Universidade do Minho
Escola de Engenharia

Willian Weber de Melo

Operational Forecasting Platforms Based on Artificial Intelligence Methods

July 2024

Operational Forecasting Platforms Based on Artificial Intelligence Methods

Willian Weber de Melo

UMinho | 2024

Fundação para a Ciência e a Tecnologia

MIT Portugal

REPÚBLICA PORTUGUESA

University of Minho
School of Engineering

Willian Weber de Melo

Operational Forecasting Platforms Based on Artificial Intelligence methods.

Doctoral Thesis
Doctoral Program in Civil Engineering

Thesis executed under the supervision of

Professor PhD. José Luís da Silva Pinho
Professor PhD. Isabel Iglesias Fernández

July 2024

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição

CC BY

https://creativecommons.org/licenses/by/4.0/

Acknowledgements

This research was supported by the Doctoral Grant SFRH/BD/151383/2021 financed by the Portuguese Foundation for Science and Technology (FCT) and with funds from the Ministry of Science, Technology, and Higher Education, under the MIT Portugal Program.

Furthermore, this thesis results from the sum of the efforts of many people who contribute to the work. I would like to especially thank:

- Professor José Pinho for supporting the initial idea of this research and for betting that I would successfully learn a completely new topic and achieve the project goals. I am grateful for his guidance during this journey.
- Professor Isabel Iglesias for accurate and constructive criticism and the capacity to always identify the blind spots of the work done. Your guidance and teaching have smoothed my experience during the PhD.
- Laís for always being there for me during all moments along this path. You were the first-row spectator who saw all my rights and wrongs during this time and witnessed my efforts to arrive at this point. I am grateful for sharing my life with you.
- My mother, Tânia, for all love, kindness, and for the role model that you are. You were always present at the most remarkable moments of my life and are the best example of endurance, resilience, and determination that I could ever have. "*Obrigado por tudo".*
- My father, George, for all the support given to help me achieve my goals. Thank you for being my professional example, sharing your experience to help me avoid mistakes, and encouraging me to pursue my dreams.
- My sister, Vanessa, for being the best sister I could ever ask for. Your example encouraged me to study abroad and stay out of my comfort zone. Thank you for all the support you gave me and for always being there for me.
- All my friends and peers for the shared moments during this time. Thank you for your friendship and helping distract me from the obstacles that appeared in my PhD journey.



STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Plataformas de previsão operacionais baseadas em métodos de inteligência artificial.

Resumo

A previsão do impacto das alterações climáticas é um dos maiores desafios científicos do século XXI. O aumento da frequência e intensidade de eventos extremos indica que devemos reforçar e promover medidas de adaptabilidade e resiliência nas nossas sociedades. As situações extremas relacionadas com a água, nomeadamente secas e inundações, são responsáveis por grande parte dos desastres naturais em todo o mundo. Assim, o desenvolvimento de metodologias que antecipem esses cenários permitem gerar conhecimento que ajudará a mitigar os seus efeitos.

Neste contexto, as ferramentas de modelação numérica são amplamente utilizadas e capazes de alcançar resultados confiáveis na previsão de eventos extremos. No entanto, quando são necessários sistemas de previsão operacionais, surgem barreiras que impedem a aplicação de modelos numéricos, nomeadamente capacidade de processamento computacional e armazenamento. Estas limitações são críticas, especialmente quando são necessários modelos de elevadas resoluções espaciais e temporais.

Nesta tese definiram-se e aplicaram-se algoritmos de inteligência artificial para reduzir os custos computacionais de previsões hidro-morfodinâmicas, desenvolvendo duas abordagens de aprendizagem supervisionada: emuladores e modelos de dados. Os emuladores usam resultados de modelos numéricos como dados de entrada e saída, enquanto os modelos de dados processam diretamente dados de campo. O primeiro método foi utilizado para emular a evolução morfológica de um estuário e uma praia, considerando, respetivamente, eventos de longo e curto prazo. O segundo método foi utilizado para forçar a fronteira fluvial de um modelo numérico estuarino, melhorando a sua capacidade de previsão e operacionalização, e também para criar um modelo de dados de evolução de perfis de praia.

Os resultados obtidos demonstram a capacidade e versatilidade dos modelos de inteligência artificial aplicados em diferentes situações com interesse para a implementação de plataformas de previsão. Na presença de séries de dados que alimentaram os modelos com qualidade e resolução adequadas foi possível realizar previsões com erros aceitáveis. Foi possível prever variáveis 1D, como caudais fluviais ou elevações de perfis de praia, e variáveis 2DH, como a erosão e sedimentação em estuários. Os resultados promissores obtidos, permitem antecipar que os modelos de dados poderão ser utilizados em complemento e alternativa aos modelos determinísticos de maior exigência computacional nos sistemas de previsão, podendo ainda contribuir para uma eficaz quantificação das respetivas incertezas.

Palavras chave: *Deep learning*; Hidrodinâmica; *Machine learning*; Morfodinâmica; Sistemas de previsão e alerta.;

Operational Forecasting Platforms Based on Artificial Intelligence methods.

Abstract

Forecasting the impact of climate change is one of the most important scientific challenges of the 21st century. The increase in the frequency and intensity of extreme events indicates that we must enhance and promote adaptability and resilience measures in our societies. Water-related extremes, namely droughts and floods, account for a great part of natural disasters worldwide. Therefore, the development of methodologies that anticipate these scenarios can generate insights that will help mitigate future risks.

In this context, numerical modelling tools are widely used and can achieve reliable results in forecast extreme events. However, when operational forecasting systems are required, barriers arise preventing the application of numerical models, namely computational processing power and storage. These limitations are critical, particularly when high-resolution models and long-term simulations are performed.

This thesis applied artificial intelligence algorithms to reduce the computational costs of hydro-morphodynamic simulations by developing two supervised learning approaches: emulators and data models. Emulators use numerical model results as the input and output data, whilst data models directly process field data to generate an output. The first method was used to emulate the morphological evolution of an estuary and beach, considering long- and short-term events. The second method was used to force the fluvial boundary of an estuarine numerical model, enhancing its forecasting and operational capacities, and to create a data model of beach profiles evolution.

The results demonstrate the capacity and versatility of artificial intelligence models applied to diverse situations, particularly for implementing operational forecasting platforms. When data series with adequate quantity and quality are used to feed these models, acceptable errors can be achieved. It was possible to predict 1D variables, such as fluvial discharge or beach profiles elevations, and 2DH variables, such as erosion and accretion in estuaries. The promising results obtained suggest that data models can be used as a complement and alternative to more computationally demanding deterministic models in forecasting systems, and can also contribute to an effective quantification of their respective uncertainties.

Keywords: Deep learning; Early warning systems; Hydrodynamic; Machine learning; Morphodynamic.

Index

Chapter 1. Introduction 2

1.1. Background 2

1.2. Objectives 4

1.3. Thesis layout 4

1.4. Publications 5

Chapter 2. State-of-the-art 9

2.1. Estuarine and coastal hydrodynamics 9

2.2. Morphodynamics 11

2.3. Numerical modelling 14

2.4. Artificial intelligence methods 15

2.4.1. Tools to implement artificial intelligence models. 22

2.5. Operational forecasting 24

2.6. References 26

Chapter 3. A data model to forecast the morphological evolution of multiple beach profiles. 33

3.1. Introduction 33

3.2. Methodology 35

3.2.1. Study area 37

3.2.2. Beach profile data 38

3.2.3. Beach profile data processing 43

3.2.4. Wave data 44

3.2.5. Random forest algorithm 47

3.2.6. Computation of beach slope and retreat/ advance 49

3.2.7. Storm conditions simulation 50

3.3. Results 50

3.3.1. Training performance 51

3.3.2. Prediction performance 52

3.3.3. Beach slope and retreat comparison 56

3.3.4. Storm simulation scenario 61

3.4. Discussion and conclusions 63

3.4.1. Data models and the importance of open data 63

3.4.2. Coastal processes and data models 64

3.4.3. Methodology limitations 65

3.4.4. Conclusions 66
3.5. References 67
Chapter 4. Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model. 72
Abstract 72
4.1. Introduction 72
4.2. Methods 75
4.2.1. Numerical model setup and data pre-processing 76
4.2.2. Input Variables 77
4.2.3. Convolutional Neural Networks 78
4.2.4. Training and testing 82
4.3. Results and Discussion 82
4.3.1. Numerical model results 82
4.3.2. Input variables correlation 83
4.3.3. Training results 84
4.3.4. Testing results 85
4.3.5. Numerical models versus data models 88
4.3.6. Future discussions 90
4.4. Conclusion 90
4.5. References 91
Chapter 5. Coastal morphodynamic emulator for early warning short-term forecasts. 96
5.1. Introduction 96
5.2. Methodology 98
5.2.1. XBeach model 98
5.2.2. Implementation of the morphodynamic emulator 103
5.3. Results and discussion 107
5.3.1. Simulation set 1 107
5.3.2. Computational performance 107
5.3.3. Influence of the grid resolution 109
5.3.4. Simulation set 2 110
5.3.5. Future of HMD modelling 116
5.4. Conclusions 117
5.5. References 119

Chapter 6. Early warning system for flood in estuarine areas: combining artificial intelligence with numerical models 123
6.1. Introduction 123
6.2. Study area 124
6.3. Methodology 126
6.3.1. Field data collecting and processing 126
6.3.2. Implementation of the data model 129
6.3.3. Implementation of the numerical model 131
6.3.4. FEWS simulation and error metrics 132
6.4. Results 133
6.4.1. Performance of the data models 133
6.4.2. Calibration of the process-based model 136
6.4.3. Performance of the FEWS platform 137
6.5. Discussion 138
6.6. Conclusions 140
6.7. References 140
Chapter 7. Interpreting machine learning-based beach profile evolution models with explainable artificial intelligence tools 144
7.1. Introduction 144
7.2. Materials and Methods 146
7.2.1. Study area 146
7.2.2. Random Forest 147
7.2.3. Data Sources 148
7.2.4. XAI methods 152
7.3. Results 153
7.3.1. Model performance 153
7.3.2. Partial dependence profiles 154
7.3.3. Permutation importance 156
7.3.4. LIME 158
7.4. Discussion and conclusions 161
7.4.1. Methodology limitations 163
7.5. References 163
Chapter 8. Conclusions and final remarks. 167
8.1. Future research 170

Abbreviations and acronyms

AE	Absolute error
AI	Artificial Intelligence
ANN	Artificial Neural Network.
BN	Bayesian Network
BSS	Bed Shear Stress
CES	Cumulative Erosion and Sedimentation
CFL	Courant–Friedrichs–Lewy
CNN	Convolutional Neural Network
COSMO	Programa de Monitorização da Faixa Costeira de Portugal Continental
CTBLC	Current Timestep Bed Level Change
DL	Deep Learning.
DT	Decision Trees
ECMWF	European Centre for Medium-Range Weather Forecast
EPR	Evolutionary Polynomial Regression
EWS	Early Warning System
FEWS	Flood Early Warning System
GPR	Gaussian Process regression
GLMV	Generalized Lagrangian Mean Velocity
GP	Genetic Programming
HMD	Hydro - morphodynamic
JONSWAP	Joint North Sea Wave Project
KDE	Kernel Density Estimation
KGE	Kling-Gupta Efficiency
LSTM	Long-short Term Memory
LW	Low Resolution
MAE	Mean Absolute Error
ME	Mean Error
MLP	Multi-Layer Perceptron.
MORFAC	Morphological Acceleration Factor
ML	Machine Learning
MSE	Mean Squared Error
NN	Neural Network
OR	Original Resolution
PCA	Principal Component Analysis

R^2	Coefficient of determination
RE	Relative Error
RF.	Random Forest
ReLU	Rectified Linear Unit
RMSE	Root Mean Squared Error
WMO	World Meteorological Organization
XAI	Explainable Artificial Intelligence
XBeach	eXtreme Beach behaviour model

Figures list

Figure 2-1. Example of convolutional operation. 18

Figure 3-1. Methodological approach: a) Methodology framework with the steps required to implement a data model . b) Architecture of the model and details on the input data and adopted hyperparameters. 36

Figure 3-2. Study area and location of the beaches where the profiles were measured. White rectangles indicate the number of profiles measured seasonally at each beach, and the rhombuses indicate the coast type of each beach. Coordinates are referred to the ETRS89 global reference system. 38

Figure 3-3. COSMO beach profiles. The yellow line indicates the first profile surveyed, and the orange line is the last profile, which was used to assess the model's performance.. 40

Figure 3-4. Beaches sediment size. The sediment samples were collected at the berm intermediate position (first column), berm edge (second column), and at the beach face (third column). a) D50 results; b) D90 results. 42

Figure 3-5. Occurrence of erosion and accretion rates in northern Portugal beaches along the seasons, computed from the COSMO beach profiles data. 43

Figure 3-6. IBI wave reanalysis data for the study period. a) Significant wave height; b) Wave period; c) Wave direction. 45

Figure 3-7. Comparison between the results predicted by M1 (blue) and M2 (green) models and the observed values (red dots). The historical data used to train the models are represented in yellow (first observation) and grey 53

Figure 3-8. Error distribution of the obtained results. a) M1 global error distribution; b) M2 global error distribution; c) M1 local error distribution; d) M2 local error distribution. 55

Figure 3-9. Beach slope computation results. The red dotted line represents the observed data, whilst the blue and green lines represent the results of M1 and M2, respectively. The slope results are presented in the white boxes at each subplot. 57

Figure 3-10. Profile advance/retreat. a) 1 m height; b) 2 m height; c) 3 m height; d) 4 m height. 59

Figure 3-11. Historical erosion/advance rates at the measured profiles. The capital letters represent the seasons in which the rates were computed, with A referring to autumn, W to winter, and S to summer and spring. 61

Figure 3-12. Storm simulation results for HM0 are equal to 4 m (_4) and to 8 m (_8). 62

Figure 4-1. Initial bathymetry of the estuary and location of selected points for erosion/sedimentation correlation analysis. 78

Figure 4-2. Architecture of the CNN. 81

Figure 4-3. Examples of the D3D results images used to train the CNN. 83

Figure 4-4. Cumulative sedimentation and erosion results of the numerical model (D3D) and the data model (CNN) 86

Figure 4-5. RMSE variation in the testing dataset. 86

Figure 4-6. Differences in the erosion/sedimentation prediction between models at selected points. 87

Figure 5-1. Topo-bathymetry considered for each XBeach domains: a) Linear topo-bathymetry domain; b) Breakwater domain; c) Groin domain. 100

Figure 5-2. Examples of images of the hydrodynamic model results used as input (GLMV and BSS) and of the morphodynamic model results used as output (CTBLC and CES) by the deep learning model. The bound values used to plot the images are in the grayscale colour bar above each image. The DL models used the first two variables as input and at least one of the last two variables as output. . 104

Figure 5-3. DL model architecture. The rectangles represent the resolution of the image in each network layer. The numbers above each layer are the number of filters, and each coloured arrow represents one different operation performed in the network. 105

Figure 5-4. Comparison between emulators using inputs with different grid resolution. 110

Figure 5-5. Emulators RMSE in the test dataset. 110

Figure 5-6. Comparison between the CES simulated by the XBeach model and the emulator in the last simulation time step. a) Breakwater domain; b) Groin domain. 112

Figure 5-7. Models mean RMSE at each validating time step. a) Error (m) of the models trained in CES data; b) Errors (mm) of the models trained in CTBLC data. 114

Figure 5-8. Mean error of the DL models in the domains. Red colours represent emulator underestimation and blue colours overestimation. White colours indicate a perfect agreement between the numerical model and the emulator. a) Error of the models (m) in forecasting the CES; b) Error of the models (mm) in forecasting CTBLC. 115

Figure 5-9. DL model error distribution. a) Distribution of the CEM errors; b) Distribution of the CTBLC errors. 116

Figure 6-1. a) Douro River localisation in the Iberian Peninsula; b) Douro River sub-basins and estuarine area; c) Numerical model domain and fluvial boundary in the Crestuma-Lever dam. 125
Figure 6-2. FEWS framework. Main steps to forecast water levels. 126
Figure 6-3. Monthly cumulated precipitation in the Douro River sub-basins. 127
Figure 6-4. Historical data of the Douro River discharge at the Crestuma-Lever dam. 128
Figure 6-5. Architecture of the implemented data models. 130
Figure 6-6. Error distribution of data models. 134
Figure 6-7. Flow forecast at the Crestuma-Lever dam during the January 2016 flood. 135
Figure 6-8. Calibration results of the hydrodynamic numerical model. 136
Figure 6-9. Water level simulation results at Cais dos Banhos using the LSTM and ATT models to generate discharge values at the fluvial boundary. 137
Figure 7-1. Study area, location of the selected beach profiles, wave data extraction poitnas and beach ty-pology. The white rectangles identify the number of profiles at each location in the map..... 146
Figure 7-2. Example of a decision tree structure. 148
Figure 7-3. COSMO program beach profile data used to train the model. 149
Figure 7-4. IBI wave reanalysis variables from Copernicus Marine Service. 150
Figure 7-5. Local model performance for the test dataset. 154
Figure 7-6. The PD profile results for M1. a) Effect of profile variables (coordinates, time, and length). b) Influence of wave action variables (HM0, wave period, and wave direction). c) Influence of sediment size variables at berm edge, intermediate position, and beach face. 155
Figure 7-7. The partial dependence profile results for M2. a) Effect of profile variables (coordinates, time, and length). b) Influence of wave action variables (HM0, wave period, and wave direction). 156
Figure 7-8. Permutation importance of variable groups. 157
Figure 7-9. Permutation importance of input variables. 158
Figure 7-10. LIME method results for shoreline points. 159
Figure 7-11. LIME method results for berm edge points. 160

Tables list

Table 2-1: Applications of emulation techniques..... 21

Table 2-2. Emulation methods applied to morphodynamics..... 22

Table 3-1. Beach sediments characteristic diameters (D50 and D90). The number following the diameter indicates where the sample was collected (1 – berm intermediate position; 2 – berm edge; 3 – beach face). 41

Table 3-2. Mean and std of beaches elevation and profile length. 44

Table 3-3. Mean value and std of the wave variables provided by the IBI reanalysis product at points P1 and P2..... 46

Table 3-4. Training results of the M1 and M2 models..... 51

Table 3-5. MAE and RMSE in the test dataset. 54

Table 3-6. Beach slope absolute error results. 58

Table 4-1. Hydrodynamic variables' correlation with cumulative erosion/sedimentation. 83

Table 4-2. Hyperparameters of the convolutional layers. 84

Table 5-1. Domain characteristics and boundary conditions of simulation sets. 101

Table 5-2. Testing conditions of the DL models. In simulation set 1, the up-sampling layer was evaluated whilst simulation set 2 assessed different combinations of inputs and outputs..... 102

Table 6-1. Performance metrics of the data models. 133

Table 6-2. Numerical model calibrated parameters. 137

Table 6-3. Performance comparison between LSTM and ATT models for forecasting water levels in Cais dos Banhos..... 138

Table 7-1. Statistical description of wave action variables. Numbers 1 and 2 refer to the wave data points from which the data were extracted..... 151

Table 7-2. Characteristic sediment sizes (D50 and D90) at the studied beaches..... 151

Table 7-3. LIME method mean, maximum and minimum weights for each input variable..... 161

CHAPTER 1.

Introduction

"Be the change that you wish to see in the world."

Mahatma Gandhi (1869 – 1948) – Indian national leader and 5 times nominated for the Nobel Peace Prize.

Chapter 1. Introduction

This chapter presents the motivation and context for this thesis. The importance of the thesis topic is given, followed by the research objectives and the structure of this thesis.

1.1. Background

Climate change and its impacts are among the most important scientific challenges of the 21st century. Stemming from the Second Industrial Revolution, the escalating concentration of greenhouse gases in the Earth's atmosphere has increased the global mean temperature, breaking records year after year. According to the World Meteorological Organization (WMO, 2024), Earth's temperature is venturing into uncharted territory.

The first registers of global warming trends were published in the first half of the 20th century, although they were considered part of the natural temperature cycles of the planet. This phenomenon was oversimplified as “the winters had gotten milder” (Spencer, 2008). Furthermore, an optimistic outlook emerged, suggesting the potential cultivation of wheat and codfish creation in northern areas, where these species have never been observed before (Spencer, 2008). However, by the end of that century, scientists had begun to publish their concerns about the consequences of greenhouse gas accumulation in the atmosphere, emphasising its potential to imperil diverse ecosystems and species.

Climate change is causing more frequent and intense extreme events, such as floods, droughts, hurricanes, and storms (Pörtner et al., 2022). These events can devastate people and infrastructure, resulting in overwhelming losses, particularly in vulnerable communities. Another effect of climate change is the rise of the mean sea level due to ice melting on the top of the mountains, glaciers, and global poles, combined with the thermal expansion of large bodies of water due to oceanic heating. This increases the severity of coastal floods and leads to food insecurity and salinity intrusion in freshwater sources.

For example, during the first quarter of 2023, a record-breaking intense rain profoundly affected the north-eastern coastal region of São Paulo, Brazil (Azevedo-Santos et al., 2023). This event, featuring the highest recorded precipitation volume in the area, led to tragic consequences, including over 40 fatalities and the displacement of more than 500 families. Concurrently, the United States confronted a dichotomy of climatic extremes: a blistering heat wave swept across the southern regions, whereas a severe winter storm affected the northern areas (NOAA, 2024). These meteorological phenomena triggered a series of disruptions, including power outages, wildfires, and road closures. Furthermore, Lisbon experienced two successive extreme floods within one week, culminating in financial losses exceeding 290 million euros.

In this context, the development of tools that can anticipate the effects of future climate scenarios and extreme events is urgently needed to mitigate their impacts and reduce losses. One of the most commonly selected approaches to understanding future scenarios on beaches and estuaries is through simulations based on deterministic numerical models. These models are developed by using mathematical descriptions of the main physical processes responsible for the dynamics of these areas. Implementing a numerical model involves the creation of a computational discretisation mirroring the study area, followed by calibration against *in-situ* data. Subsequently, the numerical model is forced with different boundary conditions to represent future scenarios. This process makes it possible to anticipate and quantify the damage resulting from distinct scenarios, enabling the strategic formulation of preventive and mitigatory measures.

The accuracy of numerical modelling methodologies depends on the grid characteristics in which the software computes the hydro-morphodynamic (HMD) variables. High-resolution models yield more reliable outcomes as the finer grid effectively captures the intricate features of the study area. However, this advantage is counterbalanced by higher computational costs, thereby constraining the evaluation of complex scenarios owing to computational resources and simulation time limitations. Conversely, low-resolution models are faster to process, although a reduction in resolution amplifies the resulting uncertainties. These problems primarily affect long-term simulations, which are essential for studying the impacts of extreme events and climate change on HMD patterns in estuarine and coastal regions.

Artificial intelligence (AI) methods can be compelling solutions to this problem. Instead of solving a set of differential equations describing the phenomena under scrutiny, AI models relate input and output data through simpler mathematical operations that depend on the employed algorithm. Although these models are limited by the data used to train them, they can be an interesting and feasible solution for reducing the computational costs of numerical model simulations.

The application of AI algorithms is rapidly increasing owing to recent advances in computer science. These algorithms have proven to be remarkably versatile and powerful tools that can be adapted to various datasets and diverse objectives. For instance, large language models, such as ChatGPT and Google Gemini, are becoming extremely popular and are startling their users because of their capacity to accurately answer specific tasks and the velocity at which the answer is given.

Furthermore, AI models are already being used in climate sciences to increase the resolution of numerical model results, reduce the computational costs of simulations, and completely replace numerical models with data models. These applications show the potential harboured by AI methods,

where the development and implementation of AI models for a crescent range of tasks becomes feasible and promising, given that the computational resources are adequate.

1.2. Objectives

This thesis aims to explore the applications of AI algorithms to HMD processes to generate insights into their implementation, limitations, and accuracy. Although AI methods have already demonstrated the capacity to be widely applied to the most diverse problems and achieve excellent results, there are still gaps in this type of methodology that need to be filled, for example:

- Can AI models achieve the same accuracy as high-resolution HMD numerical models?
- How can AI be applied to reduce the computational cost of high-resolution HMD models?
- Can AI models be applied in domains with distinct characteristics subject to various physical processes?
- How can AI be applied to foster the implementation of early warning systems?

