., 2022. Coastal wave overtopping: new Nowcast and monitoring technologies. *Coastal Eng. Proc.* 37, 1. <https://doi.org/10.9753/icce.v37.papers.1>.
- Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 785–794.
- Chiapponi, L., Addona, F., Diaz-Carrasco, P., Losada, M.A., Longo, S., 2020. Statistical analysis of the interaction between wind-waves and currents during early wave generation. *Coast. Eng.* 159, 103672. <https://doi.org/10.1016/j.coastaleng.2020.103672>.
- Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 21, 1–13. <https://doi.org/10.1186/s12864-019-6413-7>.
- Dawson, D., Shaw, J., Gehrels, W.R., 2016. Sea-level rise impacts on transport infrastructure: the notorious case of the coastal railway line at Dawlish, England. *J. Transp. Geogr.* 51, 97–109. <https://doi.org/10.1016/j.jtrangeo.2015.11.009>.
- De Chowdhury, S., Causon, D., Qian, L., Mingham, C., Chen, H., Lin, Z., Zhou, J.G., Pullen, T., Silva, E., Hu, K., 2019. Investigation of wind effects on wave overtopping at sea defences. In: *Proceedings of the Coastal Structures 2019, Hannover, Germany, 29 September–2 October 2019; Bundesanstalt für Wasserbau*. S: Karlsruhe, Germany, pp. 841–850.
- Di Leo, A., Dentale, F., Buccino, M., Tuozzo, S., Pugliese Carratelli, E., 2022. Numerical analysis of wind effect on wave overtopping on a vertical seawall. *Water*. (Basel) 14 (23), 3891. <https://doi.org/10.3390/w14233891>.
- Durbridge, S., 2021. The effect of onshore wind on wave overtopping of a vertical sea wall.
- Elbisy, M.S., 2023. Machine learning techniques for estimating wave overtopping discharges at coastal structures. *Ocean Eng.* 273, 113972. <https://doi.org/10.1016/j.oceaneng.2023.113972>.
- EurOtop, 2018. Manual on wave overtopping of sea defences and related structures: an overtopping Manual largely based on European research, but for worldwide application. www.overtopping-manual.com.
- González, S., Herrera, F., García, S., 2015. Monotonic random forest with an ensemble pruning mechanism based on the degree of monotonicity. *New. Gener. Comput.* 33, 367–388. <https://doi.org/10.1007/s00354-015-0402-4>.
- Güney, Y., Bozdoğan, H., Arslan, O., 2021. Robust model selection in linear regression models using information complexity. *J. Comput. Appl. Math.* 398, 113679. <https://doi.org/10.1016/j.cam.2021.113679>.
- Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R.K., Kumar, P., 2021. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol. Divers.* 25, 1315–1360. <https://doi.org/10.1007/s11030-021-10217-3>.
- Habib, M.A., O'Sullivan, J.J., Abolfathi, S., Salauddin, M., 2023a. Enhanced wave overtopping simulation at vertical breakwaters using machine learning algorithms. *PLoS One* 18, 1–14. <https://doi.org/10.1371/journal.pone.0289318>.
- Habib, M.A., Abolfathi, S., O'Sullivan, J.J., Salauddin, M., 2023b. Prediction of wave overtopping rates at sloping structures using artificial intelligence. In: *Proceedings of the 40th IAHR World Congress. Rivers—Connecting Mountains and Coasts*, pp. 404–413. <https://doi.org/10.3850/978-90-833476>.
- Humphrey, A., Kuberski, W., Bialek, J., Perrakis, N., Cools, W., Nuytens, N., Elakhrass, H., Cunha, P.A.C., 2022. Machine-learning classification of astronomical sources: estimating F1-score in the absence of ground truth. *Monthly Notices Royal Astron. Soc.* 577, 116–120. <https://doi.org/10.1093/mnras/577/1/116>.
- Jane, R., Simmonds, D., Gouldby, B., Simm, J., Valle, L.D., Raby, A., 2018. Coasts, marine structures and breakwaters 2017: realising the potential. Improving the representation of the fragility of coastal structures. *Brunei Seawall (Dawlish, UK)* 1, 887–899. <https://doi.org/10.1680/cmsb.63174.0887>.
- Janitzka, S., Hornung, R., 2018. Plo one. On the overestimation of random forests out-of-bag error. 13, 1–14. <https://doi.org/10.1371/journal.pone.0201904>.
- Janssen, T.T., Battjes, J.A., 2007. A note on wave energy dissipation over steep beaches. *Coast. Eng.* 54, 711–716. <https://doi.org/10.1016/j.coastaleng.2007.05.006>.
- Jennath, A., Paul, S., 2024. A multi-risk approach for projecting climate change-associated coastal flood, applied to India. *Natural Hazards* 1, 1–20. <https://doi.org/10.1007/s11069-024-06420-8>.
- Kulp, S.A., Strauss, B.H., 2019. New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. *Nat. Commun.* 1, 4844. <https://doi.org/10.1038/s41467-019-12808-z>.