This thesis proposes two approaches to answering these questions. The first is to implement a full data model, which requires extensive field observations to be configured. The second is to implement an emulator, which partially or entirely surrogates another model type, such as a numerical one. I. e., results generated by the numerical model are used to train an AI model. The following specific objectives ensure the pursuit of the main objective:

- Combining field data from multiple monitoring programs on the northern Portuguese coast.
- Generation of extensive HMD results from Delft3D (D3D) and eXtreme Beach behaviour model (XBeach) numerical models. The D3D model simulated 200 years of morphological evolution of an estuary under average conditions, whereas the XBeach model simulated the effects of an extreme event on a sandy beach.
- Development, adaptation, and implementation of emulation techniques to reduce the computational cost of numerical modelling.
- Demonstrate the efficacy of the methodology and generate a technological solution to build an operational forecasting and warning system for the case study region.
- Analysis of the platform performance, comparing its predictions for different time horizons with the observed trends.

1.3. Thesis layout

This thesis is structured based on scientific articles published in international peer-reviewed journals, aligning them with the objectives of this research. The introductory chapter emphasises the

significance of HMD modelling and its synergy with AI. This section also outlines the thesis objectives and highlights its contributions to expanding the knowledge boundaries of HMD modelling.

The second chapter presents a state-of-the-art about AI methods and HMD, intending to help to understand better some key concepts developed in the following chapters.

The third chapter introduces the first published peer-reviewed article that showcases the application of an AI model using data from diverse monitoring programs to forecast the profile evolution of beaches.

The fourth chapter addresses the implementation of a morphodynamic estuarine numerical model emulator using 200 years of a HMD simulation to generate artificial data and a deep learning (DL) architecture to simulate 2DH morphodynamic processes.

The fifth chapter demonstrates that a similar AI architecture to the one used in chapter four to emulate 200 years of HMD estuarine evolution can be applied to emulate storm effects in coastal areas, indicating that the methodology is suitable for emulating long- and short-term events.

The sixth chapter implemented two neural networks to forecast the fluvial discharge into an open boundary of an estuarine hydrodynamic deterministic model. In addition, the forecasting capacity of the neural networks considering different time horizons was evaluated, and the fluvial discharge predictions were used in a process-based numerical model to simulate a flood event. This chapter synthetizes the use of both models and proposes a framework for the modelling component of an operational forecasting platform.

The seventh chapter delved into the concept of explainable AI (XAI) and highlighted techniques that aim to improve the transparency and comprehensibility of AI models. The interpretability barrier and difficulties in understanding how AI models relate inputs to outputs reduce the wider adoption of such methods. Therefore, this chapter demonstrates how appropriate methods can enhance our understanding of an AI model and increase its reliability.

The eigthth chapter summarises the main challenges regarding the development of this thesis and discusses the similarities and differences of the works presented in chapters 3 to 7. It also provides concluding remarks, summarising the main findings and underlining the potential future developments for AI within the context of HMD modelling.

1.4. Publications

This thesis resulted in several publications in indexed peer-reviewed journals and national and international conference proceedings. Only publications in which the author was the first author were used as chapters of this thesis. The list of all publications is available below.

i. Peer-Reviewed Journals:

- Weber de Melo, W., Pinho, J., & Iglesias, I. (2024). A data model to forecast the morphological evolution of multiple beach profiles. Coastal Engineering, 104574. https://doi.org/10.1016/j.coastaleng.2024.104574. (Chapter 3).
- Weber de Melo, W., Pinho, J., & Iglesias, I. (2023). Coastal morphodynamic emulator for early warning short-term forecasts. Environmental Modelling & Software, 165, 105729. https://doi.org/10.1016/j.envsoft.2023.105729. (Chapter 5).
- Weber de Melo, W., Pinho, J., & Iglesias, I. (2022). Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model. Journal of Hydroinformatics. https://doi.org/10.2166/hydro.2022.068. (Chapter 4).
- Menten, G., Melo, W., Pinho, J., Iglesias, I., & Antunes do Carmo, J. (2023). Simulation of Saltwater Intrusion in the Minho River Estuary under Sea Level Rise Scenarios. Water, 15(13), 2313. https://doi.org/10.3390/w15132313.
- Iglesias, I., Bio, A., Melo, W., Avilez-Valente, P., Pinho, J., Cruz, M., Gomes, A., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2022). Hydrodynamic Model Ensembles for Climate Change Projections in Estuarine Regions. Water, 14(12), 1966. https://doi.org/10.3390/w14121966
- Iglesias, I., Pinho, J. L., Avilez-Valente, P., Melo, W., Bio, A., Gomes, A., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2022). Improving Estuarine Hydrodynamic Forecasts Through Numerical Model Ensembles. Frontiers in Marine Science, 9. https://doi.org/10.3389/fmars.2022.812255.

ii. Book chapters

- Weber de Melo, W., Pinho, J., Iglesias, I., Bio, A., Avilez-Valente, P., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2022). Flood Risk Assessment at the Douro River Estuary (pp. 37–49). https://doi.org/10.1007/978-981-16-5501-2_4.
- Iglesias, I., Pinho, J. L., Bio, A., Avilez-Valente, P., Melo, W., Vieira, J. L., Bastos, L., & Veloso-Gomes, F. (2023). A importância dos modelos numéricos na previsão dos efeitos das alterações climáticas e dos eventos extremos em áreas estuarinas (pp. 13–43). https://doi.org/10.34037/978-989-9053-18-2_11_1.

iii. Conference papers and communications

- Weber de Melo, W., Pinho, J. L., Iglesias, I., Gomes, A., Vieira, J., Bio, A., Bastos, L., Veloso-Gomes, F., and Avilez-Valente, P.: Estuarine morphodynamics forecast using a numerical model emulator based on deep learning methods. A first approach., EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-6277, https://doi.org/10.5194/egusphere-egu22-6277, 2022.
- Melo, W., Pinho, J., Iglesias, I., Bio, A., Avilez-Valente, P., Gomes, A., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2021). Numerical Models' Application for Morphodynamics Assessment of Climate Change Impacts in the Minho River Estuary. Environmental Smoke, (Special Issue), 1–6. https://doi.org/10.32435/envsmoke/xibesymp.1

CHAPTER 2.

State-of-the-art.

"Nobody ever figures out what life is all about, and it does not matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough."

Richard Feynman (1918 – 1988) – North American theoretical physicist and Nobel Prize laureate in 1961.

Chapter 2. State-of-the-art

This chapter summarises the knowledge that allows for a holistic view of the two main areas addressed in this thesis: HMD modelling and AI. First, estuarine and coastal HMD were explained, focusing on the main equations that drive these processes. Subsequently, the main aspects of numerical modelling methodologies are discussed, followed by an introduction to AI-based methodologies. The differences between machine learning (ML) and deep learning (DL) are discussed, presenting some of the most prominent algorithms from each category, with a special emphasis on those employed in this research. The similarities and differences between AI modelling and HMD numerical modelling are also addressed. Furthermore, operational forecasting systems are explained, focusing mainly on flood early warning systems (FEWS), which combine real-time collected data with HMD modelling methodologies.

2.1. Estuarine and coastal hydrodynamics

Coastal and estuarine hydrodynamics present complex phenomena for which knowledge is essential to protect many human settlements in these zones. Understanding and predicting hydrodynamic variables, such as water levels, extreme storm surges, current velocities, and wave energy, are necessary because of the sea level rise (SLR) and extreme events that may jeopardise these areas. Possible SLR impacts include increased flood risk, freshwater salinisation, and beach erosion (Nicholls, 1995). In addition, understanding hydrodynamic patterns is necessary to anticipate future morphodynamic changes, determine erosion-prone areas, and design defence structures (Stansby, 2013).

The hydrodynamics in these environments follow the physical laws of open-channel flow. Open channels are natural or artificial conveyance structures with a free surface (Akan, 2006a). The flow varies according to the channel geometry and characteristics, expressed by its slope, shape, constituent material, and energy difference between the upstream and downstream. Ocean circulation processes also influence estuaries and coastal areas. Oceanic currents, whose generation mechanisms are the same as those generated in estuarine regions, can be grouped according to their energy sources: wind-driven, thermohaline, and tidal currents (Pinho, 2000).

Wind-driven currents occur owing to atmospheric pressure gradients around the world. The winds blow from the high- to low-pressure regions, producing a motion effect in the surface waters. Thermohaline currents occur when a density gradient is associated with the temperature and salinity properties of water masses. The warmer and low-salinity waters are lighter than the cold and high-salinity waters, which creates convection movement. Despite the differences that seem to arise between these two types of currents, the sun is the main energy source in both processes. Finally, the tidal variation effect results from the gravitational forces exerted by the Moon and the Sun on Earth.

Tides are described as long-period waves, mainly resulting from the imbalance between the gravitational forces exerted by the Sun and the Moon with the Coriolis effect (Saes et al., 2018), and are the major source of dissipation in the deep ocean (Talley et al., 2011). Their effects can be Fourier decomposed and described according to basic tidal frequencies or tidal components, such as long-period, diurnal, and semi-diurnal (Balmforth et al., 2005). This results in tidal harmonics, which define the tidal variation around Earth's surface.

The gravitational force is an attraction force resulting from the mass of an object or body, which bends space and attracts other objects towards it. The magnitude of the attraction force is directly proportional to the mass of the body. The Coriolis effect results from Earth's rotation, which generates a centrifugal force. The action of these forces in the water masses of the planet and their variations according to the lunar and solar cycles define the tidal changes. For instance, the resulting gravitational force is stronger when the moon is closest to Earth, which characterizes a spring tide period. When it withdraws, the attraction weakens, defining the neap tide period.

Waves also play a significant role in coastal hydrodynamics. They transfer considerable energy towards the coasts, driving morphological changes and erosion (Vousdoukas et al., 2017). They can be classified as surface gravity waves, in which the restoring force is the gravity, and internal gravity waves, in which the restoring force is the buoyancy force, which results from the gravity and the difference in density between the displaced water parcel and its neighbors at the same pressure (Talley et al., 2011). Saes et al. (2018) classify waves as deep waves, which fully develop their orbital trajectory until dissipating their energy in the inferior layers, and shallow waves, which are wind driven.

Another relevant wave type affecting ocean water motion is infragravity waves, with frequencies lower than wind-generated waves. This type of wave is indirectly formed by wind, dominates the motion near the coasts, particularly during storms, and contributes considerably to sediment transport, dunes, and barrier breaching (Bertin et al., 2018). Hence, water-level variations in coastal areas result from the effects of multiple waves generated by diverse sources. In estuaries, the water level is defined by the interaction of river discharge, tidal variation, and bathymetry. The predominant behaviour will depend on the characterisation of these three factors (river flow, tidal amplitude, and bathymetry). However, if the estuary is large enough, the wind can also affect the hydrodynamic patterns and the water level.

To describe the movement of fluids, the Newton's theory was applied to the mechanics of fluids, which result in the Navier–Stokes equations (Eq. (**2-1**, (**2-2** and (**2-3**), where p is the pressure (Pa), u is the velocity component on the x direction (m/s), v is the velocity component on the y direction (m/s), w

is the velocity component on the z direction (m/s), μ is the dynamic viscosity coefficient (kg/m · s) and Fx, Fy and Fz are the volume force components by mass unit (N/kg).

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}\right)=\rho*F_x-\frac{\partial p}{\partial x}+u\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\right) \quad (2\text{-}1)$$

$$\rho\left(\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}\right)=\rho*F_y-\frac{\partial p}{\partial y}+u\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2}\right) \quad (2\text{-}2)$$

$$\rho\left(\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}\right)=\rho*F_z-\frac{\partial p}{\partial z}+u\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}+\frac{\partial^2 w}{\partial z^2}\right) \quad (2\text{-}3)$$

Substituting the local velocity components by an average value added with a fluctuation on those equations results in the Reynolds equations (Eq. (**2-4**, (**2-5** and (**2-6**), which describe the average movement of a fluid particle and where $\hat{u}$, $\hat{v}$ and $\hat{w}$ are the temporal average of the velocity components on each direction (m/s), $\hat{p}$ is the pressure temporal average (Pa) and u', $v'$and w' are the fluctuations of the velocity components along each direction (m/s). These equations form the basis of hydrodynamic formulations, which are discussed in the following sections. In addition, considering that the horizontal length scale is much greater than the vertical scale, it can be assumed that the pressure gradients are hydrostatic and that the vertical velocity component is negligible, resulting in the Saint-Venant equations after a vertical integration along the vertical direction, also known as shallow water equations. These are a special case of Reynolds equations, that consider hydrostatic pressure distribution, uniformly distributed velocity over a channel section, low average bed slope, and homogeneous and incompressible flow (Akan, 2006b).

$$\rho\left(\frac{\partial \hat{u}}{\partial t}+\frac{\partial \hat{u}\hat{u}}{\partial x}+\frac{\partial \hat{u}\hat{v}}{\partial y}+\frac{\partial \hat{u}\hat{w}}{\partial z}\right)=\rho F_x-\frac{\partial \hat{p}}{\partial x}+\mu\left(\frac{\partial^2 \hat{u}}{\partial x^2}+\frac{\partial^2 \hat{u}}{\partial y^2}+\frac{\partial^2 \hat{u}}{\partial z^2}\right)-\frac{1}{p}\left(\frac{\partial \overline{u'u'}}{\partial x}+\frac{\partial \overline{u'v'}}{\partial y}+\frac{\partial \overline{u'w'}}{\partial z}\right) \quad (2\text{-}4)$$

$$\rho\left(\frac{\partial \hat{v}}{\partial t}+\frac{\partial \hat{v}\hat{u}}{\partial x}+\frac{\partial \hat{v}\hat{v}}{\partial y}+\frac{\partial \hat{v}\hat{w}}{\partial z}\right)=\rho F_y-\frac{\partial \hat{p}}{\partial y}+\mu\left(\frac{\partial^2 \hat{v}}{\partial x^2}+\frac{\partial^2 \hat{v}}{\partial y^2}+\frac{\partial^2 \hat{v}}{\partial z^2}\right)-\frac{1}{p}\left(\frac{\partial \overline{v'u'}}{\partial x}+\frac{\partial \overline{v'v'}}{\partial y}+\frac{\partial \overline{v'w'}}{\partial z}\right) \quad (2\text{-}5)$$

$$\rho\left(\frac{\partial \hat{w}}{\partial t}+\frac{\partial \hat{w}\hat{u}}{\partial x}+\frac{\partial \hat{w}\hat{v}}{\partial y}+\frac{\partial \hat{w}\hat{w}}{\partial z}\right)=\rho F_z-\frac{\partial \hat{p}}{\partial z}+\mu\left(\frac{\partial^2 \hat{w}}{\partial x^2}+\frac{\partial^2 \hat{w}}{\partial y^2}+\frac{\partial^2 \hat{w}}{\partial z^2}\right)-\frac{1}{p}\left(\frac{\partial \overline{w'u'}}{\partial x}+\frac{\partial \overline{w'v'}}{\partial y}+\frac{\partial \overline{w'w'}}{\partial z}\right) \quad (2\text{-}6)$$

2.2. Morphodynamics

Morphodynamics investigates the morphologies and transport observed in sedimentary environments (Seminara, 1998). Morphology affects hydrodynamics in a way that influences the further evolution of the morphology (deSwart & Zimmerman, 2009; Friedrichs, 2011). Understanding how sediments are transported is essential to promote the sustainable use and management of natural coastal

resources because they are considered the most immediate contributors to the resilience of coastal stretches prone to flood risk (Karathanasi & Belibassakis, 2019; Sanchez-arcilla et al., 1998).

One of the most challenging problems encountered in open-channel hydraulics is determining the rate of movement of the bed material by a stream (Einstein, 1950). The physical explanation of sediment transport requires the determination of the unsteady motion of many solid particles with irregular shapes, driven by hydrodynamic interaction in turbulent flow and their mutual collisions (Seminara, 1998). In addition, the solution of the Navier-Stokes equations at large Reynolds numbers in a domain continuously changing in time, coupled with the solution of the equations of sediment motion, requires substantial computational resources (Seminara, 1998).

Sediments in rivers and coasts originate from interactions between the water and rocks. According to the rock type, sediments with varied sizes and characteristics are transported along the river and deposited or eroded according to the current velocity. This process can take centuries to stabilise under average conditions. Concerning coastlines, 31% of the ice-free world shoreline is sand (Luijendijk et al., 2018) from igneous rock weathering (Jackson & Short, 2020). Rivers are the main coastal sediment source, although fluvial transport is being reduced owing to dam construction, increasing the erosion risk of coastal and estuarine areas due to the less amount of sediments that reach these areas (Jackson & Short, 2020).

Sediment dynamics can be divided into four stages according to the bed shear. The first stage occurs when the shear stress cannot move the soil particles. The second stage, known as bedload transport, occurs when the shear stress is sufficient to move the particles at the bottom. In this phase, the particles move by sliding, rolling, and jumping. In the third stage, the shear stress continues increasing, and the sediments start to be suspended and transported by the flow. Finally, in the fourth stage, the shear stress is sufficiently strong to move a layer of grains of increasing thickness (Seminara, 1998).

Although suspended sediment transport is a complex phenomenon, it can be numerically computed using the 3D advection-diffusion equation (2-7), where $c^{(l)}$ is the mass concentration of the *l* sediment fraction (kg/m³), *u, v,* and *w* are the velocity components of the flow (m/s), $\varepsilon_{s,x}^{(l)}$, $\varepsilon_{s,y}^{(l)}$, and $\varepsilon_{s,z}^{(l)}$ are the eddy diffusivity coefficients of the sediment fraction for each direction (m²/s), and ${w_s}^{(l)}$is the sediment settling velocity of the sediment fraction (m/s) (Deltares, 2018; Roelvink & Reniers, 2011).

$$\frac{\partial c^{(l)}}{\partial t} + \frac{\partial u \cdot c^{(l)}}{\partial x} + \frac{\partial v \cdot c^{(l)}}{\partial y} + \frac{\partial (w - {w_s}^{(l)}) \cdot c^{(l)}}{\partial z} - \frac{\partial}{\partial x}\left(\varepsilon_{s,x}^{(l)} \cdot \frac{\partial c^{(l)}}{\partial x}\right) - \frac{\partial}{\partial y}\left(\varepsilon_{s,y}^{(l)} \cdot \frac{\partial c^{(l)}}{\partial y}\right) - \frac{\partial}{\partial z}\left(\varepsilon_{s,z}^{(l)} \frac{\partial c^{(l)}}{\partial z}\right) = 0 \qquad (2\text{-}7)$$

Sediment settling velocity depends on the soil type. Equation (2-8) determines the settling velocity in a fluid-sediment mixture as a function of the sediment concentration and non-hindered settling velocity (Deltares, 2018; Richardson & Zaki, 1997), where C_{soil} is the reference density (kg/m³), $\omega_{s,0}^{(l)}$ is the basic sediment fraction specific settling velocity (m/s), and c_s^{tot} is the mass concentration of the sediment fraction (kg/m³).

$$\omega_s^{(l)} = \left(1 - \frac{c_s^{tot}}{C_{soil}}\right)^5 \cdot \omega_{s,0}^{(l)} \tag{2-8}$$

The parameter $\omega_{s,0}^{(l)}$ depends on the soil characteristics. For cohesive sediments, such as clay, it follows equation (2-9), and for non-cohesive sediments, such as sand and gravel, it follows equation (2-10) (Deltares, 2018), where $\omega_{s,0}^{(l)}$ is the (non-hindered) settling velocity of the sediment fraction (m/s), $\omega_{s,max}^{(l)}$ is the settling velocity of the sediment fraction at salinity concentration $\mathrm{S_{max}}$, $\omega_{s,f}^{(l)}$ is the freshwater settling velocity of the sediment fraction, S is the salinity, $\mathrm{S_{max}}$ is the maximum salinity value, $s^{(l)}$ is the relative density of the sediment fraction, $D_s^{(l)}$ is the representative diameter of the sediment fraction (μm), and v is the kinematic viscosity coefficient of water (m²/s).

$$\cdot\, \omega_{s,0}^{(l)} = \begin{cases} \omega_{s,max}^{(l)} \cdot \left(1 - \cos\left(\frac{\pi \cdot S}{S_{max}}\right)\right) + \frac{\omega_{s,f}^{(l)}}{2} \cdot \left(1 + \cos\left(\frac{\pi \cdot S}{S_{max}}\right)\right), \text{when } \mathrm{S} \leq \mathrm{S_{max}} \\ \omega_{s,max}^{(l)}, \text{when } \mathrm{S} > \mathrm{S_{max}} \end{cases} \tag{2-9}$$

$$\cdot\, \omega_{s,0}^{(l)} = \begin{cases} \left(s^{(l)} - 1\right) \cdot g \cdot {D_s^{(l)}}^2 & , \boldsymbol{65\,\mu m < D_s \leq 100\,\mu m} \\ \frac{10 \cdot v}{D_s} \cdot \left(\sqrt{1 + \frac{0.01 \cdot \left(s^{(l)} - 1\right) \cdot g \cdot {D_s^{(l)}}^3}{v^2}} - 1\right) & , 100\mu m < D_s \leq 1000\mu m \\ 1.1 \cdot \sqrt{\left(s^{(l)} - 1\right) \cdot g \cdot D_s^{(l)}} & , 1000\,\mu m < D_s \end{cases} \tag{2-10}$$

Bedload transport can also be determined through analytical equations that relate the effects of macro-roughness to the availability of mobile sediments. However, the outcome of applying analytical solutions is usually overestimated (Bartels et al., 2021). Accurately simulating morphodynamic variables helps to manage coastal and estuarine natural resources better, making it possible to plan human interventions, such as dredging operations or designing defence structures. However, the uncertainties of HMD models and the computational resources required for simulations still hinder the wider adoption of this type of solution in coastal management.

The definition and modelling of sediment transport in estuaries and coastal areas are extremely complex because of the combination of currents and wave effects. Waves interact with currents, modifying

bed shear stress, bed ripples, sediment mobility, and near-bed current transporting sediments (Roelvink & Reniers, 2011), requiring simplification of these processes in modelling attempts. In coastal areas, waves dominate sediment transport both onshore and offshore. For example, wave action is the main process of restoring the beach after erosion induced by storms, bringing sand back onshore (Roelvink & Reniers, 2011). For offshore transport, the return flow generated by waves compensates for the onshore-directed Stokes drift and phase coupling between wave groups and long waves (Roelvink & Reniers, 2011).

The use of numerical models can present reliable results in simulating HMD, but long-term simulations, which are necessary to understand morphodynamic processes, require huge computational resources, which limits the application of this type of methodology. At each simulation time step, the numerical model computes the hydrodynamic variables resulting from the boundary conditions for each considered physical process (waves, winds, and currents) and the resulting sediment transport of each component. Subsequently, the bathymetry of the domain is updated, which changes the hydrodynamic pattern of the next simulation time step, resulting in a feedback loop. Hence, the computational costs are proportional to the resolution of the model, the number of time steps, and the number of considered physical processes.

2.3. Numerical modelling

Numerical modelling is one of the methodologies that are currently the most selected for simulating HMD processes. They have been used to predict future water levels for risk assessment (Iglesias et al., 2019; Li et al., 2016), to forecast plume dynamics under different conditions (Des et al., 2019; Gong et al., 2020; Sousa et al., 2014), and to simulate sediment transport (Portela, 2008), among other topics, achieving significant maturity, but still requiring substantial computational resources (Elko et al., 2014).

Numerical modelling-based methods consist of creating a numerical representation of the study area, which is commonly defined by a computation grid and bathymetry of the region. The grid is a set of points referenced to a coordinate system where the software computes the HMD variables. The higher the number of points in the grid, the higher the resolution of the model and the details of the results. However, more processing power is required for the simulation. Bathymetry is the depth of each grid point, usually based on observed data and interpolated to the grid points. The numerical methods applications allow for solving Reynolds equations or shallow water equations, previously presented in this chapter, by applying the finite difference equations using Taylor series approximations to the equation's partial terms (Akan, 2006b).

To achieve reliable results, a numerical model should be consistent, convergent, and stable (Akan, 2006b). A model is consistent if the governing equations are obtained from the finite difference equations as the space and time increments approach zero (Franz & Melching, 1997). It will be convergent if, as the time and space increments are reduced, the results approach the limiting values that match the true solution (Akan, 2006b), and the method will be stable if round-off and truncation errors do not accumulate, causing the solution to diverge.

After implementing the model, i.e. discretising the study area and defining the bathymetry and grid, the parameters of the model must be calibrated to ensure that the model's results are consistent. The calibration step involves adjusting the parameter values, such as the friction coefficients or tidal harmonic constituents, to reduce the error between the observed data and simulated results. In this procedure, automatic tools such as OpenDA software can be used to determine the optimal combination of parameter values. OpenDA has already been demonstrated to be a valuable tool for speeding up the calibration procedure while allowing reliable results to be achieved (Garcia et al., 2015; Melo et al., 2020).

Despite the capacity of numerical modelling systems to simulate complex patterns, they are limited by available computational resources. Long-term simulations, which are essential for understanding the effects of climate change or coastal and fluvial morphology evolution, are limited by the available computational resources necessary to store their results and by the computation time, considering that no numerical instability occurs during the simulations. These limitations constrain the application of numerical solutions on a larger scale, particularly in areas that do not have adequate computational resources for implementation.

There are several options of software capable of modelling hydrodynamic patterns, such as Delft3D, openTelemac-Mascaret, and Mike21 (Capellán-Pérez et al., 2016; Des et al., 2019; Iglesias et al., 2019; Shen et al., 2018; Xiang et al., 2019), which are appropriate for modelling estuarine HMD. The XBeach software is among the most popular options for modeling coastal areas (Roelvink et al., 2009). These solutions can achieve satisfactory results with acceptable computational costs, depending on the simulation length and processes considered.

2.4. Artificial intelligence methods

The term AI can be understood as a computational approach inspired by human intelligence capable of adapting according to circumstances (Ertel, 2017a), or the theory and development of computer systems that can perform tasks that require human intelligence (Kapelan et al., 2020). The use of AI has been increasing in recent decades, and it is currently receiving investments and incentives with the recent deployment of large language models such as ChatGPT, Google Gemini, and Claude.

What was once considered a technology of a far future is currently reachable due to advances in computational processing capacity and high-level programming languages, allowing us to implement many kinds of AI algorithms with a few lines of code without needing advanced programming skills.

AI can be subdivided into two categories: ML and DL. ML refers to coding programs that automatically adjust their performance according to the given data (Irizarry, 2019), i. e., they learn according to their experience and the provided data. ML algorithms, in turn, can be subdivided into three groups: supervised, unsupervised, and reinforcement learning.

Regarding supervised learning algorithms, the user defines the input and output data, and the algorithm adjusts its parameters according to the relationship between these two datasets. In unsupervised learning methods, the data are given to the algorithm, which discovers patterns in the dataset. Reinforcement learning consists of giving rewards to the model if it performs a task correctly, otherwise, it is penalised. The method selection hinges on the goals of the model and available data. In this thesis, only supervised approaches were applied.

The ML algorithm used in this thesis was the Random Forest, which consists of an ensemble method using multiple decision trees (DT). DT is a popular ML algorithm because of its simplicity, efficiency, and interpretability (Ertel, 2017b; Kubat, 2017). This algorithm uses a series of decision rules that result from input and output data (Pedregosa et al., 2011). However, DTs are easily overfitted, which means that the accuracy of the model significantly decreases when the input data are outside the training data range. To overcome this problem, the Random Forest algorithm combines the results of multiple DTs. This approach does not guarantee that the ensemble result will be better than the best DT but ensures that the result will have a lower bias.

DL algorithms are a particular case in ML that covers and extends the capacity of ML algorithms by combining multiple layers and making the algorithm capable of solving more complex tasks at the cost of increasing the complexity of the models and the computational resources necessary to execute them (Wani et al., 2020). This advancement from ML to DL was possible because of the backpropagation algorithm, which optimised the training of AI algorithms with multiple layers (Aggarwal, 2018; Wani et al., 2020).

Neural networks (NNs) are an example of DL algorithms originating from human brain understanding (Sataloff et al., 2006). It consists of neurons that process information separately and simultaneously, mimicking the functioning of the brain. The first studies on NNs were registered in 1943, in which neurons were modelled to perform logic functions (Landahl et al., 1943). In 1986, a significant discovery in the field was the multilayer perceptron (MLP) trained with the backpropagation learning

algorithm (De et al., 1986). This algorithm, which is an application of dynamic programming, can determine the optimal weight set (Aggarwal, 2018; Sataloff et al., 2006). It is divided into two phases, the forward phase, in which the inputs for a training instance are fed into the NN, and the backward phase, which aims to learn the gradient of the loss function for different weights (Aggarwal, 2018).

NNs are defined by neurons in multiple layers that process input data and generate an output according to an activation function. The activation function represents a linear or nonlinear mapping from input to output (Sataloff et al., 2006). A neuron transfers its information by synapsis, which is considered in this algorithm as the weight of each neuron on the output. The mathematics behind a neural network consider that the neuron i carries the following calculation in a time step, and the activation potential is accomplished by the summation of the weighted output values of all incoming connections (Ertel, 2017a), considering Equation (2-11), where w_i is the synaptic weight and x_j is the value of the input.

$$\sum_{j=1}^{n} w_{ij} \cdot x_j \tag{2-11}$$

Different weight values result in different outputs. Therefore, it is fundamental to determine the combination of values that results in the smallest difference between the NN output and the desired result. The weight values are updated according to the learning rate during the training process. Low learning rates increase the computational time required to achieve optimal results, whereas high values cannot approximate the optimal result (Aggarwal, 2018). Hence, starting with a higher learning rate is recommended to quickly initialise the interactive process and transition to lower values to converge towards the optimal result.

After computing these weighted sums, an activation function (2-12) is applied to the result, which varies according to the analysed data. This function should be differentiable and should satisfy a simple differential equation, permitting the evaluation of weight increments using the chain rule for partial derivatives (Elliott, 1993). Common activation functions include the sigmoid function, Rectified Linear Unit (ReLU), linear function, and hyperbolic tangent (tanh). Changes in the activation function modify the input data processing in the network, and it is fundamental to fine-tune this hyperparameter to optimise the performance of the network according to the objective of the model.

$$x_i = f(\sum_{j=1}^{n} w_{ij} \cdot x_j) \tag{2-12}$$

NNs are versatile and can combine different layers to perform different operations according to the user's needs. This is the case of the convolutional layer, which is designed to work with tri-dimensional information, such as images. Each pixel in an image has a spatial location along the horizontal and vertical

axes, and a value. This type of layer characterizes a convolutional neural network (CNN), whose architecture is named after this operation. Each layer of a CNN is determined by height, width, and depth. The depth of a CNN is the number of feature maps in each hidden layer and the values of the pixels on an RGB colour scale (Aggarwal, 2018). Hence, this type of layer is adequate for processing the numerical model results in two dimensions and satellite imagery.

During a convolutional operation, the filters slide over the input image, overlapping all the positions. At each position, a dot product is performed between the pixel values of that area of the image and the values of the filter, resulting in an output. Convolutional operations are equivariant to translation, meaning that the computed pixel value variation in the image is processed similarly regardless of its spatial location (Aggarwal, 2018). Figure 2-1 shows an example of a convolutional operation.

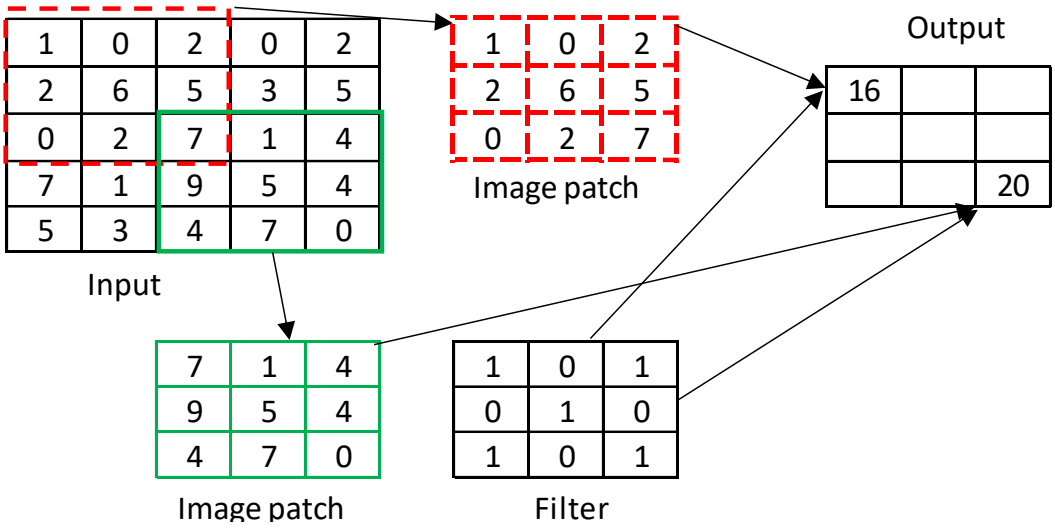


Figure 2-1. Example of convolutional operation.

Another relevant layer used in NN, specially dedicated to predicting time-series, is the long-short term memory (LSTM). LSTM networks were designed to solve the vanishing gradient problem of recurrent neural networks during backpropagation (Aggarwal, 2018). This type of architecture is ideal for processing time-series data and can deal with large data sequences. Further details and the equations that describe LSTM networks are presented in **Erro! Fonte de referência não encontrada.**.

Furthermore, AI models have another type of configuration to be defined, known as hyperparameters. Hyperparameters are values defined by the user that determine how the input is processed. For instance, the number of neurons and the activation function of an NN are hyperparameters, whereas the weight values are parameters. The main difference between these two concepts is that parameter values are adjusted during the training phase, whilst hyperparameters are

values that are not estimated from the data. Fine-tuning the values of the hyperparameters is necessary to optimise the network's computational performance and the model's accuracy.

In the case of CNNs, the number and size of filters is a hyperparameter of each convolutional layer that is defined before training the model, whereas the values of the weights in each filter are the parameters defined during the training of the network (Agrawal & Mittal, 2020; Noriega, 2005). In addition, using different activation functions in convolutional layers can result in completely different outputs. For example, using a sigmoid activation function results in outputs found in the interval (0, 1). Therefore, if negative values are important for the result, another function must be used to guarantee that the output matches the objective of the model. However, this function can be useful for filtering specific aspects of an image, such as sharp edges or vertical and horizontal lines. The same logic can also be applied to any NN.

During the training phase and the selection of hyperparameters, evaluating whether the model is hyperfitted is fundamental. Hyperfit occurs when the model presents excellent results in the training dataset but inferior performance in the validation dataset. This occurs because the model learns to predict most of the training dataset, but when it is tested on different data, it simply does not know how to consider the inputs. Hence, hyperfitted models do not present an acceptable generalising capacity. This problem generally occurs when the model architecture is overcomplex.

The opposite condition is an underfitted model, which occurs when the model is not well trained or complex enough to reproduce the data behaviour. To solve both problems, the hyperparameters of the model must be modified. A method to verify the learning condition of the model is by plotting the learning curves, which is a graphic of the accuracy (y-axis) or error in each epoch (x-axis) for both the validation and training datasets. Ideally, the learning curves will exhibit decrescent behaviour according to the epochs, stabilising in the final iterations. The error values in the training dataset should be equal to or slightly lower than the validation dataset's. If the error values in the validation dataset begin to increase in the last epochs, the model is overfitted. Conversely, if the error in the validation dataset does not approximate the training error, the model is underfitted. This method is particularly effective in DL models because the number of hyperparameters is higher than in ML models.

At this point, the equivalence between AI methods and numerical modelling must be clarified to provide a better understanding to readers unfamiliar with one of these topics. In numerical modelling methodologies, field observations are necessary to calibrate the model. The values of some parameters of the models, such as the bottom friction coefficient and tidal harmonic constituents, are adjusted to

reduce the difference between the results of a numerical model simulation and the observed field data. The correct choice of parameter values is fundamental for reducing uncertainties in the model results.

In AI methodologies, field data are also used to adjust the parameters of the model, a process called training instead of calibration. Moreover, the training objective is the same as the calibration: finding the best parameter values that reduce the model error when simulating the observed data.

Despite these similarities, the logic behind the numerical and data modelling is slightly different. While the equations that describe the physics of the problems are known in numerical modelling methods, in AI methods, these equations are determined according to the input and output data. Each has advantages and disadvantages, but both are completely trustworthy if the data used to calibrate the models have adequate quality (Maier et al., 2023).

Over the past few decades, AI methods have been believed to function as black-box solutions. It was assumed that it was impossible to know what the algorithm was computing, only to evaluate whether the results were correct or incorrect. However, this hypothesis was rejected, and it is possible to understand how NNs relate inputs to outputs (Maier et al., 2023). If certain principles are adopted, the results of the AI models are as consistent as those of the numerical models.

One of the advantages of AI models relies exactly on how it computes the input data. The input and output are associated with relatively simple mathematical operations, which vary according to the algorithm used in the case of ML, or the layers in the case of DL. Numerical models solve a set of differential equations according to the physics of the problem and the number of dimensions considered, which requires more processing capacity, particularly if the models have a high spatial resolution. Therefore, AI can be an interesting option for reducing the computational resources required by numerical models by partially or completely replacing them, known as emulation.

Emulation techniques, also known as surrogate modelling or metamodeling, are tools and/or procedures for developing computationally more efficient surrogates of high-fidelity models (Razavi et al., 2012). An emulator is a statistical approximation of a simulator (O'Hagan, 2006). This solution can optimise the resources necessary to simulate future scenarios. Kizhisseri *et al.* (2006) combined field and synthetic data from a numerical model to correlate velocity flux with suspended sediment concentration, discovering useful mathematical relationships between dependent and independent variables. Combining the strengths of data-driven and physics-based methods in a hybrid model can increase the accuracy and speed of numerical models (Goldstein & Coco, 2015). Razavi *et al.* (2012) divided these methods into two broad families: the response surface surrogate, which is a statistical or

empirical data-driven model emulating high-fidelity model responses, and lower-fidelity physically based surrogates, which are simplified models of the original system.

Many AI algorithms can be used to emulate other models, such as Evolutionary Polynomial Regression (EPR) applied by Giustolisi and Savic (2009) to predict groundwater levels based on total monthly rainfall, Artificial Neural Network (ANN) selected by Tayfur and Guldal (2006) to estimate total suspended sediments in natural streams, Convolutional Neural Network (CNN) used by Chen *et al.* (2021) to construct a flood forecasting, or Gaussian Process Regression (GPR), an emulator developed by Parker *et al.* (2019) to rapidly predict estuarine variables related to flooding. Additionally, Chen *et al.* (2018) configured a set of ANNs to surrogate an estuarine water quality model, Antolínez *et al.* (2015) introduced a framework to reduce the dimensionality of wave driver inputs to morphodynamics models using Principal Component Analysis (PCA), Bárcena *et al.* (2016) used the K-means clustering approach to analyse stratification patterns, and Ravichandran et al.(2021) used a gradient boost tree (GBT) classification model to detect leaks in water mains. Table 2-1 summarises some applications of emulation techniques.

Table 2-1: Applications of emulation techniques.

Reference	Applied method	Application
Antolinez (2015)	PCA	To model long-term shoreline evolution
Bárcena *et al.* (2016)	K-means	Stratification patterns analysis
Chen *et al. (2018)*	ANN	Estuarine water quality model
Chen *et al.* (2021)	CNN	Forecasting flood processes
Giustolisi and Savic (2009)	EPR	Predict future groundwater levels
Parker *et al.* (2019)	GPR	Predict flood-related estuarine variables
Racivhandran *et al.* (2021)	GBT	Leak detection in water distribution systems
Tayfur and Guldal (2006)	ANN	Estimate suspended sediments in streams

These studies demonstrated that surrogate techniques could reduce the computational efforts of high-resolution models, which are necessary to deal with large temporal and spatial scale simulations (Antolínez et al., 2015, 2016). These simulations are crucial for understanding the effects of climate change on hydrodynamic patterns and the evolution of fluvial morphology. In addition, the large amount of data that has been generated in recent decades makes it feasible to create methods that can produce knowledge from raw data.

Another interesting AI technique used to surrogate numerical models is the Bayesian network (BN), which utilises knowledge about the independence of variables to simplify models (Ertel, 2017a). According to the same author, configuring a BN proceeds in two steps: the network structure's design and the probabilities of events. BNs have already been applied to hindcast morphological changes, achieving high accuracy in forecasting (Wilson et al., 2015) and predicting coastal hazards for sandy coasts (Poelhekke et al., 2016). Genetic Programming (GP) has also been used to assess suspended sediments' velocity flux and predict settling velocity (Goldstein & Coco, 2014; Kizhisseri et al., 2006). Table 2-2 presents some applications of emulation techniques applied to morphodynamics.

Table 2-2. Emulation methods applied to morphodynamics.

Reference	Applied method	Application	Results
Yoon *et al.* (2013)	ANN	Prediction of time-dependent sediment suspension	Improvement of erosion prediction capability by 40%
Kizhisseri *et al.* (2006)	GP	Correlate velocity flux with suspended sediment concentration	Maximum error of 0.035 m^2/s between observed and simulated data of sediment load
Mafi et al. (2013)	DT	Prediction of longshore sediment transport	Correlation = 0.95 Bias = -0.006
Goldstein and Coco (2014)	GP	Prediction of settling velocity	Normalized root square mean error = 0.26 Correlation = 0.97
Wilson *et al.* (2015)	BN	Hindcast of morphological changes	81.9% accuracy in predicting volumetric changes of an island
Poelhekke *et al.* (2016)	BN	Predicting coastal hazards for sandy coasts	81% - 85% accuracy in predicting onshore hazards 91% - 95% accuracy in predicting no significant onshore hazards

Finally, this section presented a literature review of AI methods, explaining important concepts that can help to understand better the methodology developed in this thesis. The next chapters focus on the application of these methods to different datasets, highlighting the application of the methods, the characteristics of the input and output data, and how the performance of the models was evaluated.

2.4.1. Tools to implement artificial intelligence models.

With the increasing popularity of AI models in all areas of knowledge, many programming languages have created appropriate libraries to implement them. For instance, the Caffe library is a

popular AI framework which works in C++, the scikit-learn is an open-source ML library which works in Python (Pedregosa et al., 2011), and the TensorFlow framework (Abadi et al., 2016), adequate for implementing more complex ML and DL models, is also an open-source library that works in Java and Python.

Python is one of the most popular programming languages for AI algorithms, but others, such as Julia and R, also provide a wide range of libraries. Python is considered an easy language to learn, has a large ecosystem of scientific libraries, and the code is executed immediately in the console without the compilation step to machine language (Irizarry, 2019).

The choice of language and framework to adopt will depend on the user's experience with the language and if the language supports the requirements of the AI model. The scikit-learn library provides excellent ML models, although it is currently impossible to customize them fully. For instance, the activation function of a multilayer perceptron (MLP) network implemented using scikit-learn must be selected according to the available options. Conversely, the TensorFlow library is more adequate when customisation is required to implement more complex models. It allows the customisation of activation functions or sets different operations for each input and posteriorly combines them inside the network architecture to generate the output.

In addition, it is necessary to manipulate and organise the data to implement an AI model successfully. This step consists of removing outliers or anomalous values, characterising the data using descriptive statistics and graphics, splitting the data into training and testing datasets to calibrate and validate the model, and normalising the data, which is an optional task that will be necessary according to the algorithm used. Only then will AI libraries be used to implement this model. Previous data analysis is fundamental to ensure the data model generates adequate results.

To perform these tasks, the Python language provides libraries that are automatically coupled with TensorFlow and scikit-learn models. The NumPy library allows the manipulation of multidimensional array objects, computation of shape manipulation, and mathematical and statistical operations (Harris et al., 2020). The Pandas library is an efficient option for manipulating data frames such as tables. It provides functions that can easily organise, analyse, read, and export data in different formats, such as csv and xlsx (Pandas development team, 2023). Matplotlib and Seaborn are the libraries used to create graphics and figures, which help visualise the data and assess the results of the AI models. The first is a more comprehensive library, which allows the creation of figures which are easily edited (Hunter, 2007), while the second is focused on statistical data visualisation optimised to datasets integrated with Pandas data frames (Waskom, 2021).

Therefore, the implementation of an AI model can be divided into three activities:

- Data cleaning and organising: Consists of assessing data quality and quantity to ensure that it will be possible to implement the AI model. In this task, graphics can help to see the characteristics of the dataset, and correlations can be computed to guarantee that the inputs and outputs are related. If the input and output data are not correlated, the chances of achieving success when implementing the AI model are reduced. Moreover, transforming the input variables may help achieve good results, which must be performed at this stage. This activity consumes a significant portion of the time required to implement the AI model.
- AI model implementation: In this phase, an AI solution is defined. The AI algorithm is selected, the training and testing datasets are split, and the architecture of the model is chosen (in the case of DL models). The hyperparameters of the models are fine-tuned during this activity.
- Analysis of the results: After training the AI model, the capacity of the model to process data different from the training dataset must be evaluated to ensure that the model can generalise and is neither hyperfitted nor underfitted. Metrics such as the root mean squared error (RMSE), mean squared error (MSE), absolute error (AE), and coefficient of determination (R^2) are computed to compare the simulated results with the observed data.

2.5. Operational forecasting

Operational forecasting is generally described as the use of real-time collected data in models that forecast the evolution of variables of interest in different time horizons. This type of system can be widely used to predict any variable, such as atmospheric temperature, ocean currents, air quality, water levels, and earthquakes impacts. To establish an operational forecasting system, it is crucial to promptly assess all accessible information and send it to the predictor as soon as possible (Giaiotti & Stel, 2007). Suppose the lag between collecting data and concluding the simulation is substantial. In that case, the practical usefulness of the system is compromised, leading to an inability to incorporate the simulation results in the decision-making process.

When applied to forecasting natural disaster outcomes, operational forecasting systems are known as early warning systems (EWS). The implementation of EWS is a central part of the Sustainable Development Goal (SDG) 11 of the United Nations (UN) - Make cities and human settlements inclusive, safe, resilient, and sustainable, and is strongly interconnected with three other SDGs: 1) SDG 3.d –

Strengthen the capacity of all countries for EWS, risk reduction and management of national and global health risks; 2) SDG 13.3 - Improve education, awareness-raising and human and institutional capacity on climate change mitigation, adaptation, impact reduction and early warning; 3) SDG 15.1 – ensure the conservation, restoration and sustainable use of terrestrial and inland freshwater ecosystems and their services, in particular forests, wetlands, mountains and drylands, in line with obligations under international agreements. According to Bertola et al. (2023), 95% of extreme floods that occurred in Europe between 1999 and 2021 could have been anticipated by sharing and pooling data across the continent, demonstrating the importance of EWS.

EWS are composed of five pillars (WMO, 2022): 1) disaster risk knowledge that aims to identify and map hazards and vulnerabilities on a local scale; 2) observations and forecasting that monitor risk-related parameters and implement forecasting and warning systems; 3) warning dissemination and communication that focuses on the dissemination of warnings to reach everyone in the risk area with concise information; 4) preparedness and response capabilities that are concerned with the population reaction during a risk scenario; and 5) monitoring and evaluation to promote a continuous assessment of the EWS, looking forward to identifying possible flaws and improving the system.

Each pillar is relevant to a fully operational EWS, and combining all the pillars ensures an efficient and robust system. The first pillar is essential for surveying the most vulnerable areas and risk exposure at a local scale. This component ensures the usefulness of the EWS at the community level by accurately identifying the most urgent hazards. The second pillar is the basis for forecasting future scenarios. Hazard-related parameters are identified, and the modelling components are implemented in this phase. In this thesis, we specifically focus on the implementation of this pillar for flood early warning systems (FEWS). The third component consists of communicating disaster occurrences to the population and ensuring that the warnings reach everyone at risk with clear information. The fourth pillar consists of defining emergency plans for disaster situations and guaranteeing that the population knows how to respond. Lastly, the fifth pillar ensures that all components work together adequately.