- Lashley, C.H., Brown, J.M., Yelland, M.J., van der Meer, J.W., Pullen, T., 2023. Comparison of deep-water-parameter-based wave overtopping with wirewall field measurements and social media reports at Crosby (UK). *Coast. Eng.* 179, 104241. <https://doi.org/10.1016/j.coastaleng.2022.104241>.
- Lerma, A.N., Bulteau, T., Elineau, S., Paris, F., Durand, P., Anselme, B., Pedreros, R., 2018. Natural Hazards and Earth System Sciences. High-resolution marine flood modelling coupling overflow and overtopping processes: framing the hazard based on historical and statistical approaches. 18, 207–229. <https://doi.org/10.5194/nhess-18-207-2018>.
- McGlade, M., Valiente, N., Brown, J., Stokes, C., Poate, T., 2024. Investigating Appropriate Artificial Intelligence Approaches to Reliably Predict Coastal Wave Overtopping and Identify Process Contributions. *Ocean Modelling*.
- Murakami, K., Maki, D., Ogino, K., 2019. Effect of wind velocity on wave overtopping. In: *Proceedings of the 10th International Conference on Asian and Pacific Coasts (APAC 2019), Hanoi, Vietnam, 25–28 September*.
- NNRCMP, 2024. National Network of Regional Coastal Monitoring Programmes. <https://coastalmonitoring.org/> [Accessed 1st June 2024].

- Nti, I.K., Nyarko-Boateng, O., Aning, J., 2021. Performance of machine learning algorithms with different K values in K-fold cross-validation. *Int. J. Inf. Technol. Comput. Sci.* 13, 61–71. <https://doi.org/10.5815/ijitcs.2021.06.05>.
- Nwankpa, C.E., 2020. Advances in optimisation algorithms and techniques for deep learning. *Adv. Sci. Technol. Eng. Syst. J.* 5, 563–577. <https://doi.org/10.25046/aj050570>.
- Özçift, A., 2011. Random forests ensemble classifier trained with data resampling strategy to improve cardiac arrhythmia diagnosis. *Comput. Biol. Med.* 41, 265–271. <https://doi.org/10.1016/j.combiomed.2011.03.001>.
- Park, K.M., Shin, D., Chi, S.D., 2019. Variable chromosome genetic algorithm for structure learning in neural networks to imitate human brain. *Applied Sciences* 9, 3176–3177. <https://doi.org/10.3390/app9153176>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., 2011. *Scikit-learn: Mach. learn. Python 12*.
- Pullen, T., Allsop, W., Bruce, T., Pearson, J., 2009. Field and laboratory measurements of mean overtopping discharges and spatial distributions at vertical seawalls. *Coast. Eng.* 56 (2), 121–140. <https://doi.org/10.1016/j.coastaleng.2008.03.011>.
- Prieto, A., Prieto, B., Ortigosa, E.M., Ros, E., Pelayo, F., Ortega, J., Rojas, I., 2016. Neural networks: an overview of early research, current frameworks, and new challenges. *Neurocomputing.* 214, 242–268. <https://doi.org/10.1016/j.neucom.2016.06.014>.
- Rae, M., Van Hove, M., Göpfert, A., 2023. Effect of climate related flooding on health and healthcare worldwide. *Br. Med. J.* 381, 1131. <https://doi.org/10.1136/bmj.p1331>.
- Ramezan, C.A., Warner, T.A., Maxwell, A.E., Price, B.S., 2021. Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. *Remote Sens. (Basel)* 13, 368–369. <https://doi.org/10.3390/rs13030368>.
- Rigatti, S.J., 2017. Random forests. *J. Insur. Med.* 47, 31–39. <https://doi.org/10.17849/insm-47-01-31-39.1>.
- Salles, T., Gonçalves, M., Rodrigues, V., Rocha, L., 2015. Broof: exploiting out-of-bag errors, boosting and random forests for effective automated classification. In: *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval*, pp. 353–362.
- Salauddin, M., Pearson, J.M., 2019. Wave overtopping and toe scouring at a plain vertical seawall with shingle foreshore: a physical model study. *Ocean Eng.* 171, 286–299. <https://doi.org/10.1016/j.oceaneng.2018.11.011>.
- Schmidhuber, J., 2015. Deep learning in neural networks: an overview. *Neural Netw.* 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>.
- Shen, Z., Yang, H., Zhang, S., 2021. Neural network approximation: three hidden layers are enough. *Neural Netw.* 141, 160–173. <https://doi.org/10.1016/j.neunet.2021.04.011>.
- Stokes, K., Poate, T., Masselink, G., King, E., Saulter, A., Ely, N., 2021. Forecasting coastal overtopping at engineered and naturally defended coastlines. *Coast. Eng.* 164, 103827. <https://doi.org/10.1016/j.coastaleng.2020.103827>.
- Suzuki, T., Altomare, C., Veale, W., Verwaest, T., Trouw, K., Troch, P., Zijlema, M., 2017. Efficient and robust wave overtopping estimation for impermeable coastal structures in shallow foreshores using SWASH. *Coast. Eng.* 122, 108–123. <https://doi.org/10.1016/j.coastaleng.2017.01.009>.