The forecasting component of FEWS assembles several data types and combines them with models that forecast the future state of variables of interest. In the case of predicting floods in an estuary, the parameter that we must predict is the water level along the estuary. Therefore, it is necessary to implement a model that reliably simulates water levels in the study area by forcing the boundaries of the model under future conditions. Hence, it is fundamental to forecast the evolution of the boundary variables.

The marine boundary is relatively simple to configure because of the existence of global tidal models that provide future water levels using tidal harmonic constituents, such as TPXO Global Tidal Models, considering that other ocean water level anomalies do not influence the study area. Conversely, the major problem is forecasting fluvial discharge, the intensity of which depends on more complex factors, namely precipitation, past discharges, and basin characteristics.

Hydrological models can simulate fluvial discharge based on precipitation and topographic data. However, the complexity of this phenomenon creates obstacles that are extremely challenging to model, thereby resulting in overcomplex models. Differences in soil physical characteristics, soil use and occupation, basin slope, and vegetation influence fluvial discharge. Hence, the larger the hydrographic basin, the higher the complexity of the model. This problem becomes even more complex when hydropower plants are considered, adding the energy consumption variable to the fluvial discharge equation. Hydrological models are extremely useful for understanding how anthropogenic actions affect runoff patterns, although the uncertainties and computational resources required by this type of model can be obstacles to implementing FEWS.

During the execution of an EWS, field data should be registered and sent to the predictor, which uses this data to update its state and forecast the evolution of the monitored risk variable. If a threat is identified in this simulation, the emergency teams and population should receive the warning, and the emergency plan must be executed to reduce potential losses. Consequently, the longer the predictor processing time, the faster the emergency response must be. Hence, the accuracy and computational time of the forecasting component of an EWS are critical variables.

In this context, optimising the modelling component is fundamental for balancing accuracy and processing power. High-resolution models may result in more precise results but require more computational capacity. In contrast, low-resolution models can provide fast answers, and their lack of skill may overestimate or underestimate the risk. Therefore, the model should have a high resolution in the most endangered areas to ensure that the results are accurate in critical locations, whilst lower resolutions are recommended in areas where the risk is irrelevant.

2.6. References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., ... Zheng, X. (2016). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. http://arxiv.org/abs/1603.04467

Aggarwal, C. C. (2018). Neural Networks and Deep Learning. In *Artificial Intelligence*. Springer International Publishing. https://doi.org/10.1007/978-3-319-94463-0

Agrawal, A., & Mittal, N. (2020). Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy. *The Visual Computer*, *36*(2), 405–412. https://doi.org/10.1007/s00371-019-01630-9

Akan, A. O. (2006a). Fundamentals of open-channel flow. In *Open Channel Hydraulics* (pp. 1–23). Elsevier. https://doi.org/10.1016/B978-075066857-6/50002-3

Akan, A. O. (2006b). Introduction to unsteady open-channel flow. In *Open Channel Hydraulics* (Vol. 1, pp. 315–359). Elsevier. https://doi.org/10.1016/B978-075066857-6/50009-6

Antolínez, J. A. A., Méndez, F. J., Camus, P., Gonzélez, E. M., & Ruggiero, P. (2015). A CLIMATE-BASED MULTIVARIATE WAVE EMULATOR FOR LONG-TERM MORPHODYNAMIC SIMULATIONS. *The Proceedings of the Coastal Sediments 2015*, 1–16. https://doi.org/10.1142/9789814689977_0201

Antolínez, J. A. A., Méndez, F. J., Camus, P., Vitousek, S., González, E. M., Ruggiero, P., & Barnard, P. (2016). A multiscale climate emulator for long-term morphodynamics (MUSCLE-morpho). *Journal of Geophysical Research: Oceans*, *121*(1), 775–791. https://doi.org/10.1002/2015JC011107

Azevedo-Santos, V. M., Daga, V. S., Tonella, L. H., Ruaro, R., Arcifa, M. S., Fearnside, P. M., & Giarrizzo, T. (2023). Brazil undermines flood resilience. *Science*, *382*(6669), 383–383. https://doi.org/10.1126/science.adl0891

Balmforth, N. J., Llewellyn-Smith, S., Hendershott, M., & Garrett, C. (2005). 2004 program of study : tides. In *2004 Program of Study : Tides*. Woods Hole Oceanographic Institution. https://doi.org/10.1575/1912/98

Bárcena, J. F., García-Alba, J., García, A., & Álvarez, C. (2016). Analysis of stratification patterns in river-influenced mesotidal and macrotidal estuaries using 3D hydrodynamic modelling and K-means clustering. *Estuarine, Coastal and Shelf Science*, *181*, 1–13. https://doi.org/10.1016/j.ecss.2016.08.005

Bartels, G. K., Castro, N. M. dos R., Collares, G. L., & Fan, F. M. (2021). Performance of bedload transport equations in a mixed bedrock–alluvial channel environment. *CATENA*, *199*(April 2020), 105108. https://doi.org/10.1016/j.catena.2020.105108

Bertin, X., de Bakker, A., van Dongeren, A., Coco, G., André, G., Ardhuin, F., Bonneton, P., Bouchette, F., Castelle, B., Crawford, W. C., Davidson, M., Deen, M., Dodet, G., Guérin, T., Inch, K., Leckler, F., McCall, R., Muller, H., Olabarrieta, M., … Tissier, M. (2018). Infragravity waves: From driving mechanisms to impacts. In *Earth-Science Reviews* (Vol. 177, pp. 774–799). Elsevier B.V. https://doi.org/10.1016/j.earscirev.2018.01.002

Bertola, M., Blöschl, G., Bohac, M., Borga, M., Castellarin, A., Chirico, G. B., Claps, P., Dallan, E., Danilovich, I., Ganora, D., Gorbachova, L., Ledvinka, O., Mavrova-Guirguinova, M., Montanari, A., Ovcharuk, V., Viglione, A., Volpi, E., Arheimer, B., Aronica, G. T., … Zivkovic, N. (2023). Megafloods in Europe can be anticipated from observations in hydrologically similar catchments. *Nature Geoscience*, *16*(11), 982–988. https://doi.org/10.1038/s41561-023-01300-5

Capellán-Pérez, I., Arto, I., Polanco-Martínez, J. M., González-Eguino, M., & Neumann, M. B. (2016). Likelihood of climate change pathways under uncertainty on fossil fuel resource availability. *Energy & Environmental Science*, *9*(8), 2482–2496. https://doi.org/10.1039/C6EE01008C

Chen, C., Hui, Q., Xie, W., Wan, S., Zhou, Y., & Pei, Q. (2021). Convolutional Neural Networks for forecasting flood process in Internet-of-Things enabled smart city. *Computer Networks*, *186*(June 2020), 107744. https://doi.org/10.1016/j.comnet.2020.107744

Chen, Roy, S. B., & Hutton, P. H. (2018). Emulation of a process-based estuarine hydrodynamic model. *Hydrological Sciences Journal*, *63*(5), 783–802. https://doi.org/10.1080/02626667.2018.1447112

De, R., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 1: Foundations*, 318–362.

de Swart, H. E., & Zimmerman, J. T. F. (2009). Morphodynamics of Tidal Inlet Systems. *Annual Review of Fluid Mechanics*, *41*(1), 203–229. https://doi.org/10.1146/annurev.fluid.010908.165159

Deltares. (2018). *Delft3D-FLOW: Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments. User Manual*.

Des, M., DeCastro, M., Sousa, M. C., Dias, J. M., & Gómez-Gesteira, M. (2019). Hydrodynamics of river plume intrusion into an adjacent estuary: The Minho River and Ria de Vigo. *Journal of Marine Systems*, *189*(May 2018), 87–97. https://doi.org/10.1016/j.jmarsys.2018.10.003

Einstein, H. A. (1950). *The bed-load function for sediment transportation in open channel flows* (Issue 1026). US Government Printing Office.

Elko, N., Feddersen, F., Foster, D., Hapke, C., Mcninch, J., Mulligan, R., Ozkan-haller, H. T., Plant, N., & Raubenheimer, B. (2014). The future of nearshore processes research. *Coastal Forum*, *December*.

Elliott, D. L. (1993). A Better Activation Function for Artificial Neural Networks. In *ISR Technical Report* (Vol. 13, Issue 96).

Ertel, W. (2017a). *Introduction to Artificial Intelligence*. Springer International Publishing. https://doi.org/10.1007/978-3-319-58487-4

Ertel, W. (2017b). *Introduction to Artificial Intelligence*. Springer International Publishing. https://doi.org/10.1007/978-3-319-58487-4

Franz, D. D., & Melching, C. S. (1997). *Full Equations (FEQ) model for the solution of the full, dynamic equations of motion for one-dimensional unsteady flow in open channels and through control structures*.

Friedrichs, C. T. (2011). Tidal Flat Morphodynamics. In *Treatise on Estuarine and Coastal Science* (Vol. 3, pp. 137–170). Elsevier. https://doi.org/10.1016/B978-0-12-374711-2.00307-7

Garcia, M., Ramirez, I., Verlaan, M., & Castillo, J. (2015). Application of a three-dimensional hydrodynamic model for San Quintin Bay, B.C., Mexico. Validation and calibration using OpenDA. *Journal of Computational and Applied Mathematics*, *273*, 428–437. https://doi.org/10.1016/J.CAM.2014.05.003

Giaiotti, D. B., & Stel, F. (2007). Atmospheric Convection: Research and Operational Forecasting Aspects. In D. B. Giaiotti, R. Steinacker, & F. Stel (Eds.), *Atmospheric Convection: Research and Operational Forecasting Aspects* (Vol. 475). Springer Vienna. https://doi.org/10.1007/978-3-211-69291-2

Giustolisi, O., & Savic, D. A. (2009). Advances in data-driven analyses and modelling using EPR-MOGA. *Journal of Hydroinformatics*, *11*(3–4), 225–236. https://doi.org/10.2166/hydro.2009.017

Goldstein, E. B., & Coco, G. (2014). A machine learning approach for the prediction of settling velocity. *Water Resources Research*, *50*(4), 3595–3601. https://doi.org/10.1002/2013WR015116

Goldstein, E. B., & Coco, G. (2015). Machine learning components in deterministic models: hybrid synergy in the age of data. *Frontiers in Environmental Science*, *3*(APR), 1–4. https://doi.org/10.3389/fenvs.2015.00033

Gong, W., Chen, L., Zhang, H., Yuan, L., & Chen, Z. (2020). Plume Dynamics of a Lateral River Tributary Influenced by River Discharge From the Estuary Head. *Journal of Geophysical Research: Oceans*, *125*(2), 1–25. https://doi.org/10.1029/2019JC015580

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science |& Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Iglesias, I., Venâncio, S., Pinho, J. L., Avilez-Valente, P., & Vieira, J. M. P. (2019). Two Models Solutions for the Douro Estuary: Flood Risk Assessment and Breakwater Effects. *Estuaries and Coasts*, *42*(2), 348–364. https://doi.org/10.1007/s12237-018-0477-5

Irizarry, R. A. (2019). Introduction to Data Science. In *Introduction to Data Science*. https://doi.org/10.1201/9780429341830

Jackson, D., & Short, A. (2020). *Sandy beach morphodynamics*. Elsevier.

Kapelan, Z., Weisbord, E., & Babovic, V. (2020). *Digital Water Artificial Intelligence Solutions for the Water Sector*. International Water Association. https://iwa-network.org/wp-content/uploads/2020/08/IWA_2020_Artificial_Intelligence_SCREEN.pdf

Karathanasi, F. E., & Belibassakis, K. A. (2019). A cost-effective method for estimating long-term effects of waves on beach erosion with application to Sitia Bay, Crete. *Oceanologia*, *61*(2), 276–290. https://doi.org/10.1016/j.oceano.2018.12.001

Kizhisseri, A. S., Simmonds, D., Rafiq, Y., & Borthwick, M. (2006). An Evolutionary Computation Approach to Sediment Transport Modelling. *Coastal Dynamics 2005*, *1963*, 1–14. https://doi.org/10.1061/40855(214)81

Kubat, M. (2017). An Introduction to Machine Learning. In *An Introduction to Machine Learning*. https://doi.org/10.1007/978-3-319-63913-0

Landahl, H. D., McCulloch, W. S., & Pitts, W. (1943). A statistical consequence of the logical calculus of nervous nets. *The Bulletin of Mathematical Biophysics*, *5*(4), 135–137. https://doi.org/10.1007/BF02478260

Li, C., Cheng, X., Li, N., Du, X., Yu, Q., & Kan, G. (2016). A Framework for Flood Risk Analysis and Benefit Assessment of Flood Control Measures in Urban Areas. *International Journal of Environmental Research and Public Health*, *13*(8), 787. https://doi.org/10.3390/ijerph13080787

Luijendijk, A., Hagenaars, G., Ranasinghe, R., Baart, F., Donchyts, G., & Aarninkhof, S. (2018). The State of the World's Beaches. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-24630-6

Mafi, S., Yeganeh-Bakhtiary, A., & Kazeminezhad, M. H. (2013). Prediction formula for longshore sediment transport rate with M5' algorithm. *Journal of Coastal Research*, *165*(65), 2149–2154. https://doi.org/10.2112/SI65-363.1

Maier, H. R., Galelli, S., Razavi, S., Castelletti, A., Rizzoli, A., Athanasiadis, I. N., Sànchez-Marrè, M., Acutis, M., Wu, W., & Humphrey, G. B. (2023). Exploding the myths: An introduction to artificial neural networks for prediction and forecasting. In *Environmental Modelling and Software* (Vol. 167). Elsevier Ltd. https://doi.org/10.1016/j.envsoft.2023.105776

Melo, W., Pinho, J., Iglesias, I., Bio, A., Avilez-Valente, P., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2020). Hydro- and Morphodynamic Impacts of Sea Level Rise: The Minho Estuary Case Study. *Journal of Marine Science and Engineering*, *8*(6), 441. https://doi.org/10.3390/jmse8060441

Nicholls, R. J. (1995). Coastal megacities and climate change. *GeoJournal*, *37*(3), 369–379. https://doi.org/10.1007/BF00814018

NOAA, N. C. for E. I. (2024). *Monthly National Climate Report for Annual 2023*. https://www.ncei.noaa.gov/access/monitoring/monthly-report/national/202313.

Noriega, L. (2005). Multilayer perceptron tutorial. *School of Computing. Staffordshire University*, 1–12.

O'Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety*, *91*(10–11), 1290–1300. https://doi.org/10.1016/j.ress.2005.11.025

Pandas development team, T. (2023). *pandas-dev/pandas: Pandas*. Zenodo. https://doi.org/10.5281/zenodo.7979740

Parker, K., Ruggiero, P., Serafin, K. A., & Hill, D. F. (2019). Emulation as an approach for rapid estuarine modeling. *Coastal Engineering*, *150*(March), 79–93. https://doi.org/10.1016/j.coastaleng.2019.03.004

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel V. and Thirion, B., Grisel, O., Blondel, M., Prettenhofer P. and Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pinho, J. L. (2000). *Aplicação de modelação matemática ao estudo da hidrodinâmica e da qualidade da água em zonas costeiras*.

Poelhekke, L., Jäger, W. S., van Dongeren, A., Plomaritis, T. A., McCall, R., & Ferreira, Ó. (2016). Predicting coastal hazards for sandy coasts with a Bayesian Network. *Coastal Engineering*, *118*, 21–34. https://doi.org/10.1016/j.coastaleng.2016.08.011

Portela, L. I. (2008). Sediment transport and morphodynamics of the Douro River estuary. *Geo-Marine Letters*, *28*(2), 77–86. https://doi.org/10.1007/s00367-007-0091-1

Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., & Rama, B. (2022). *IPCC, 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. https://doi.org/10.1017/9781009325844

Ravichandran, T., Gavahi, K., Ponnambalam, K., Burtea, V., & Mousavi, S. J. (2021). Ensemble-based machine learning approach for improved leak detection in water mains. *Journal of Hydroinformatics*, 1–17. https://doi.org/10.2166/hydro.2021.093

Razavi, S., Tolson, B. A., & Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, *48*(7). https://doi.org/10.1029/2011WR011527

Richardson, J. F., & Zaki, W. N. (1997). Sedimentation and fluidisation: Part I. *Chemical Engineering Research and Design*, *75*(1 SUPPL.), S82–S100. https://doi.org/10.1016/S0263-8762(97)80006-8

Roelvink, D., & Reniers, A. (2011). *A Guide to Modeling Coastal Morphology* (Vol. 12). WORLD SCIENTIFIC. https://doi.org/10.1142/7712

Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., & Lescinski, J. (2009). Modelling storm impacts on beaches, dunes and barrier islands. *Coastal Engineering*, *56*(11–12), 1133–1152. https://doi.org/10.1016/j.coastaleng.2009.08.006

Saes, R., Valentim, S., Buchmann, F., de Morais, P., Bezerra, M., Piedras, F., Gaeta, J., Moreira, L., Sahm, B., & Damasceno, É. (2018). *Princípios de oceanografia*. EDUFPI.

Sanchez-arcilla, A. A., Jimenez, J. A., Valdemoro, H. I., Summer, F., & Sanchez-arcilla, A. (1998). The Ebro Delta : Morphodynamics and Vulnerability. *Journal of Coastal Research*, *14*(3), 754–772.

Sataloff, R. T., Johns, M. M., & Kost, K. M. (2006). *Neural Networks in a Softcomputing Framework*. Springer-Verlag. https://doi.org/10.1007/1-84628-303-5

Seminara, G. (1998). Stability and Morphodynamics. *Meccanica*, *33*(1), 59–99. https://doi.org/10.1023/A:1004225516566

Shen, X., Lee, B. J., Fettweis, M., & Toorman, E. A. (2018). A tri-modal flocculation model coupled with TELEMAC for estuarine muds both in the laboratory and in the field. *Water Research*, *145*, 473–486. https://doi.org/10.1016/J.WATRES.2018.08.062

Sousa, M. C., Vaz, N., Alvarez, I., Gomez-Gesteira, M., & Dias, J. M. (2014). Influence of the Minho River plume on the Rias Baixas (NW of the Iberian Peninsula). *Journal of Marine Systems*, *139*, 248–260. https://doi.org/10.1016/j.jmarsys.2014.06.012

Spencer, R. W. (2008). *The Discovery of Global Warming*. Harvard University Press.

Stansby, P. K. (2013). Coastal hydrodynamics – present and future. *Journal of Hydraulic Research*, *51*(4), 341–350. https://doi.org/10.1080/00221686.2013.821678

Talley, L. D., Pickard, G. L., Emery, W. J., & Swift, J. H. (2011). Gravity Waves, Tides, and Coastal Oceanography. In *Descriptive Physical Oceanography* (pp. 223–244). Elsevier. https://doi.org/10.1016/B978-0-7506-4552-2.10008-3

Tayfur, G., & Guldal, V. (2006). Artificial neural networks for estimating daily total suspended sediment in natural streams. *Nordic Hydrology*, *37*(1), 69–79. https://doi.org/10.2166/nh.2005.031

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., & Feyen, L. (2017). Extreme sea levels on the rise along Europe's coasts. *Earth's Future*, *5*(3), 304–323. https://doi.org/10.1002/2016EF000505

Wani, M. A., Bhat, F. A., Afzal, S., & Khan, A. I. (2020). *Introduction to Deep Learning* (pp. 1–11). https://doi.org/10.1007/978-981-13-6794-6_1

Waskom, M. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

Weber de Melo, W., Pinho, J., Iglesias, I., Bio, A., Avilez-Valente, P., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2022). *Flood Risk Assessment at the Douro River Estuary* (pp. 37–49). https://doi.org/10.1007/978-981-16-5501-2_4

Wilson, K. E., Adams, P. N., Hapke, C. J., Lentz, E. E., & Brenner, O. (2015). Application of Bayesian Networks to hindcast barrier island morphodynamics. *Coastal Engineering*, *102*, 30–43. https://doi.org/10.1016/j.coastaleng.2015.04.006

WMO, W. M. O. (2022). *Early Warnings for all - The UN Global Early Warning Initiative for the Implementation of Climate Adaptation - Executive Action Plan 2023-2027*. https://library.wmo.int/idurl/4/58209

WMO, W. M. O. (2024). *State of the Global Climate 2023*. https://library.wmo.int/idurl/4/68835

Xiang, Y., Fu, Z., Meng, Y., Zhang, K., & Cheng, Z. (2019). Analysis of wave clipping effects of plain reservoir artificial islands based on MIKE21 SW model. *Water Science and Engineering*, *12*(3), 179–187. https://doi.org/10.1016/J.WSE.2019.08.002

Yoon, H.-D., Cox, D. T., & Kim, M. (2013). Prediction of time-dependent sediment suspension in the surf zone using artificial neural network. *Coastal Engineering*, *71*, 78–86. https://doi.org/10.1016/j.coastaleng.2012.08.005

CHAPTER 3.

A data model to forecast the morphological evolution of multiple beach profiles.

Article published in July 2024 in the Coastal Engineering journal.

Weber de Melo W, Pinho J, Iglesias I. A data model to forecast the morphological evolution of multiple beach profiles. Coastal Engineering 2024:104574. https://doi.org/10.1016/j.coastaleng.2024.104574

"Ensinar não é transferir conhecimento, mas criar as possibilidades para a sua própria produção ou a sua construção."

Paulo Freire (1921 – 1997) – Brazilian educator and philosopher. Quote from Pedagogy of Autonomy.

Chapter 3. A data model to forecast the morphological evolution of multiple beach profiles.

Abstract

Beaches are a natural defense against extreme events, such as storms and hurricanes, whose intensity and frequency are expected to increase in the future due to climate change. In this context, models that forecast the morphological evolution of coastal areas can be used to anticipate the effects of future scenarios, allowing early action to mitigate the damage caused by extreme events. Hence, this study included data from three different monitoring programs in data models to simulate the seasonal morphological evolution of several Portuguese beaches. Two different data models were implemented using the Random Forest algorithm. One was fed with profile data and wave conditions while the other considered also sediment size data. Both models achieved suitable performances, but the inclusion of sediment data reduced the model errors and variance, and thus improved model performance. It was demonstrated that combining data from multidisciplinary campaigns can be a solution to generate reliable and robust morphological forecasting models.

3.1. Introduction

Coastal zones shelter about 10% of the world's population. These regions are among the most threatened areas exposed to the effects of storms and fluvial and coastal floods. Additionally, in the actual context of climate change, it is expected a sea level rise and an increase in the frequency and intensity of extreme events due to the augment of the global mean temperature (McGranahan et al., 2007; Pörtner et al., 2022). In the last decade, most European coastal cities experienced a sea level rise, which increased the flood risk exposure of these populations and their vulnerability (UN-Habitat, 2022).

In coastal zones, beaches play a crucial role as natural barriers in protecting upland structures and infrastructure against the impact of extreme events, namely storms and wave action (Mil-Homens et al., 2022). The wave height and frequency, jointly with the coastal currents, the sediment size, and the beach type (dissipative, intermediate, or reflective), are the main morphodynamic drivers at coastal zones (Aagaard et al., 2013; Prodger et al., 2017). The front dunes at the shoreface dissipate part of the wave energy, reducing the wave sediment transport capacity and erosion on the backshore. However, their stability depends on a delicate sediment balance, considering all sediment particles entering or leaving the system. If this sum is positive, the area is being accreted; otherwise, it is being eroded.

Beach morphological evolution has been the focus of much research in the last decade, including studies that represent current and projected erosive/accretional patterns at several locations using different tools. Vousdoukas et al. (2020) investigated the trends of shoreline dynamics considering the

effects of different climate change scenarios resulting from the analysis of synthetic data series generated by the Monte Carlo method, Karunarathna et al. (2012) implemented a cross-shore beach profile evolution predictor by merging a data-driven approach with a diffusion-type model, Gunawan et al. (2023) implemented a data model to forecast shoreline changes using satellite images, and Kalligeris et al. (2020) implemented two process-based numerical models to forecast the beach profile evolution in California. All the mentioned works identified that numerical models had low skill when predicting multiple beach profiles with a single set of parameters.

In this context, promoting monitoring projects in coastal areas is paramount to tracking coastal morphological evolution. While long-term observations can increase the knowledge about the causes of the morphological change and be a reliable source of information to decision-makers (Benveniste et al., 2019), short-term observations are indispensable to monitoring the response of highly dynamic beaches and the effects of extreme events (Jayson-Quashigah et al., 2019). The data collected by monitoring projects is fundamental for the application of modelling methodologies, such as numerical and data modelling, which are highly dependent on the quality and quantity of field data.

Numerical modelling methodologies rely on monitoring data to calibrate the models and ensure their results are valid (Biolchi et al., 2022; Matheen et al., 2021; Suzuki and Cox, 2021). However, these tools need considerable computational resources, which will be even more extensive if different sediment classes are considered along the computational domain (Prodger et al., 2017; Weber de Melo et al., 2022). Morphodynamic numerical models, such as Delft3D and XBeach (Deltares, 2018; Roelvink et al., 2009), solve the shallow water equations coupled with sediment transport equations, updating the bathymetry after each hydrodynamic timestep. Although these models can achieve excellent results, their computational costs tend to be high, particularly for high-resolution models, which can be a critical restriction for adopting this solution as the main forecasting tool for operational systems.

Data modelling, which is based on artificial intelligence (AI) methods such as machine learning (ML) and deep learning (DL) techniques, depends not only on the data quality but also on its quantity to achieve adequate performance metrics (Razavi et al., 2012; Weber de Melo et al., 2023; Ye et al., 2021). These models have lower computational costs than numerical models to produce accurate solutions (Antolínez et al., 2016; Gharagozlou et al., 2022; Weber de Melo et al., 2023), even if different sediment classes are considered in the computation. AI-based models compute simpler equations resulting from training the model's parameters using input and output data. Therefore, data models should be considered an alternative and valuable tool for implementing methodologies capable of forecasting the morphological evolution of beaches.

AI algorithms have already been applied to assess shoreline evolution (McAllister et al., 2022), but most applications use satellite imagery as the main data source. Luijendijk et al. (2018) implemented a classification and regression tree model to determine the water-land transition using images from the Google Earth Engine, and Chen et al. (2023) implemented a physical-informed neural network to map the nearshore bathymetry also using satellite remote sensing data. These studies demonstrate the potential of AI to study the evolution of coastal morphology. However, few studies have combined these methods with data collected in-situ. While remote-sensing data is an easy and affordable method to monitor large areas, the resolution of publicly available images is restricted to 10 m (McAllister et al., 2022), which is inadequate to monitor the evolution of beach profiles with sufficient accuracy. This strongly limits the options for those who cannot afford to buy or measure beach profiles with higher-resolution approaches.

Hence, this study contributes to filling the gap in the applications of AI algorithms using in-situ measured beach profiles. In this study, the outcomes of multiple coastal monitoring programs were merged to implement two data models for forecasting the morphological evolution of multiple seasonal beach profiles along the northern Portuguese coast. The data models were implemented using the random forest (RF) algorithm. The first model (M1) used beach profile historical data, wave variables, and sand size data as inputs, and the second model (M2) used the same data as M1 but omitted sediment size data to understand the importance of this variable in the forecast. The outputs of both models were twelve beach profiles from eight different beaches in northern Portugal. Lastly, the forecast of beach slope and beach retreat/advance was computed to assess the reliability of the data model results. In addition, the models were tested under storm conditions to evaluate their performance outside the training data range.

3.2. Methodology

This study applied the RF algorithm to develop two data models that are able to predict the seasonal beach profile evolution due to wave action. This section describes the study area, the sources and characteristics of the data, and the main steps to implement the RF models.

The methodology framework (Figure 3-1) summarizes the main tasks required to implement an AI model, as well as the specifications adopted in this study. The first task included the identification of data sources, the data collection, and the characterization of the study area. The second task addresses data processing, ensuring that the data type, values, and format are adequate for the selected algorithm. The data is usually normalized, outliers are removed, and features are combined to reduce data that may not be relevant to the model's objective. The necessity of normalization depends on the ML algorithm

used. In the model implementation task, an algorithm and a training technique are defined according to the characteristics of the data and the algorithm capacity. The last task is focused on validating the obtained results using different metrics. Details about how each task was executed in this study are described in the following sections.

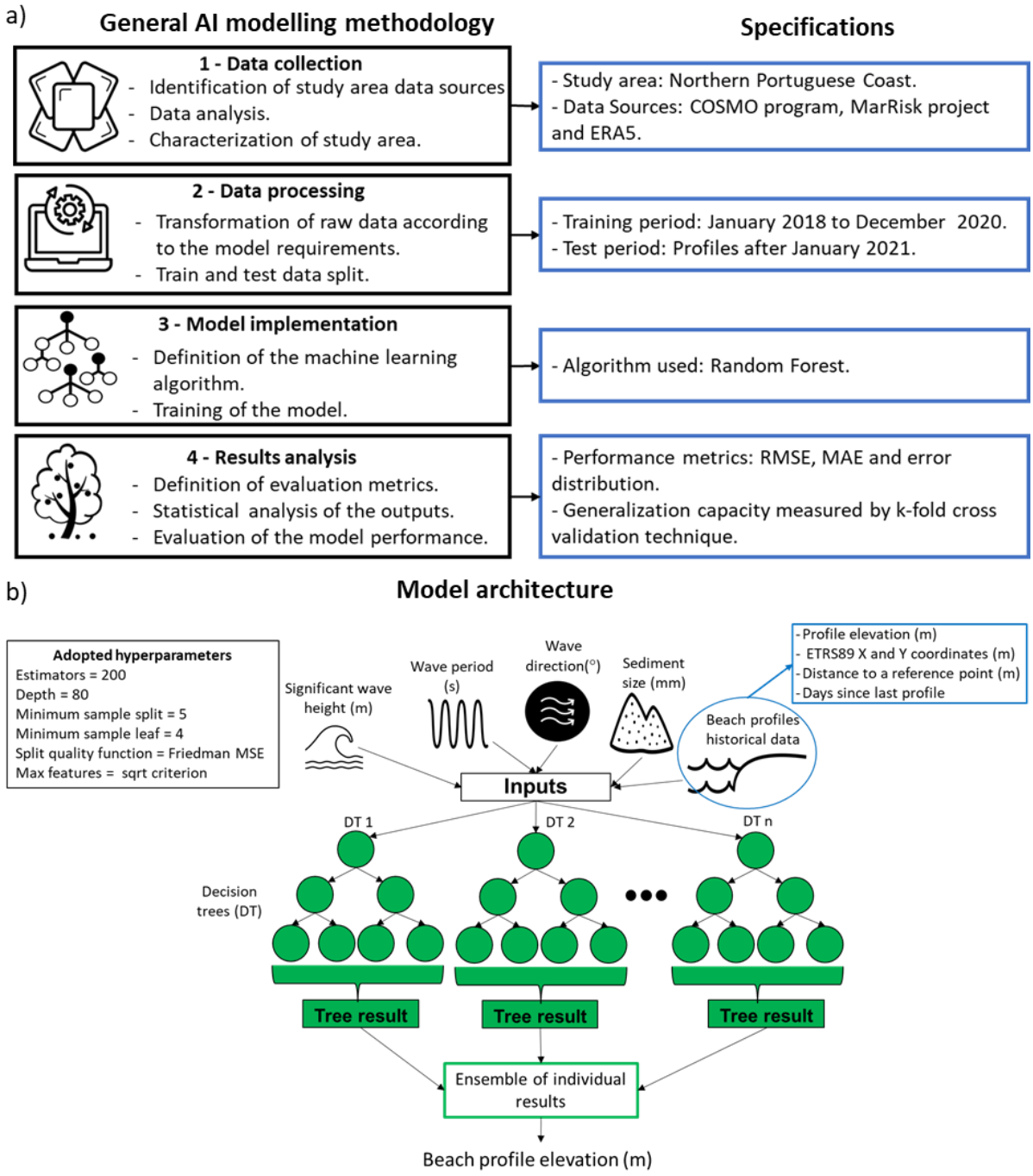


Figure 3-1. Methodological approach: a) Methodology framework with the steps required to implement a data model . b) Architecture of the model and details on the input data and adopted hyperparameters.

3.2.1. Study area

Northern Portugal has a Mediterranean climate dominated by the Azores' high-pressure system, with wet winters and dry summers (Sanderson et al., 2023). The region is characterized by high mesotidal conditions and wave-dominated dynamics, with waves exhibiting a predominantly northwesterly orientation. The significant wave heights in the region typically range between 2 m and 3 m, while the wave periods generally range between 8 seconds and 12 seconds (Alveirinho Dias and Nittrouer, 1984; Bio et al., 2020a). During storm events, low-pressure systems generated over the Atlantic Ocean can cross the region with associated S-SW winds. Extremely energetic conditions are produced during these events, with wave heights that can reach between 9 m and 12 m, although waves between 3 m and 6 m are most common during those events (Viitak et al., 2021). During these highly energetic conditions, which normally take place during winter, it was demonstrated that the Portuguese coastline is highly susceptible to severe erosion (Bio et al., 2022).

The Portuguese coast includes a wide range of morphological and sedimentary environments, such as beaches, cliffs, estuaries, lagoons, and barrier islands. Since the 19th century, an overall regressive trend of the coastline has been documented as being caused by human intervention (dams and coastal engineering structures construction, sand extraction, and agricultural practices, Santos et al., 2021). The NW Portuguese coast has been historically defended against the sea since the early 20th century. However, natural and artificial barriers can affect wave patterns, causing localized reversals in the prevalent longitudinal drift, mainly directed to the south.

This erosion has led to dune loss, the transformation of some sandy beaches into gravel beaches, and inefficient transverse structures (Granja and Pinho, 2012). Inside the Portuguese littoral, the NW coast is characterized by cliffs, low rocky coasts, sandy and gravel beaches, and extensive linear beaches backed by coastal dune systems. It was demonstrated that this region was prone to severe erosion rates that reached 8 m/yr (Ponte Lira et al., 2016). However, in the medium-term (last decade), the NW Portuguese shoreline presented a relatively stable behaviour with a trend towards erosion in recent years. This trend, combined with the projected sea level rise, changes in wave climate, and increased frequency and strength of extreme events, suggests that erosion and shoreline retreat could be a severe problem (Bio et al., 2022).

Northern Portugal's coastline sediments originated from igneous and metamorphic rocks in the Precambrian and Palaeozoic eras (Alveirinho Dias and Nittrouer, 1984). The sediment deposits of the region have a predominant marine origin below 50 m height and a fluvial origin between 50 m and 130 m (Araújo, 2000), with a transition character between 37 m and 60 m (Pedrosa, 2013).

3.2.2. Beach profile data

Two different datasets were selected to define the beach profiles that will feed the data models. The COSMO program monitored several beach profiles along the Portuguese coast. Meanwhile, the MarRisk project defined the coast type and the beach sediment sizes at several coastal locations. The combination of both datasets resulted in beach profile data from 12 sites on the NW coast of Portugal. These profiles are distributed along a coastal stretch of approximately 90 km in length between Moledo Beach, at Caminha municipality, and Paramos Beach, at Espinho municipality (Figure 3-2). The beach profiles were selected by considering available data from two monitoring programs, the COSMO program (APA, 2023) and the MarRisk project (Bio et al., 2020b).

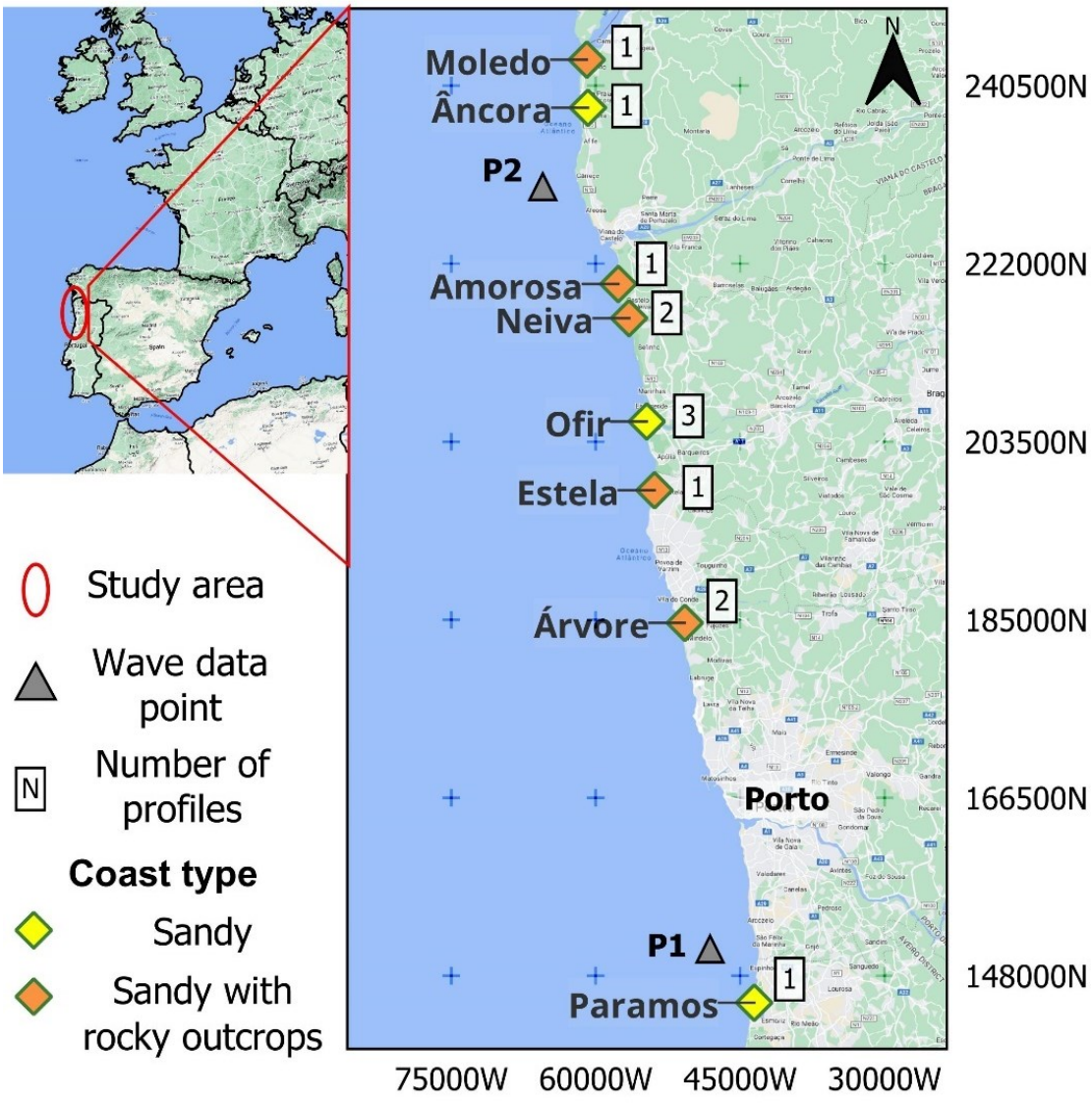


Figure 3-2. Study area and location of the beaches where the profiles were measured. White rectangles indicate the number of profiles measured seasonally at each beach, and the rhombuses indicate the coast type of each beach. Coordinates are referred to the ETRS89 global reference system.

The COSMO project provided beach profiles measured between 2018 and 2021, with observations approximately every three months, resulting in 11 to 12 profiles for each location. Five parameters were selected from these surveys: the elevation referred to the mean sea level (Z_MSL, m), the x-coordinate (m), the y-coordinate (m), the distance to a reference point in the profile (m) and the time since the first observation (days). The coordinates were referenced using the ETRS89 reference system. The reference point was defined considering the profile below 6 m (MSL), corresponding to the profile extension exposed to wave action under average conditions. The horizontal errors in the profile survey varied from 1.1 cm to 2.5 cm, and the vertical errors ranged from 1.8 cm to 4.9 cm. The coordinates provide information about the geographical location of the profiles, allowing us to predict the elevation of points that were not previously measured as long as the coordinates are in the same range as the training coordinates.

Beach profile data is presented in Figure 3-3, highlighting the different morphodynamic patterns of the beaches. Profiles such as the ones measured in Moledo, Âncora, and Árvore P2 revealed stronger morphodynamics than the ones observed in Neiva P2 and Paramos, whose profiles were relatively stable along the monitored period. Also, some profiles are wider than others. For instance, Neiva P2 extends to approximately 70 m, whereas Paramos extends to almost 200 m.

The active profile on some beaches begins precisely at the first point, at 6 m elevation (MSL). In contrast, others reveal an active profile below lower elevations, such as Neiva P2 and Ofir P3. Note that the COSMO program measured two profiles in Neiva and Árvore beaches and three profiles in Ofir beach. In the beaches that presented more than one profile at different locations, these profiles were identified as P1, P2, and P3. Most of the profiles are in an advance erosive state when the complete profile is observed (not shown), presenting erosion cliffs in the frontal dune, except at Moledo, Âncora, and Árvore P1, which presented a beach profile with a frontal dune, berms of variable extension, and a pronounced face.

Data from the MarRisk project were selected to define the coast type, classifying it as sandy or sandy with rocky outcrops, and the beach sediment sizes for each location were obtained from sand samples collected at three different points for each profile (Table 3-1). These locations are identified by the number that follows the sediment size, in which "1" refers to a berm intermediate position, "2" refers to the berm edge, and "3" refers to the beach face. Beach sediment sizes were analyzed either by sieving and digital images (Pinho et al., 2024, in preparation)

It must be stressed that the COSMO program data had two profiles in Neiva and Árvore beaches and three profiles in Ofir, as mentioned before, but the MarRisk program only analysed the sediment size

in a single profile of each one of these two beaches. Hence, the same sediment characteristics were assumed for the two profiles at Neiva beach and the three profiles at Ofir beach. Árvore beach is in the transition between the coast types, resulting in different classifications for each profile on that beach. It was assumed that the sand in these locations had the same D50 and D90, which correspond to the median diameter and the size in which 90% of the sediment mass is smaller, respectively. It was assumed that the sediment data remained constant over time due to the lack of sand samples collected in different seasons. The use of sediment data is relevant to include in the model the influence of the sediments on the beach morphology under a similar wave action.

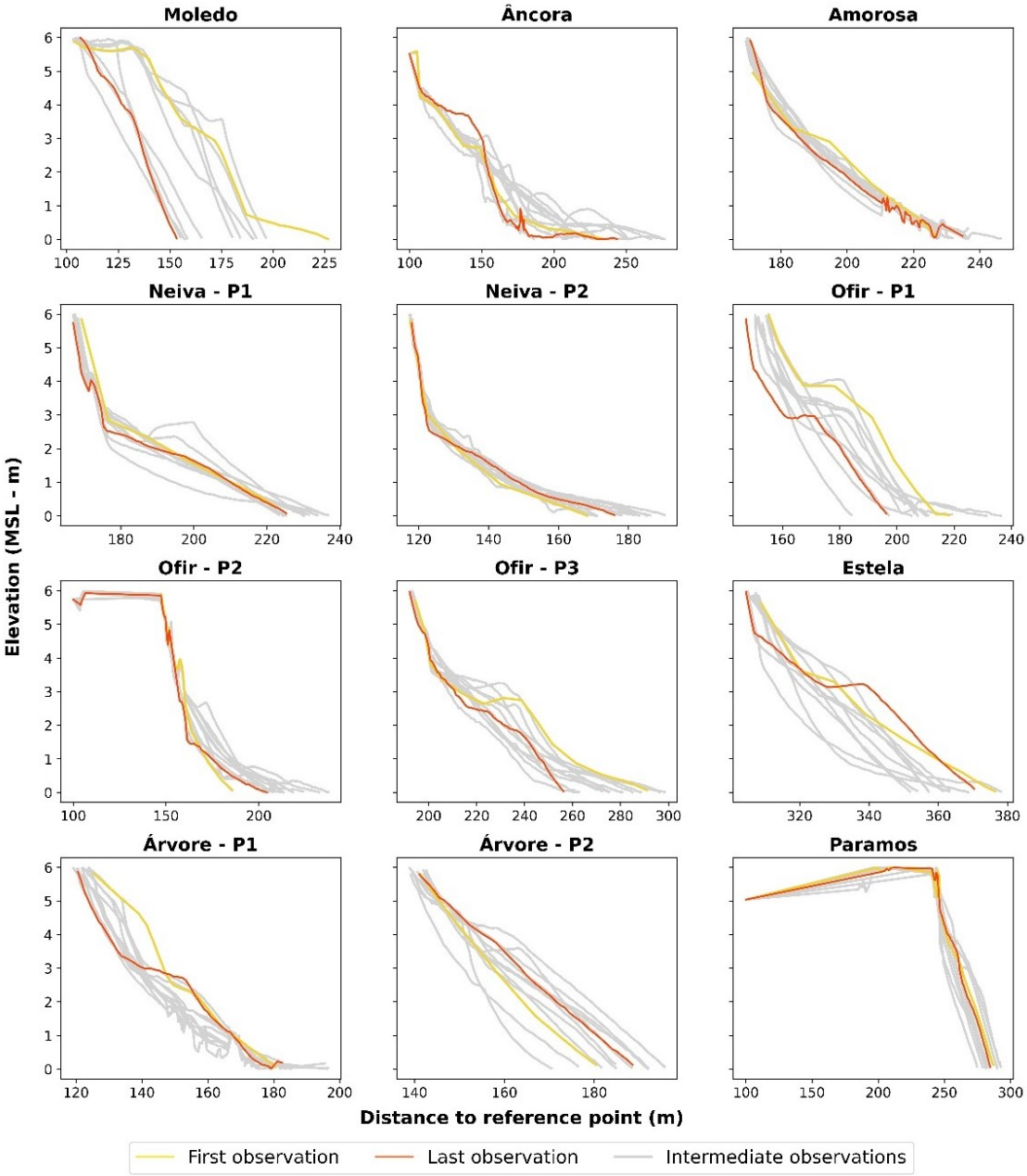


Figure 3-3. COSMO beach profiles. The yellow line indicates the first profile surveyed, and the orange line is the last profile, which was used to assess the model's performance..

Table 3-1. Beach sediments characteristic diameters (D50 and D90). The number following the diameter indicates where the sample was collected (1 – berm intermediate position; 2 – berm edge; 3 – beach face).

Beach	D50_1 (mm)	D50_2 (mm)	D50_3 (mm)	D90_1 (mm)	D90_2 (mm)	D90_3 (mm)	Coast type
Moledo	0.64	0.69	0.77	0.95	0.99	1.68	Sandy with rocky outcrops
Âncora	0.58	0.50	0.50	0.92	0.91	0.92	Sandy
Amorosa	0.37	1.48	0.42	0.61	4.67	2.58	Sandy with rocky outcrops
Neiva P1	0.36	0.29	0.22	0.76	0.49	0.44	Sandy with rocky outcrops
Neiva P2							
Ofir P1	0.35	0.45	0.27	0.49	1.38	0.75	Sandy
Ofir P2							
Ofir P3							
Estela	0.42	0.56	0.79	0.82	1.76	2.76	Sandy with rocky outcrops
Árvore P1	0.64	0.72	0.34	0.98	1.92	0.58	Sandy with rocky outcrops
Árvore P2							Sandy
Paramos	0.54	0.66	0.70	0.94	0.97	1.29	Sandy

The summary of the sediment size data and the profile locations are presented in Figure 3-4. The sandy beaches had the lowest D50 and D90 with a more uniform granulometric distribution. Meanwhile, the sandy beaches with rocky outcrops had higher sediment sizes.

Results of both COSMO program and MarRisk project allowed the identification of four cliff profiles (VP Âncora, Ofir P1, Ofir P2, and Paramos) and eight Dean profiles, according to usual beach profiles classification (Jackson and Short, 2020). These classifications were based on the topographic survey conducted by the COSMO program and by using photogrammetric elevation data, aerial photography, and on-site observations collected in the MarRisk project.