- Uçar, M.K., Nour, M., Sindi, H., Polat, K., 2020. The effect of training and testing process on machine learning in biomedical datasets. *Math. Probl. Eng.* 1, 28336236. <https://doi.org/10.1155/2020/28336236>.
- Van Dongeren, A., Ciavola, P., Martinez, G., Viavattene, C., Bogaard, T., Ferreira, O., Higgins, R., McCall, R., 2018. Introduction to RISC-KIT: resilience-increasing strategies for coasts. *Coast. Eng.* 132, 2–9. <https://doi.org/10.1016/j.coastaleng.2017.10.007>.
- Van Gent, M.R.A., van der Bijl, R.J., Wolters, G., Wüthrich, D., 2024. The maximum influence of wind on wave overtopping at seawalls with crest elements. *Coasts, Marine Structures and Breakwaters 2023: Resilience and Adaptability in a Changing Climate*. Emerald Publishing Limited, pp. 1007–1021.
- Vousdoukas, M.I., Mentaschi, L., Hinkel, J., Ward, P.J., Mongelli, I., Fang, J., Feyen, L., 2018. Economic motivation for raising coastal flood defences in Europe. *Nat. Clim. Chang.* 8, 776–780. <https://doi.org/10.1038/s41558-018-0266-7>.
- Whittaker, C.N., Fitzgerald, C.J., Raby, A.C., Taylor, P.H., Borthwick, A.G.L., 2018. Extreme coastal responses using focused wave groups: overtopping and horizontal forces exerted on an inclined seawall. *Coast. Eng.* 140, 292–305. <https://doi.org/10.1016/j.coastaleng.2018.08.004>.
- Wongvorachan, T., He, S., Bulut, O., 2023. A comparison of under sampling, oversampling, and SMOTE methods for dealing with imbalanced classification in educational data mining. *Information* 14, 54. <https://doi.org/10.3390/info14010054>.
- Wright, L.G., Onodera, T., Stein, M.M., Wang, T., Schachter, D.T., Hu, Z., McMahon, P.L., 2022. Deep physical neural networks trained with backpropagation. *Nature* 601, 1–13. <https://doi.org/10.5281/zenodo.4719150>.
- Yacouby, R., Axman, D., 2020. Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. *Proc. First Workshop Eval. Comparison NLP Syst.* 14, 79–91. <https://doi.org/10.18653/v1/2020>.
- Yang, M., Lim, M.K., Qu, Y., Li, X., Ni, D., 2023. Deep neural networks with L1 and L2 regularization for high dimensional corporate credit risk prediction. *Expert. Syst. Appl.* 213, 118873. <https://doi.org/10.1016/j.eswa.2022.118873>.
- Yelland, J.Brown, Cardwell, C., Jones, D., Pascal, R., Pinnell, R., Pullen, T., Silva, E., 2023. A system for in-situ, wave-by-wave measurements of the speed and volume of coastal overtopping. *Commun. Eng.* 2, 9 s44172-023-00058-3.
- Young, I.R., Vinoth, J., Zieger, S., Babanin, A.V., 2012. Investigation of trends in extreme value wave height and wind speed. *J. Geophys. Res.* 117, 1–14. <https://doi.org/10.1029/2011JC007753>.
- Yue, P., Shangguan, B., Hu, L., Jiang, L., Zhang, C., Cao, Z., Pan, Y., 2022. Towards a training data model for artificial intelligence in earth observation. *Int. J. Geogr. Inf. Sci.* 36, 2113–2137. <https://doi.org/10.1080/13658816.2022.2087223>.
- Zanuttigh, B., van der Meer, J.W., Pezzutto, P., Andersen, T.L., 2016. Update of the EurOtop Neural Network Tool: improved prediction of wave overtopping. In: *Proceedings of the Coastal Engineering Conference*, pp. 1–8.
- Zhang, C., Zheng, J., Wang, Y., Demirbilek, Z., 2011. Modeling wave-current bottom boundary layers beneath shoaling and breaking waves. *Geo-Marine Lett.* 31, 189–201. <https://doi.org/10.1007/s00367-010-0224-9>.
- Zheng, C.W., Li, C.Y., 2015. Variation of the wave energy and significant wave height in the China Sea and adjacent waters. *Renewable Sustainable Energy Rev.* 43, 381–387. <https://doi.org/10.1016/j.rser.2014.11.001>.
- Zou, Q.P., Chen, Y., Cluckie, I., Hewston, R., Pan, S., Peng, Z., Reeve, D., 2013. Ensemble prediction of coastal flood risk arising from overtopping by linking meteorological, ocean, coastal and surf zone models. *Q. J. Roy. Meteorol. Soc.* 139, 298–313. <https://doi.org/10.1002/qj.2078>.

Further reading

- Peng, Z., Zou, Q.P., 2011. Spatial distribution of wave overtopping water behind coastal structures. *Coast. Eng.* 58, 489–498. <https://doi.org/10.1016/j.coastaleng.2011.01.010>.