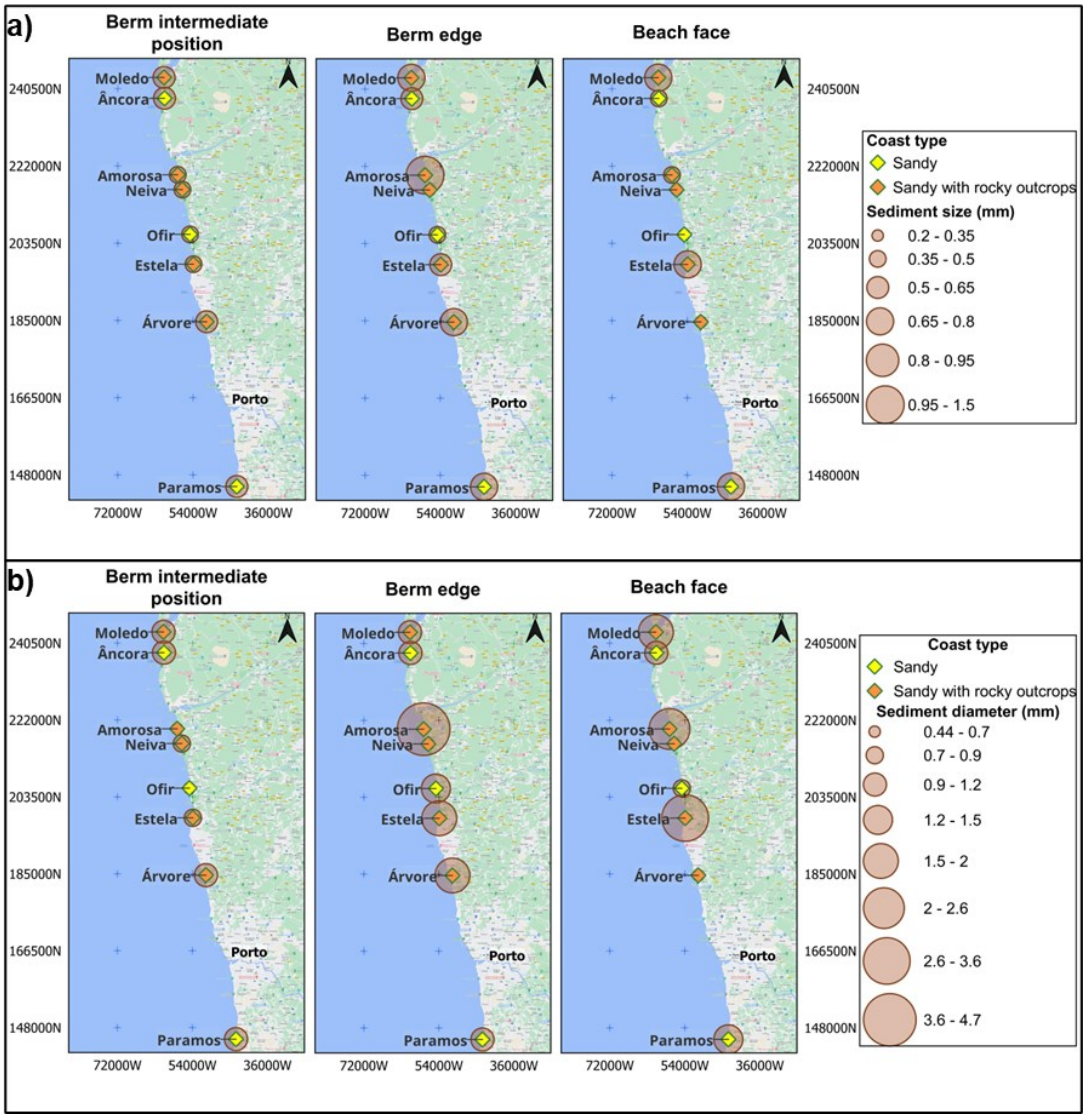


Figure 3-4. Beaches sediment size. The sediment samples were collected at the berm intermediate position (first column), berm edge (second column), and at the beach face (third column). a) D50 results; b) D90 results.

The erosion/accretion trends of the selected beaches were presented using the profile data from the COSMO program (Figure 3-5). The rates were computed by considering sequentially measured profiles, the first performed in one season and the last performed in the following season. Some profiles presented erosional and other accretional rates, but the sum of the rates for each seasonal pair is presented in Figure 3-5. From this data, it can be inferred that accretional rates prevail between winter and spring and from spring to summer; meanwhile, erosion rates increase from autumn to winter and between summer and autumn. The beach accretion on the northern Portuguese coast is more intense for seasons with lower wave energy (winter to spring and spring to summer), reaching values of more than 60% of the occurrences along the monitored period. In contrast, erosion can occur during seasons

with more intense wave energies (summer to autumn and autumn to winter), reaching frequencies between 58% and 69% along the monitored period.

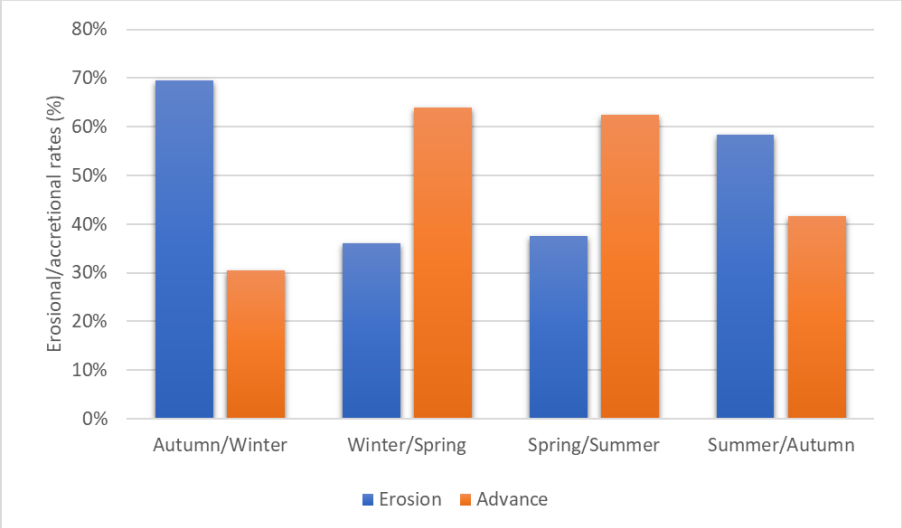


Figure 3-5. Occurrence of erosion and accretion rates in northern Portugal beaches along the seasons, computed from the COSMO beach profiles data.

3.2.3. Beach profile data processing

Beach profile data was divided into two sets. The first set, which covers the period between January 2018 and February 2021, was used to train the data model. The remaining data, from February 2021 to August 2021, was used to evaluate the model's performance. This partition resulted in ten to eleven profiles for each location for the model training and one for the model validation.

The profile variables used in the RF models were the Z_MSL (m), x-coordinate (m), y-coordinate (m), distance to a reference point in the profile (m), and number of days since a reference date (01/01/2018) computed using the field campaign dates. The reference point in each profile was located 100 m before the first point in the profile, which presents a 6 m elevation (MSL). When the profiles have several points with 6 m elevation, only the points below the furthest 6 m elevation point were selected. This elevation value was adopted as the upper limit of the tide and wave action under average conditions (Fontán-Bouzas et al., 2022). This assumption restrained the available profile data to the area affected by tides and waves, reducing the amount of data to be passed to the RF model and filtering it to the profile extension under the influence of the coastal hydrodynamics. Thus, aeolian processes affecting the berm and frontal dunes dynamics will be left out of the analysis, as they are not the focus of this work.

Considering that the model was trained using the data of all the profiles simultaneously, the x and y coordinates are key parameters to provide the location information of the profiles along the study area. They allow the model to simulate the profiles at any given coordinate as long as the values are in the range of the profile coordinates. The distance to a reference point in the profile, which can also be

referred to as the profile length, provided information on the potential magnitude of the wave effect at each profile point.

The mean and standard deviation (std) of the beach profile elevation and length are presented in Table 3-2. It can be observed that the elevation oscillated between 1.47 m and 3.70 m, and the profile's mean length varied between 142 m and 335 m. These metrics indicate that the profiles did not have any similar pattern because they were chosen according to the availability of data.

Table 3-2. Mean and std of beaches elevation and profile length.

Profile	Mean elevation (m)	Elevation std (m)	Mean profile length (m)	Profile length std (m)
Moledo	3.7	1.9	142.2	24.5
Âncora	1.7	1.5	173.1	42.0
Amorosa	2.1	1.6	202.5	18.9
Neiva P1	1.9	1.4	197.0	18.6
Neiva P2	1.5	1.3	146.6	18.4
Ofir P1	2.3	1.6	182.9	19.9
Ofir P2	2.0	1.9	175.9	27.2
Ofir P3	2.1	1.6	235.5	26.6
Estela	2.4	1.7	334.6	17.7
Árvore P1	2.2	1.7	152.4	17.8
Árvore P2	2.8	1.6	163.5	14.2
Paramos	3.3	1.9	252.7	30.2

3.2.4. Wave data

Waves are the most relevant driver of morphological changes for the chosen profile extension in the study area. The wave data was included in the data models to provide them with information related to the hydrodynamics of the study area when trying to forecast the profile evolution. Wave data for this study was extracted from the Atlantic -Iberian Biscay Irish (IBI) Ocean Wave Reanalysis product managed by the E.U. Copernicus Marine Service Information (Copernicus Marine Service, 2022). Reanalysis datasets perform hindcast simulations that combine historical observations with climate models to generate consistent time series data. The temporal resolution of this database is hourly, containing records between January 1993 and December 2021 with a spatial resolution of 0.027° x 0.027°. The significant wave height (HM0), wave period (WP), and wave direction (WD) variables were obtained for the period between 01/01/2018 and 31/12/2021, in agreement with the beach profile data. The wave data from the IBI product was extracted at two points, the first located near Paramos Beach (P1) and the other located between Âncora and Amorosa beaches (P2, Figure 3-2). Figure 3-6. represents the evolution of wave variables for the selected study period.

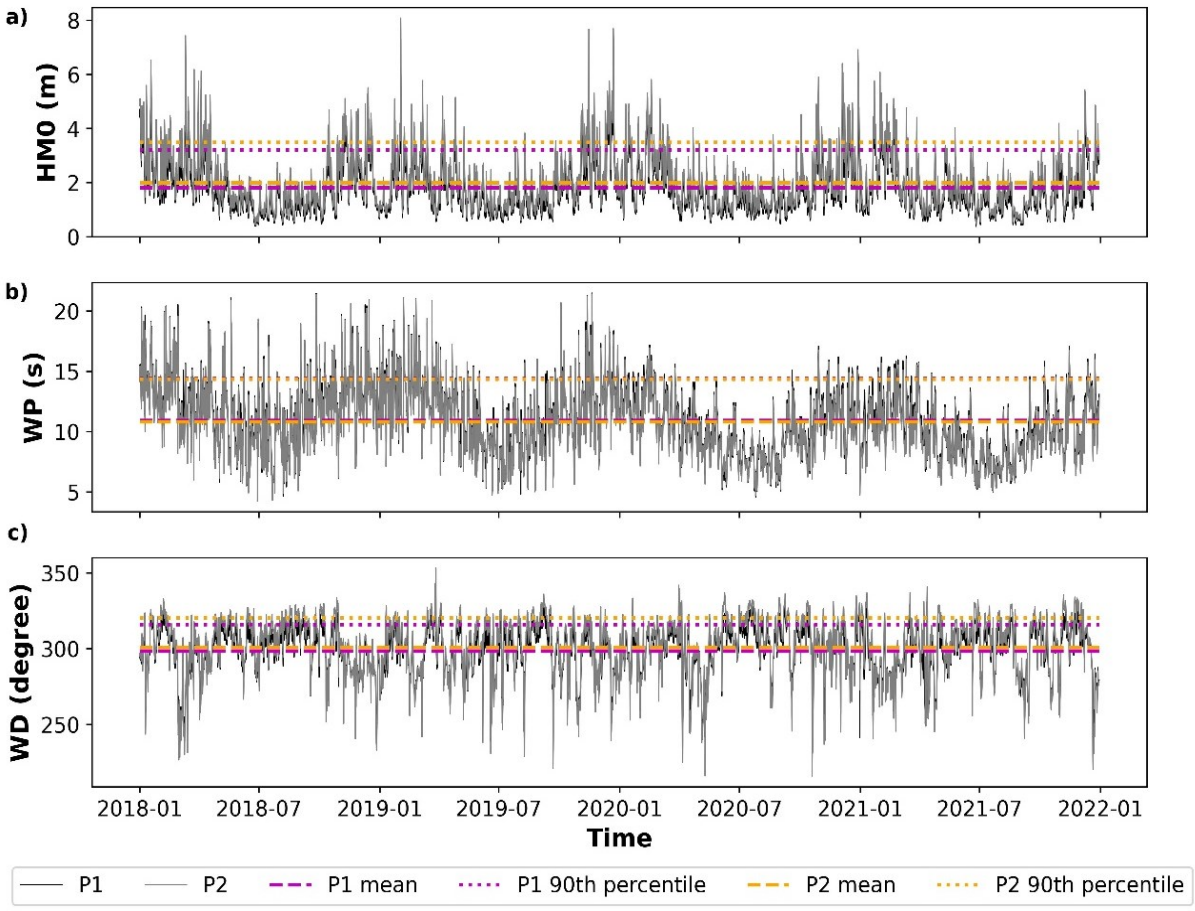


Figure 3-6. IBI wave reanalysis data for the study period. a) Significant wave height; b) Wave period; c) Wave direction.

The wave data at both points present slight differences for HM0 but are identical for WP and WD for the analysed period. For WP and WD parameters, P1 and P2 lines overlap, and the mean and 90th percentile values are almost the same. The mean and the std values of the time series are presented in Table 3-3. The mean HM0 is 27 cm higher at P2, but the mean WP and WD have similar values. During the observation period, more energetic events during winter periods can be identified. The HM0 reached almost 8 m during the winters of 2019 and 2020, and the WP exceeded 20 s. Despite these values, it should be noted that there were no extreme storms during this period compared to historical data (Vieira et al., 2020).

It must be highlighted that the grid resolution of the IBI product is not fine enough to get accurate details for each one of the considered beach profiles. Therefore, P1 and P2 points were defined to provide wave data near the northern and southern limits of the study area (Figure 3-2). The model learns how to process these inputs during the training, attributing the relevance of each considered input.

Table 3-3. Mean value and std of the wave variables provided by the IBI reanalysis product at points P1 and P2.

	P1		P2	
Variable	**Mean**	**Std**	**Mean**	**Std**
HM0 (m)	1.81	0.97	2.08	1.06
WP (s)	10.9	2.7	10.7	2.7
WD (degree)	298.4	15.7	302.2	19.3

The distance between the beach profiles to P1 and P2 may indicate that some locations are more exposed to the influence of one of these points. The relevance of wave characteristics at P1 and P2 for each beach profile is determined during the training of the models. The parameters that associate the wave conditions at P1 with the output will present higher values when predicting a profile closer to P1, while the wave conditions at P2 will have lower weights when simulating a profile closer to P1. Hence, the associations that the model creates between the inputs and outputs are calibrated during the training phase.

Although the choice of only two data points may increase the uncertainties of the model, adding more wave data points would not significantly provide more information to the model because, as it was demonstrated in Figure 3-6 and Table 3-3, the selected wave parameters present similar values and behaviour for both locations. Only two points were chosen to simplify the model and decrease the number of input variables and the computational cost. However, areas with higher wave data variability would need more points, and the adequate number of points for each study should be carefully selected to provide reliable data to the model and thus forecast the morphodynamic variables accurately.

The resulting wave data time series were then combined with the profile and the sediment size data. Particularly, the data model was fed with the 90th percentile of HM0 and WP and the mean WD, which were computed considering the time intervals between each survey. Hence, the wave variables summarized the hydrodynamic conditions during the interval between observations.

The 90th percentile was adopted because wave energy is one of the most important parameters for sediment transport and beach morphodynamics, which is directly associated with wave height and period (Jackson and Short, 2020). Therefore, if the mean HM0 value was considered instead of the 90th percentile, the effect of higher waves could be underestimated.

The input variables should have an adequate frequency according to the morphological output variable timescale to be predicted. For daily morphological changes, the water level variation is important for beach profile dynamics, especially for analysis of dynamics of morphological patterns that are present on the beach face and that are linked to cycles of spring and neap tides (Gomes et al., 2018; Jackson

and Short, 2020). This is also the case for extreme events that could have a significant impact depending on the storm surge water level. However, for the seasonal timescale considered in this study, the tide effects were assumed to be less important than the wave action variables that were considered, with average values for the forecast period. This way, the model's simplicity was privileged since considering water levels will increase the model complexity.

Besides, the ocean water level variation due to tide and storm surge was not considered because these variables should present similar values along the study area. The seasonal frequency of the profile measurements did not capture tidal influence on the beach profile dynamics. In addition, profiles were not performed just before and after a storm event where its associated storm surge can have a strong impact. Hence, only wave action data were considered as the main drivers for the seasonal morphodynamic processes under simulation in this study. It is important to notice that tide and storm surge variability are key variables when modelling short-term morphological changes, considering hourly to daily timescales.

3.2.5. Random forest algorithm

The RF algorithm is an ensemble learning method that combines the results of multiple base models to reduce the results error. RF has the advantage of delivering satisfactory results even with few and no normalized data. This is an advantage for the data used in this work, which come from multiple sources. Moreover, RF uses a Decision Tree (DT) as a base model, an algorithm suitable for classification and regression problems. DTs were vastly used for different classification problems, for instance, to identify the category of ships according to their radar profile and to recognise numerical digits (Breiman et al., 2017). For the water resources scientific area, DTs were applied to estimate the flooding probability (Chen et al., 2008) and to predict the wave runup on vertical piles (Kazeminezhad & Etemad-Shahidi, 2015).

However, DTs models are easily hyper-fitted, which means they can present an excellent performance in the training dataset but have inferior performance when using different data. RFs can overcome this problem by combining the results of many DTs, which reduces the individual bias of each model and, consequently, the overfitting probability. Despite that, using RFs models does not guarantee that the result of the ensemble will be better than the result of the best individual model. However, it was demonstrated that model ensembles present smaller bias and variance than the individual solutions, thus improving the final prediction's accuracy, reliability, and consistency (Iglesias et al., 2022). Thus, the use of RFs models will decrease the likelihood of choosing a model with deficient performance (Cutler et al., 2012).

The RF models were implemented in Python version 3.10.9 using the open-source ML library Scikit-learn version 1.2.1 (Pedregosa et al., 2011) and the *RandomForestRegressor* function. The hyperparameter values were defined in an iterative process, which consisted of training the model and evaluating its performance. The model training consists of using the *fit()* function, which automatically adjusts the weights of the model to minimize the error between observed data and model output. During this process, the relevance of each input is determined, as well as the relevance of each DT to the output.

The models were trained and validated using the k-fold cross-validation method, which divided the dataset into five consecutive groups. This method uses one fold to evaluate the model performance and k-1 folds to train it. The process involves repeatedly refining the hyperparameters of the model and evaluating the results until a suitable level of performance is achieved. The data before January 21st, 2021, were used to train the model, and the remaining data were used for testing, resulting in ten to eleven profiles in the training dataset and one in the test dataset. The number of profiles considered depended on the location. At Estela, Moledo, and Paramos beaches, eleven profiles were considered in total, whereas twelve were considered for the other locations. The performance of the model was measured by the root mean squared error (RMSE), the mean absolute error (MAE), the coefficient of determination (R2), and relative error (RE) indicated by equations (**3-1**, (3-2, (**3-3** and (**3-4** respectively:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3-1}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{3-2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3-3}$$

$$\text{RE} = \frac{| y_i - \hat{y}_i |}{y_i} \tag{3-4}$$

where n is the number of observations, y_i is the observed value, y^_i is the predicted value and y ¯ is the mean of the observed values. The maximum error was also determined for each validation fold.

Additionally, two different models were implemented with the same specifications but different input variables. The M1 model was fed with the beach profiles, the wave conditions, and the sediment size data. The M2 model considered only the waves and the profile data to assess the impact of including sediment size data on the profile evolution forecast. Figure 3-1 presents the architecture of the RF model and the chosen hyperparameters.

The data used for training and testing were organised to enable the models to predict the profile at t+1 using the elevation of the current profile at t. The input variables included the current profile's x-y coordinates, elevation, number of days until the next campaign, HM0 and wave period 90th percentiles, mean wave direction (M2), and sediment characteristic diameters (D50 and D90 -M1). The statistics of the wave data were calculated between t and t+1.

The final values of the RF hyperparameters were 200 estimators, the maximum depth of the DTs was 80, the minimum sample split was 5, and the minimum sample leaf was 4. The function to measure the quality of the splits was the Friedman mean squared error, and the maximum number of features to consider during each split used the sqrt criteria. These values were determined iteratively by observing how each hyperparameter affects the model's error. The models were trained, and a performance metric baseline was determined. After that, the value of a single hyperparameter was changed, maintaining the values of the others, and its impact on the error was determined. When the minimal error for that configuration was determined, a second hyperparameter was assessed, following the same steps. The other hyperparameters used the default value. After training, the models' performance was evaluated using previously unseen data relative to the period after February 2021.

Although the sediment values were assumed to be constant at each location, the sediment data helped the model to differentiate each profile, considering that it was trained using all beach profiles simultaneously. This way, during the training phase, the model can learn how to associate the sediment size with each location and how the size values affect the morphological changes locally. More details on how the decision trees consider each input can be found in the literature (Cutler et al., 2012; Irizarry, 2019; Kubat, 2017). Ideally, a time series of the evolution of the beach profile granulometry at each location could also be used to achieve higher accuracy in modeling the morphological evolution of the cross-shore sections. Still, a simultaneous collection of sediment samples with the beach profile measurements is required.

3.2.6. Computation of beach slope and retreat/ advance

After predicting the beach profiles with the data models, the resulting beach slope was computed to assess the differences between the observed and simulated slopes. The slope was determined considering the distance between the points with elevations of 0 m and 3 m (MSL). Moreover, these elevations were used as a reference to determine the beach width in each profile. This interval was selected according to the wave height values in the region, focusing on the elevations where the wave action is most relevant to the morphological change. The choice of reference heights directly influences the slope error. Nevertheless, these values are adequate for the profiles studied.

The profile lengths were divided into 5 m segments. At each segment, the slope was computed considering the elevation of each stretch's initial and final points. The elevation values were interpolated using the *interp1d* function from the SciPy library (Virtanen et al., 2020). Then, the mean slope of each profile was computed considering the slope of each segment.

The retreat/advance of the beach profile during the last monitoring period was computed considering four elevation references (1 m, 2 m, 3 m, and 4 m, MSL). As the profile elevation results were not necessarily equal to the adopted reference values, these elevations were added to the results of the models, and the correspondent distance was linearly interpolated between the closest values around the reference elevation. With the interpolated distance values, the advance/retreat was determined by computing the difference between the distances of the most recent and the previous profile at each elevation. For instance, if the distance difference at 1m height is positive, it indicates accretion, whereas negative values represent erosion.

3.2.7. Storm conditions simulation

The model's capacity to forecast the storm's effects was also qualitatively evaluated. Two different scenarios were considered. The first scenario adopted an HM0 of 4 m and a WP of 15 s. It is the same scenario that Vieira et al. (2020) defined to study the performance of coastal protection structures on the Portuguese coast. The second scenario increased the HM0 to 8 m, maintaining the same value for the WP. The WD was the same for both scenarios, ranging between 286 and 305, according to the profile location.

The main objective of the proposed simulations is to produce insights about the model response using HM0 values in the range of the training data (~4 m) and additionally demonstrate the capacity of the model in predicting the morphological evolution considering wave conditions outside the training data range (the maximum HM0 was slightly under 8 m). Considering that RF is a multi-dimensional space, it is necessary to use additional techniques to inspect if they are considering each input variable adequately. Through this simulation, it is possible to evaluate if the model behaves as expected during a simulation of a storm event that was not covered in the training data, even if the available data is not adequate for short-term extreme events simulations, as explained before in section 3.2.444.

3.3. Results

In this section, the training performance of the models is presented, followed by an assessment of their test results. After, the results of the data models' computation of the beach slope, their retreat/advance, and simulation of a storm event are discussed.

3.3.1. Training performance

The results of the model training using the k-fold cross-validation method are presented in Table 3-4. M1 performed satisfactorily across all metrics, with practically equal R2 values in the five considered folds. The RMSE and MAE also indicate that the model results had low errors. However, the maximum error range was from 0.96 m to 1.65 m, with a 1.33 m mean, which reveals that the implemented model may not be suitable for all the considered locations used to train the model.

Table 3-4. Training results of the M1 and M2 models.

		Fold					
Model	**Metric**	**1**	**2**	**3**	**4**	**5**	**Mean**
M1	**R^2**	0.985	0.987	0.986	0.986	0.985	0.986
	Max error (m)	1.19	0.96	1.52	1.65	1.33	1.33
	RMSE (m)	0.04	0.04	0.04	0.04	0.04	0.04
	MAE (m)	0.15	0.15	0.148	0.14	0.15	0.15
M2	**R^2**	0.986	0.987	0.987	0.986	0.986	0.986
	Max error (m)	1.42	1.34	1.79	1.67	1.39	1.52
	RMSE (m)	0.04	0.04	0.04	0.04	0.04	0.04
	MAE (m)	0.14	0.14	0.14	0.14	0.14	0.14

M2 followed the M1 trend in RMSE, MAE, and R2. Nonetheless, M2 maximum error was higher than M1, ranging from 1.34 m to 1.79 m, with a mean of 1.52 m. Hence, M1 performed better due to the lower maximum error. However, the other metrics did not present significant differences, which allowed us to conclude that the maximum error was associated with a few points in the profile data, as will be further discussed in the next section.

Both models performed highly, with similar R2, RMSE, and MAE in all folds. However, the M1 maximum error was lower than that obtained with M2, with a mean difference of 20 cm. This result indicates that the sediment size data was useful in reducing the maximum error of the model in the training dataset. However, it did not lead to significant improvements in other metrics. Additionally, the models' performance demonstrated their accuracy regardless of the fold used for evaluation. The models have good generalising capacity and are not biased; consequently, they are reliable for forecasting using previously unseen data.

3.3.2. Prediction performance

The model performance to forecast the profiles of the last field campaign is presented in Figure 3-7. The best performances occurred in Amorosa and Neiva P1 beaches, with RMSE varying from 0.18 m to 0.26 m for both considered models, M1 and M2. The worst performance occurred in Moledo, with a RMSE higher than 1.35 m. Additionally, M1 had a better performance than M2 in eight of the 12 locations considered, namely at Âncora, Amorosa, Neiva P1 and P2, Ofir P3, Estela, Árvore P2 and Paramos, but M1 underperformed M2 in Moledo, Ofir P1, and Árvore P1. Both models presented equivalent results for the Ofir P2 profile, and the most significant improvement occurred in Neiva P2, in which M1 RMSE was 0.17 m and M2 RMSE was 0.49 m. The average M1 RMSE for all the considered profiles was 0.51 m, and the average M2 RMSE was 0.55 m.

Additionally, both models achieved good results despite the temporal variability of the profiles. It highlights that the algorithm correctly reproduced the results in dynamic beaches, and not only in beaches with more stable trends. It can be observed that the M1 forecast at Amorosa, Neiva P1, and Neiva P2 overlapped with the observed data, having slight deviations. The results in Neiva P1 must also be emphasised due to higher variability in the historical data, demonstrating that the model prediction capacity is not limited to less dynamic areas.

In the case of the Moledo profile, it must be noted that it is located very close to the Minho River estuary mouth, an important sediment source to the coastal platform that was not considered as input data. The narrow mouth of this estuary increases the current velocity, intensifying sediment transport from the estuary to the ocean (Melo et al., 2020). Therefore, the river flow may be a relevant variable that was not included in this study, and the AI models did not have enough information to forecast the profile evolution accurately in this location. Quantifying the contribution of fluvial inflows in the area may improve the performance in the locations close to estuary outlets.

Moreover, there is no clear pattern between the beaches in which the models performed well or not. The RMSE achieved in sandy beaches (Âncora, Ofir, and Paramos) ranged from 0.37 m to 0.73 m, whereas in sandy beaches with rocky outcrops, the RMSE varied from 0.17 m to 0.63 m (excluding Moledo). Hence, the lowest errors were obtained for sandy beaches, but the highest errors were not significantly different. In addition, proximity to the wave data points, geographical location, and sediment size did not affect the model's performance. These influences can be rejected when observing the performance of the models in simulating profiles at the same beaches (Neiva, Ofir, and Árvore), in which error had considerable local variability.

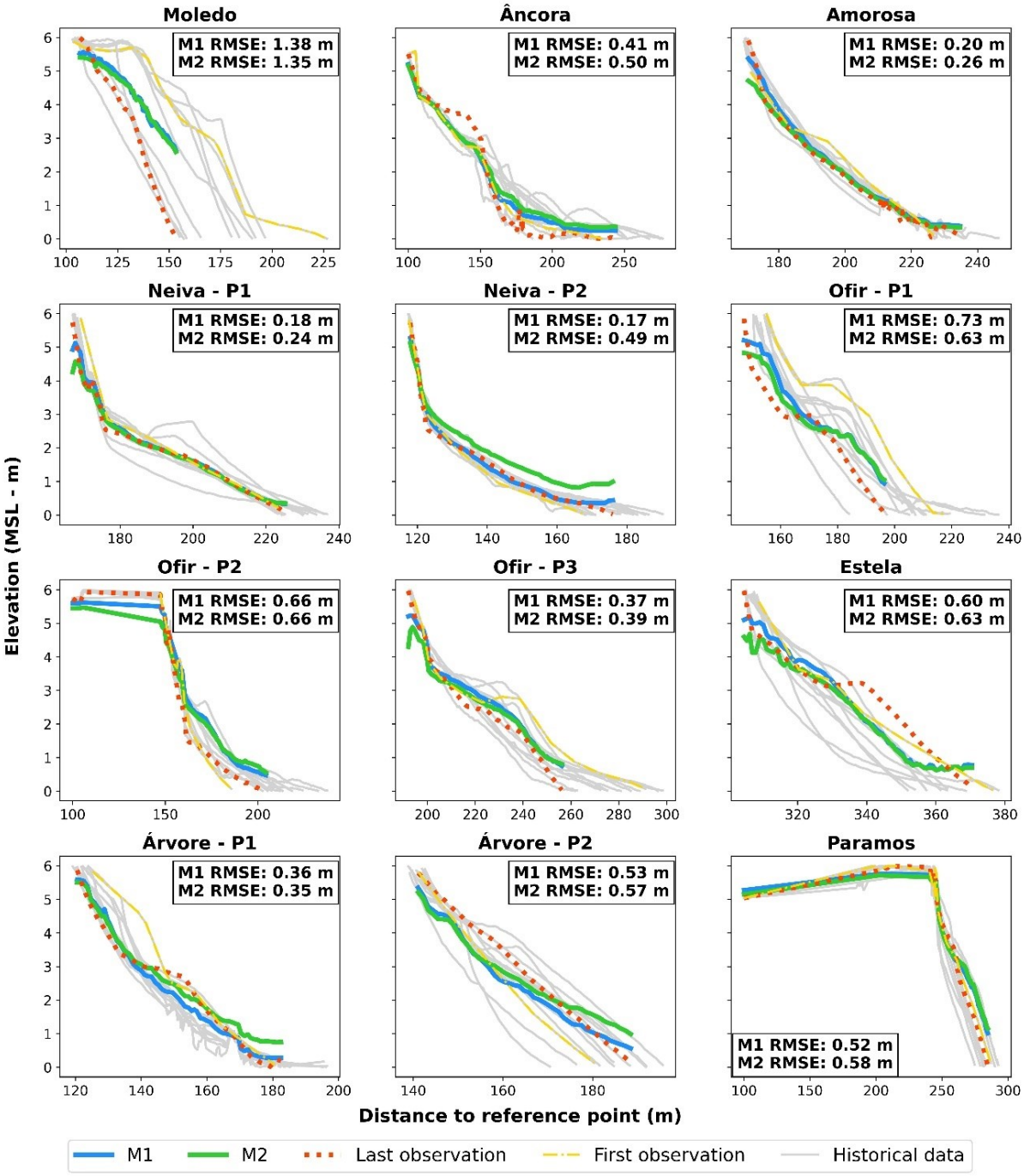


Figure 3-7. Comparison between the results predicted by M1 (blue) and M2 (green) models and the observed values (red dots). The historical data used to train the models are represented in yellow (first observation) and grey.

Despite that, it is possible to use explainable AI (XAI) frameworks that help to interpret the relationship between inputs and outputs. This concept aims to make the models more intelligible to humans (Gunning et al., 2019). Adopting such frameworks could enhance the reliability of AI models, allowing physical interpretations of the results and increasing their acceptance by domain experts. For instance, it would be possible to identify how the model associates each profile to each wave data point

and which variables are the most relevant to the output. However, this study is not focused on that component of AI methodologies.

Furthermore, it was identified that the coast stretch type slightly influenced the M1 model error and had an effect on the averaged RMSE. The lowest errors were obtained in profiles measured at sandy beaches with rocky outcrops (Neiva P1, Neiva P2, and Amorosa). The average RMSE from the beach profiles located in sandy stretches with rocky outcrops was 0.48 m, whilst the average RMSE in the profiles in sandy stretches was 0.54 m. Regarding M2 results, an average RMSE of 0.55 m was obtained for sandy stretches with rocky outcrops and 0.56 m for sandy stretches. This result also highlights that the consideration of sediment size data on the models improved their performance. The accuracy of M1 increased mainly at the beginning of the profiles, as observed in the results of Amorosa, Ofir P1, Ofir P2, and Estela.

The capacity of the data models to forecast the evolution of beach profiles was compared with results obtained in previous studies. Using a diffusion model, Karunarathna et al. (2012) achieved an MAE range between 0.1 m and 0.5 m to predict profile changes, and Otto et al. (2021) used a spatiotemporal-functional approach to forecast the beach profile evolution, achieving RMSE ranging between 0.25 m and 0.9 m. The average MAE in the test dataset obtained in this study was 0.45 m for M1 and 0.49 m for M2 and, excluding Moledo, which is clearly an outlier, the MAE range is between 0.12 m and 0.62 m for M1 and between 0.14 m and 0.63 m for M2 (Table 3-5). So, the obtained values are in the same range as those obtained by previous authors (Karunarathna et al., 2012; Otto et al., 2021).

The averaged RMSE for M1 was 0.54 m and for M2 0.58 m, and, excluding Moledo again, the RMSE range for M1 was from 0.17 m to 0.66 m and for M2 from 0.26 m to 0.66 m. Lee et al. (2022) implemented a numerical model to simulate long-term changes in beach profiles, achieving an RMSE of 3.7 m, a clearly higher value than the ones obtained in this study. It can be concluded that the errors reached in this study are similar or even lower to those obtained by other authors using different techniques, reinforcing the utility of the data models implemented here.

Table 3-5. MAE and RMSE in the test dataset.

	MAE (m)		RMSE (m)	
Profile	**M1**	**M2**	**M1**	**M2**
Moledo	1.11	1.08	1.38	1.35
Âncora	0.35	0.45	0.41	0.50
Amorosa	0.16	0.14	0.20	0.26
Neiva P1	0.11	0.13	0.18	0.24
Neiva P2	0.12	0.54	0.17	0.49
Ofir P1	0.62	0.55	0.73	0.63
Ofir P2	0.58	0.63	0.66	0.66

Ofir P3	0.31	0.32	0.37	0.39
Estela	0.50	0.49	0.60	0.63
Árvore P1	0.30	0.29	0.36	0.35
Árvore P2	0.46	0.51	0.53	0.57
Paramos	0.42	0.48	0.52	0.58
Average MAE	0.46	0.50	0.55	0.59

To further investigate the models' capacity to predict the observed data, the error distribution was assessed by computing the difference between the observed data and the models' results (Figure 3-8). The error distribution plot is computed by determining the kernel density estimate plot to visualise the data distribution. This figure was done using the *kdeplot* function available in the Seaborn library (Waskom, 2021), which produces a continuous density estimation of the analysed variable using a Gaussian kernel. This methodology is suitable for comparing the models' errors. Density plots exhibit how frequent the values are in the analysed dataset.

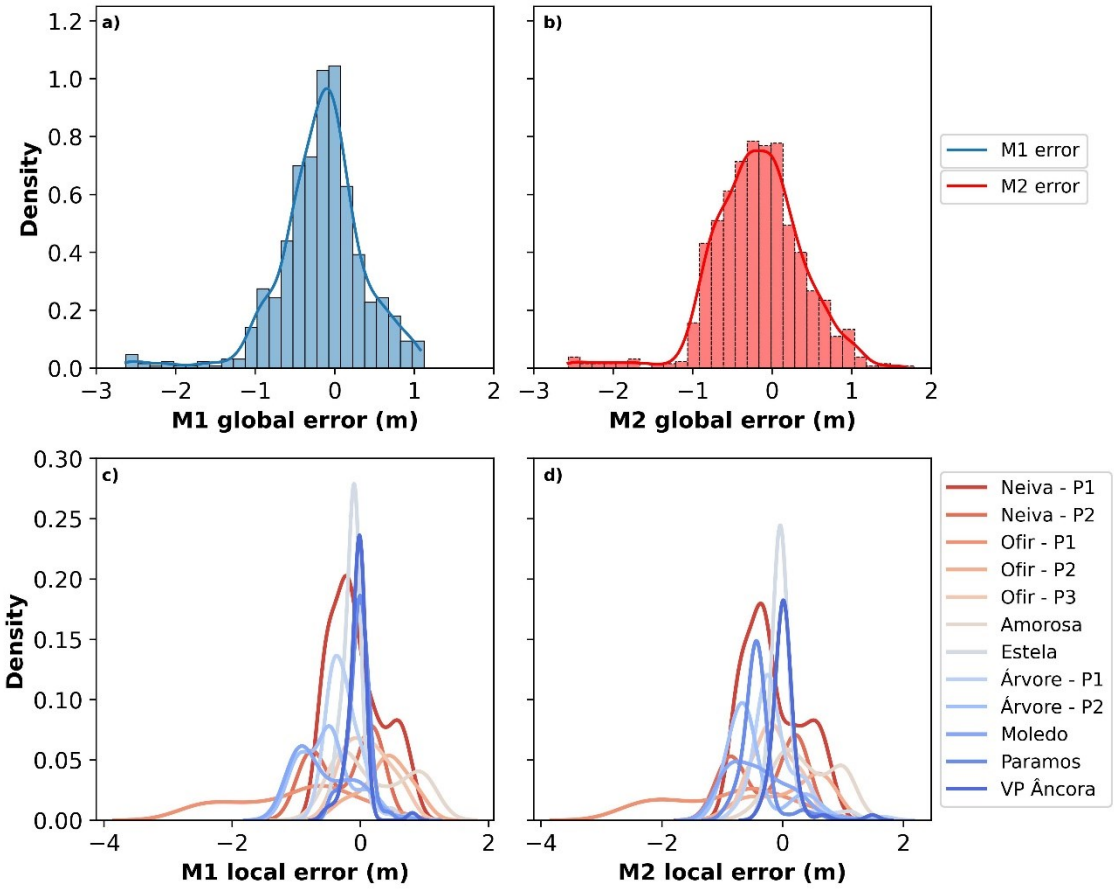


Figure 3-8. Error distribution of the obtained results. a) M1 global error distribution; b) M2 global error distribution; c) M1 local error distribution; d) M2 local error distribution.

These results were analysed by considering a global performance (all locations) and a local performance (assuming only the results at each location). Both models had a symmetric global error

distribution centred near zero, indicating that most errors were near zero. However, the M1 density plot distribution indicates that M1 results had lower variance when compared to M2.

The values in the extremities of the plot indicate the highest and lowest error values obtained. Both models' maximum overestimations (negative values) were close to -2.7 m, and underestimations were close to 1 m for M1 and 2 m for M2. However, the low density of errors indicates that these values are uncommon. The global errors considering a 5% confidence interval were computed, removing the values in the extremities with frequency lower than 2.5%.

When analysing the local error distribution, a similar behaviour between both models can be identified, with a narrower tendency in the M1 results and more dispersion for M2. The highest error occurred in Ofir P2, which appeared in the profile's first 50 m (Figure 3-7). Furthermore, both models had a normal error distribution, with the majority of values between -1 m and 1 m, and a mean of approximately 0 m. Both models best result was in Estela, in which the error distribution is a peaked curve centred near 0, indicating a good agreement between models and data.

Therefore, both models performed satisfactorily, achieving acceptable errors and low variance. The inclusion of sediment size data produced better results, indicating that it was an important variable for beach profile forecasting with data models. Moreover, the sediment size data helped to reduce the error in most beaches, regardless of the sand beach granulometric distribution.

The higher error observed in the forecast of sandy beach profiles can be explained due to their more dynamic behaviour. The rocky outcrops increase the stability of the coastal stretch, being more resilient to meteo-oceanographic conditions than the sandy stretches. The beach profile surveys were, perhaps, not long enough to capture significant modifications in these areas, whereas sandy beaches can experience significant changes in shorter periods. This could explain why M1 results in stretches with rocks were better. The model could learn its behaviour with the available data as the pattern changes slower. At the same time, a longer time series may be necessary to improve the model performance in more dynamic areas. However, the model learned the profile evolution of less stable beaches, as can be observed for Neiva P1 results.

3.3.3. Beach slope and retreat comparison

Beach slopes were computed to assess whether the data models could reproduce this profile feature. The slope and width results are depicted in Figure 3-9. As expected, the models' performance aligns with the beach profile's forecast performance. The best results were obtained in Amorosa, Neiva P1, and Neiva P2, while the worst was in Moledo. It must be noted that the stretches shown in this figure correspond to the beach faces of the most recent profiles presented in Figure 3-7. For instance, in the

Estela profile, the beach face corresponds to the stretch in which the models had lower performance when compared to the beach berm. These results clearly influence the slope error at this location. Despite that, the width error of both models was only 10 m.

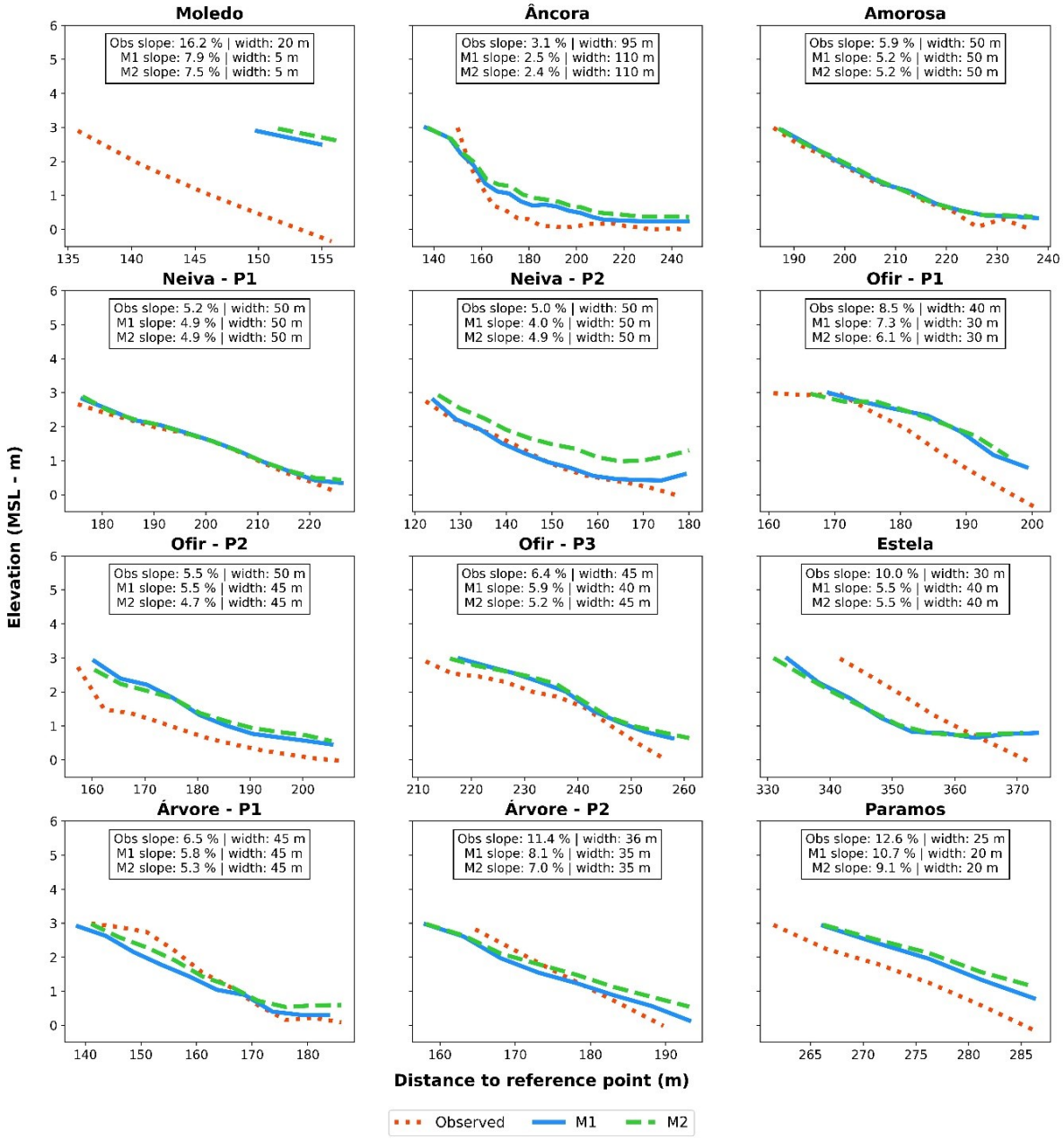


Figure 3-9. Beach slope computation results. The red dotted line represents the observed data, whilst the blue and green lines represent the results of M1 and M2, respectively. The slope results are presented in the white boxes at each subplot.

The absolute error was computed to evaluate the slope error, while the relative error was utilized to assess the width error when comparing the models' performance in relation to beach slope and width (The beach width errors exhibited similar behaviour. Both models achieved perfect results in Amorosa, Neiva P1, Neiva P2, and Árvore P1. The mean relative errors for M1 and M2 are 19% and 18%,

respectively. The width results for both models were identical for all profiles, except at Ofir P3, where M2 outperformed M1.

Table 3-6). Both models results were acceptable in reproducing the beach slope, with an absolute error lower than 2.4%. These results also showed that M1 overcame M2. Seven beaches had a slope error lower than 1% for M1 results, while M2 had five beaches within this error range. Additionally, M2 outperformed M1 only at Amorosa.

The beach width errors exhibited similar behaviour. Both models achieved perfect results in Amorosa, Neiva P1, Neiva P2, and Árvore P1. The mean relative errors for M1 and M2 are 19% and 18%, respectively. The width results for both models were identical for all profiles, except at Ofir P3, where M2 outperformed M1.

Table 3-6. Beach slope absolute error results.

	Slope				
Beach	**Observed**	**M1**	**M2**	**M1 absolute error**	**M2 absolute error**
Moledo	16.41%	8.82%	8.75%	7.59%	7.66%
Âncora	3.17%	2.32%	2.25%	0.86%	0.92%
Amorosa	5.92%	5.72%	5.50%	0.20%	0.42%
Neiva P1	5.16%	5.55%	5.37%	0.39%	0.20%
Neiva P2	5.07%	5.13%	4.29%	0.06%	0.78%
Ofir P1	8.18%	7.71%	6.57%	0.46%	1.61%
Ofir P2	5.79%	6.87%	6.05%	1.09%	0.26%
Ofir P3	6.38%	5.65%	5.18%	0.73%	1.20%
Estela	9.98%	4.18%	4.23%	5.79%	5.75%
Árvore P1	6.75%	5.97%	5.37%	0.78%	1.39%
Árvore P2	11.30%	7.22%	6.28%	4.08%	5.02%
Paramos	12.50%	9.32%	8.91%	3.18%	3.60%
MAE				2.10%	2.40%

The slope predictions with errors lower than 1% presented the same pattern as the local RMSE. The best performances considering both metrics were obtained in Amorosa, Neiva P1, and Neiva P2, indicating that the results of the data models can help in the computation of relevant beach variables, such as the slope. It is important to note that all beach slopes were estimated considering the slope between 0 m and 3 m elevations, which may not be adequate for all considered beaches. However, interpolating the elevation every 5 m along the profile length and calculating the mean slope smooths the influence of the chosen elevation interval. This simplifies the analysis of the profile slope and focuses on the general performance of the model. When applying this methodology to a specific location, the best

elevations should be carefully chosen, ensuring that the slope between the highest and lowest points is linear.

Finally, the profiles retreat/advance considering different reference elevations also demonstrated that M1 performs better than M2 (Figure 3-10. Profile advance/retreat. a) 1 m height; b) 2 m height; c) 3 m height; d) 4 m height.Figure 3-10). Positive values represent an advance toward the ocean, while negative values represent a retreat. The best results occurred in Neiva P1, although the beach advance magnitudes at most profiles differ significantly. From a qualitative point of view (advance or retreat), both models correctly forecasted the behaviour of nine profiles at 2 m elevation and seven profiles at 3 m and 4 m height. The missing bars at some profiles at 1 m elevation indicate that the models did not have values at that height.

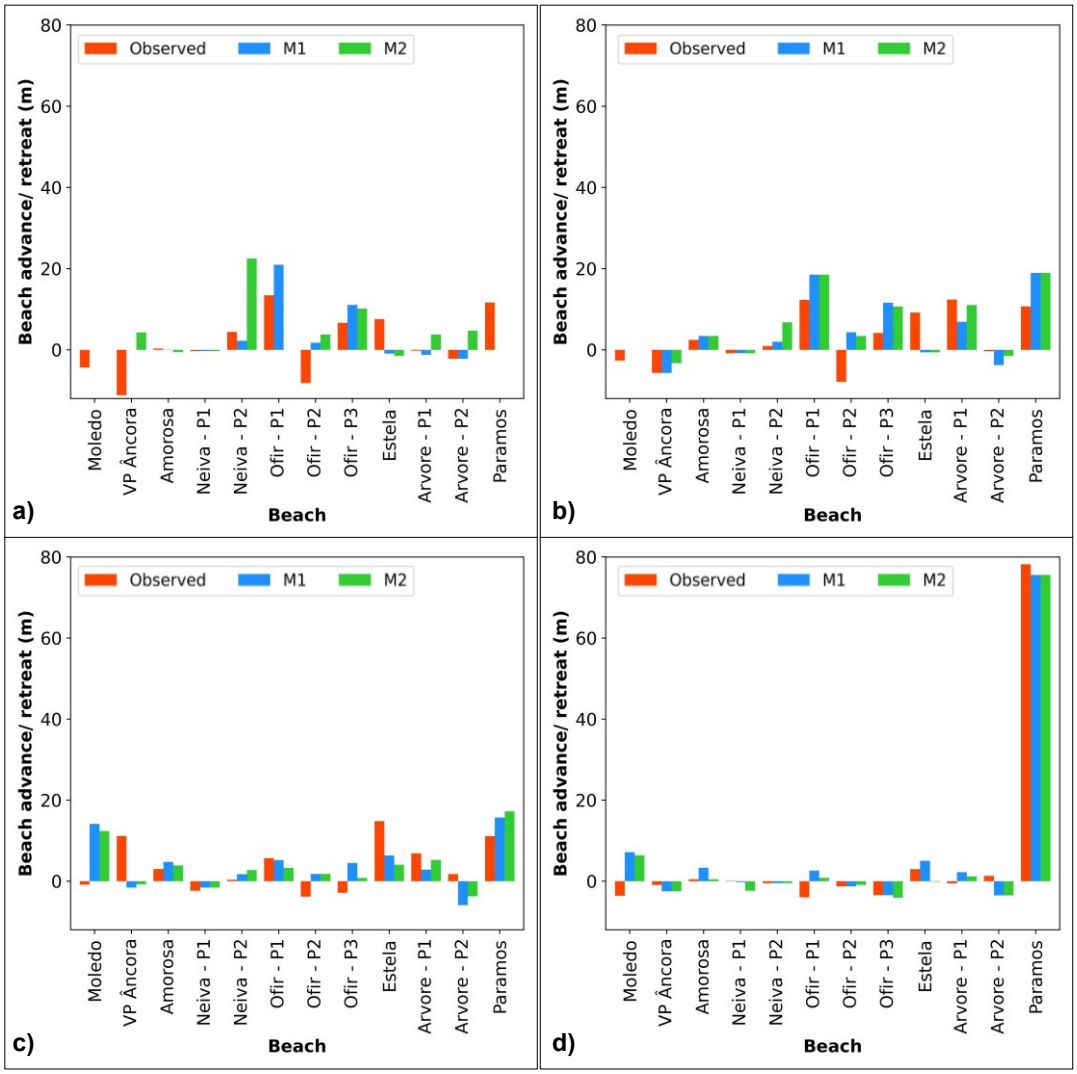


Figure 3-10. Profile advance/retreat. a) 1 m height; b) 2 m height; c) 3 m height; d) 4 m height.

The higher errors at 1 m elevation can be related to poor accuracy of the model for that elevation, even for profiles with low RMSE and MAE. The performance of the models increased at the other elevations, probably because the lower profile section was more dynamic than the others, requiring more data for better forecasting by the RF model. Therefore, the advance and retreat results were not accurate enough if it were necessary to know the exact profile dynamics. Still, both models can indicate the behaviour of the profile. In addition, the advance/retreat results followed the same accretion/erosion behaviour observed in Figure 3-5. Considering the results in which both models forecasted the same pattern of the observed trend, there are nine eroded and sixteen accreted profiles, representing 36% and 64%, respectively. There is a predominance of accretion over erosion in both model results.

These findings suggest that both models were effective in capturing significant features from the simulated profiles, particularly between 1 and 4 m in height. Both beach slope, width, and profile advance/retreat were effectively replicated on the majority of the beaches under consideration. As anticipated, the highest errors were observed in areas where the profile simulation also yielded higher errors.

In order to compare the behaviour of the profiles, the seasonal advance/retreat rate of the beach profiles was determined using the data from COSMO program (Figure 3-11). There is no identifiable morphodynamic behaviour considering all beaches. For instance, Amorosa presented a stable profile during all periods, whilst Moledo presented severe erosion after winters. At Árvore P2, the data showed erosion after the spring-summer surveys and accretion after winter. These results demonstrate that the RF model reproduced this variable with acceptable accuracy despite the absence of a clear pattern considering all beaches.

Although the quantitative error of the model is still considerable for some of the assessed beaches (e.g., Paramos and Estela), the qualitative prediction (advance/retreat) was accurate, indicating that the model can be reliable in indicating the trend. Besides, the presented results show a stable trend, except for Paramos. There is a balance between eroded and accreted profiles, with low profile changes. However, it must be stressed that the profile's advance/retreat was computed considering the time interval between the last two field surveys (approximately four months), which does not show the long-term trend of the profiles, only seasonal trends. The same occurred when analysing the beach slope results, though the quantitative errors were lower.

Hence, the models results were useful to estimate the seasonal morphological evolution of profile features, achieving satisfactory error metrics. In the case of the profile advance/retreat, a classification

model would have achieved excellent metrics, considering that the model correctly identified most of the accretion/erosion patterns.

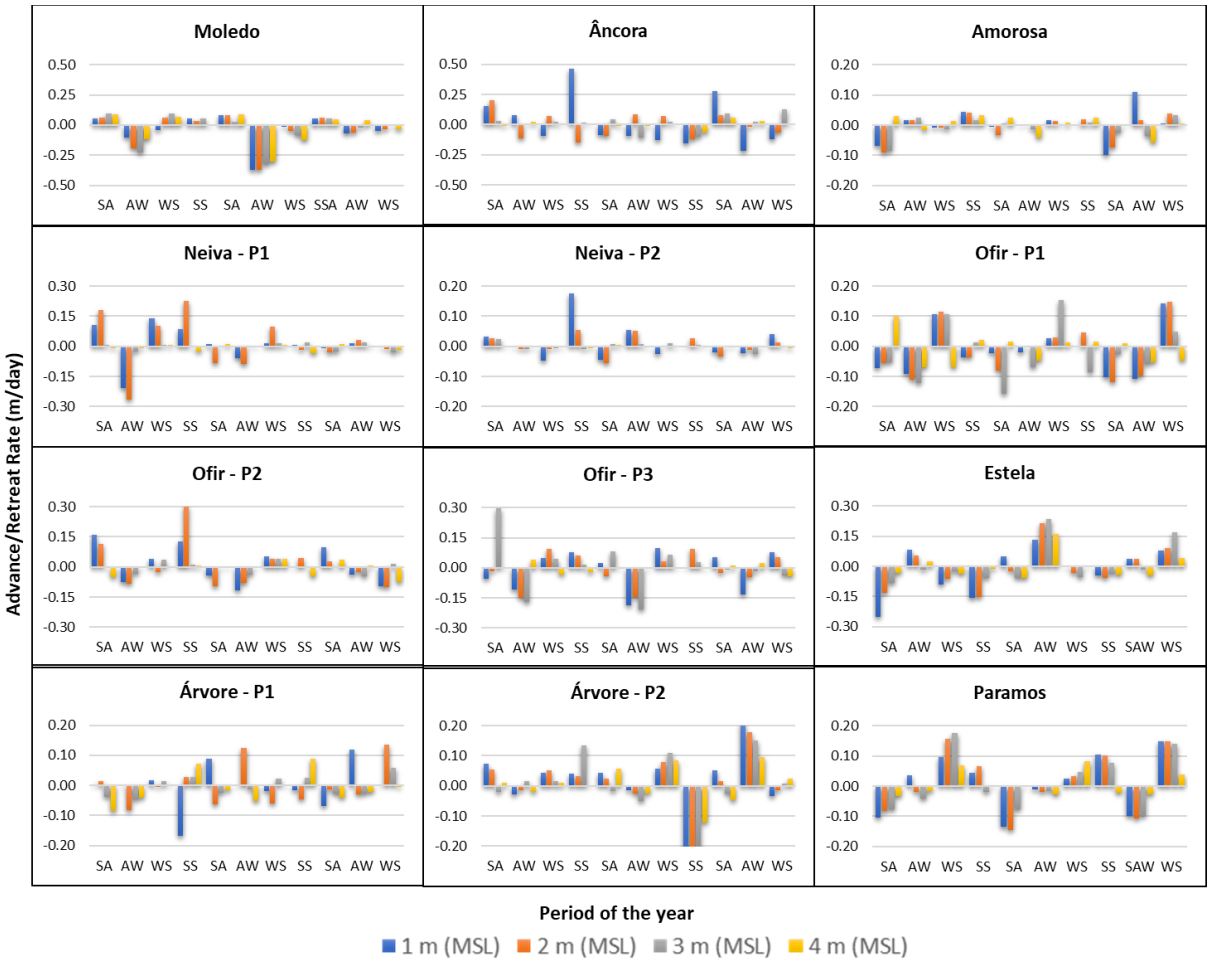


Figure 3-11. Historical erosion/advance rates at the measured profiles. The capital letters represent the seasons in which the rates were computed, with A referring to autumn, W to winter, and S to summer and spring.

3.3.4. Storm simulation scenario

The models' performance under storm conditions is presented in Figure 3-12. The "_4" and "_8" refer to the simulation results with HM0 equal to 4 m and 8 m respectively. The previous profile line is the profile state before the storm simulation and serves as a baseline to assess the storm effect. As expected, increasing the HM0 value resulted in sharper profile changes when comparing the same model. It is also worth noting that M2 results simulated more intense changes compared to M1, as can be seen in Neiva P2 and Estela. Although the results were relatively close, the only scenario in which M1 forecasted more severe changes was in Ofir P1 and Árvore P2.

The model simulated the storm results differently at each location. At Moledo, Âncora, Neiva P2, Ofir P1, Ofir P3, and Árvore P1, the results showed an accretion predominance, while at the other profiles,

erosion patterns prevailed. To evaluate these results, it would be necessary to use process-based models to determine if the outputs are realistic.

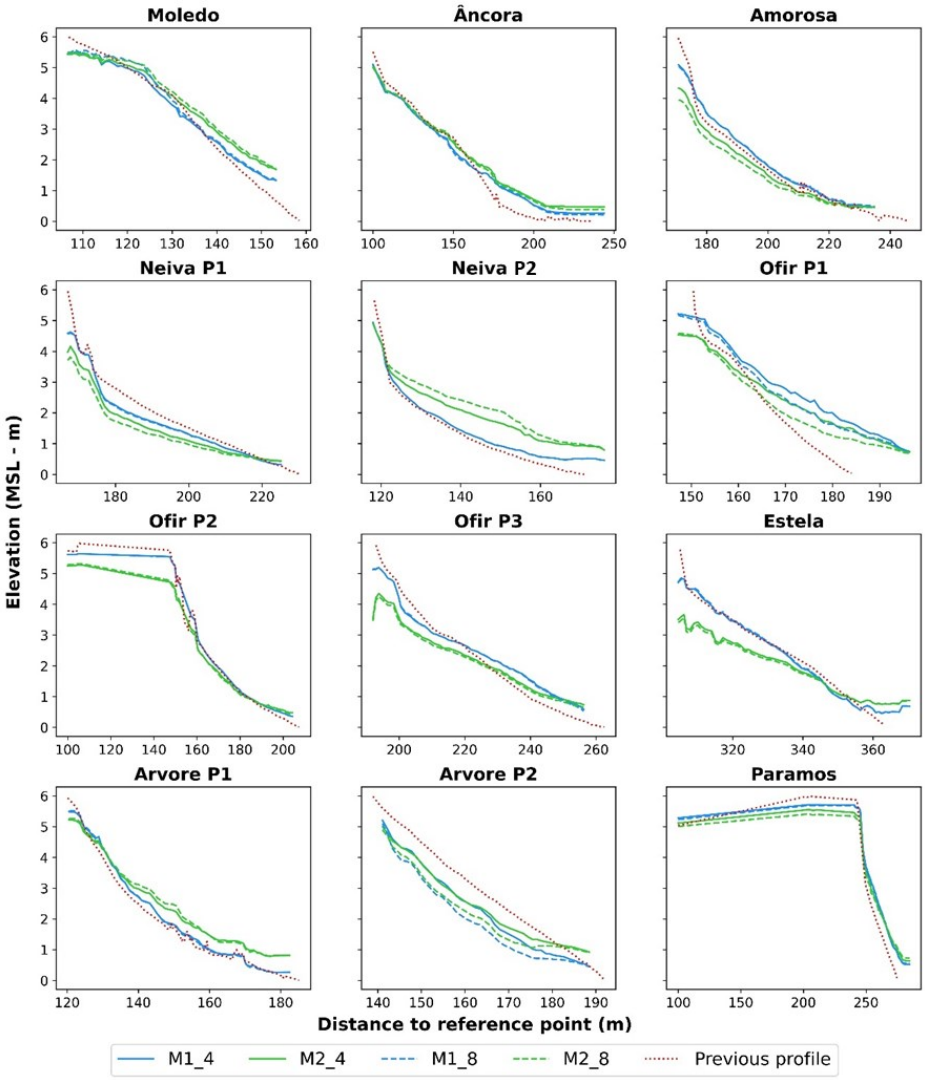


Figure 3-12. Storm simulation results for HM0 are equal to 4 m (_4) and to 8 m (_8).

In addition, the result of the 8 m HM0 simulation is proof of the robustness of data models. Although the HM0 value was never that high in the training dataset, the results are according to what was expected, showing sharper profile changes regardless of the model. However, it must be highlighted that when a model simulates conditions outside the training/calibration range, its uncertainties increase, and its performance diminishes. Hence, the results obtained in this section should be carefully considered. For instance, the model simulation resulted in more accreted profiles than eroded, which is not expected after storm events. However, considering the historical data, the simulation occurred between winter and spring, a period with accretion predominance (Figure 3-5). Hence, the models reproduced the same morphological patterns for these seasons.

3.4. Discussion and conclusions

This section discusses the main challenges of this study, considering the use and processing of data from multiple sources and with different formats. Subsequentially, the limitations of this study's methodology are reviewed, and the main conclusions are presented.

3.4.1. Data models and the importance of open data

The latest advances in AI techniques have transformed data into one of the most valuable but still underused products. Relatively simple models can not only process different types of data but also discover patterns that may result in significant insights for the progress of beach morphodynamics and, consequently, for the management of coastal areas. Hence, local monitoring programs can efficiently gather relevant information from any area of interest. However, there are still some barriers that must be overcome to implement this kind of solution.

Although global monitoring programs such as Copernicus provide essential data, they cannot accurately represent local features and dynamics due to the temporal and spatial resolution of the data. Local monitoring programs have the advantage of knowing the current situation of a specific study area, allowing the identification of critical points and key data to be observed/measured and also the needed timestep for those observations.

In this context, monitoring programs must provide open access to their data to allow the implementation of more robust data models that can merge relevant types of information. The models developed in this study combined numerical model reanalysis results from the Copernicus Marine Service, coastline profiles from the COSMO program, and sediment size data from the MarRisk project. However, one of the challenges of this study was to organise the data of the three sources according to date and location. Each one presents different time horizons and uses different data formats, which must be pre-processed to adequately implement a data model.

Nevertheless, it must be stressed that data models can integrate multidisciplinary data, making it possible to consider water quality, chemical, biological, urban, and anthropogenic data jointly with physical variables like the ones considered in this study. Interdisciplinary approaches are essential to promote more integrated coastal management (de Schipper et al., 2020). It is just a question of determining which knowledge areas influence the phenomena under scrutiny and how each can collaborate to solve the problem from a more holistic point of view.

Therefore, creating open-access systems that concentrate all data types would be important in implementing transversal data models. This would allow any stakeholder to manipulate these data according to their necessities, goals, and expertise, which could be extremely relevant to those entities

that cannot afford to invest in monitoring systems or field surveys. Moreover, part of the necessary work to implement a data model could be reduced by just adopting a pattern in the data types, and centralizing the data would facilitate its standardization, thereby reducing redundant costs. For example, under the COSMO and MarRisk projects, beach profile data were measured on the Northern Portugal coastline. Some beaches were surveyed only in one of the projects, others in both, and others in none.

Finally, it would be pertinent for national and international agencies to take the initiative to create a central multidisciplinary database. It would ensure that all the information could be accessed through a trustworthy agent, reducing the time required to find the desired data if it exists. The Copernicus Marine Service is an excellent example of this kind of platform, providing atmospheric, marine, land, and climate data. However, it is still not integrated into national databases. While it would be challenging to incorporate such efforts in an international platform, national agencies could start organizing their data to facilitate integrating national and international data sources for future projects, which would be extremely useful for implementing multidisciplinary data models.

3.4.2. Coastal processes and data models

This study combined beach profile survey data, sediment characteristics, and wave variables to implement an RF model capable of forecasting the morphological evolution of beach profiles. The achieved results are promising and demonstrate the capacity of data models to reproduce complex morphological patterns. A single model learned to differentiate the patterns at each location and forecasted consistent results using data from the profiles.

One of the main limitations of ML models is interpretability. Whilst drawing conclusions from numerical models is easy, they may underestimate some physical processes. For example, considering a beach with vegetation at the base of the dunes, the morphological evolution will be influenced by the main drivers, such as wave and wind actions, and the presence of vegetation. If the modeler ignores these features, uncertainties will be associated with the model. This problem can be solved by increasing the details and resolution of the models, but it also greatly increases the modelling efforts. For instance, the same can be considered for the granulometry of beach sediments. It varies for each profile, and sometimes, even within the same profile, simplifications are adopted to make the models feasible.

Data models have the advantage of intrinsically considering these features, as the influence of vegetation and sediment size would be reflected in the field data. This contribution does not need to be explicit, as is the case with numerical modelling methodologies. Whereas numerical models can easily extrapolate physical conditions, data models are limited to the training variables. If no data are acquired during an extreme event, the data model will be inappropriate for simulating such conditions.

Furthermore, XAI can solve the interpretability barrier of data models. It consists of a series of methods and tests that determine the relevance of each input to the output, thereby helping to understand the limitations of such methodologies. Therefore, by adopting these concepts, it will be possible to compare process-based models with data models, increase knowledge about the limitations of each model, and allow the creation of adequate modelling solutions for each case. The use of XAI frameworks is strongly encouraged in future research.

3.4.3. Methodology limitations

The data models implemented in this study considered wave characteristics, sediment size, and beach profile variables. However, the morphological evolution of beach profiles is also influenced by other variables that were not considered. For instance, at higher elevations, the wind can be an important factor to be considered in sediment transport, mainly for finner sediment beaches.

Additionally, as previously explained, one of the sources of the errors obtained in Moledo could be the lack of consideration of the Minho River discharge. Conversely, fluvial sediment transport's importance could also explain some of the errors in other beach profiles. One of the best results was in Neiva, which is only affected by the discharges of a small river. In northern Portugal, five main rivers can contribute to the fluvial sediment transport towards the ocean: the Minho River (influence Moledo and Âncora beaches), the Lima River (influence Amorosa beach), the Cávado River (influence Ofir beach), the Ave River and the Douro River (both affecting Árvore beach). Considering the river plumes' influence on the data models could reduce the errors mainly at the lower profile elevations.

Unfortunately, most river flow dischargers were unavailable for the period selected in this study and could not be incorporated into the implemented models. Depending on the data availability, two possible solutions can be considered to overcome this constraint. The first is to access data from environmental agencies that monitor the discharges of rivers, which directly feed this data to the data models. The second is the use of hydrological models to simulate these data based on historical precipitation records.

Furthermore, the COSMO data covered the morphological evolution of the profiles for approximately five years, which is a short period for long-term morphological process characterization. Hence, the models implemented in this study may not be adequate for forecasting long-term profile changes, only seasonal profiles. The data captured some trends in the dynamics of the profiles, such as in Âncora, Estela, Árvore P2, and Ofir P1. However, both models achieved acceptable metrics at these locations. It is necessary to apply the models to a longer-measured profile time series to assess the

capacity of the model to reproduce long-term morphodynamic patterns, but unfortunately, no data are available to accomplish this task.

Furthermore, the methodology could be further developed to simulate short-term morphological processes, such as the morphological changes induced by extreme events. The role of water level variations during daily intervals are more evident than during longer periods. For example, the effects of a storm surge will be highlighted if the data time interval is adequate for capturing this event. However, if the time window is too long, storm-induced water level fluctuations effects on the beaches morphology may be obscured. To capture short-term beach morphodynamics, it is necessary to have either a longer data series of beach profile measurements that include extreme events or a selective data series that precisely registers the morphologies before and after extreme events.

Another limitation of this model is the lack of data about beach nourishment during the study period. Human interventions significantly modify the beach profiles, changing their natural evolution dynamic. Beach fill operations have been increasing in Portugal in the last decades. According to Pinto et al. (2020), the northern Portuguese coast had more than 30 interventions between 1957 and 2017. Therefore, data about structural interventions on beaches could also increase the performance of data models, but this data was not available to the authors. However, quantitative data regarding this type of intervention is usually restricted to those who execute and supervise it. The few available open data are mostly qualitative, for instance, the date of the intervention and the place, but not its exact location and the nourishment volumes, which would be the information needed to be included in the data models.

3.4.4. Conclusions

This study combined data from different sources to create a model capable of forecasting the evolution of multiple beach profiles considering the influence of wave and sand size. Two different models were implemented using the RF algorithm, in which M1 used wave, beach profiles, and sediment size variables as inputs, whereas M2 considered the same variables except for the sediment size. The output variable was the elevation relative to the mean sea level, which was compared with observational datasets to depict the accuracy of the data models.

The training results showed that both models had low error metrics, achieved similar values, and demonstrated the suitability of the data models for representing beach morphology. The average RMSE was 0.04 m, and the MAE was lower than 0.15 m for both models. These metrics indicated that the difference in the model input data was negligible. The type of profile did not influence the performance of the models.

However, when the forecasting capacity of the models was tested, M1 outperformed M2. Considering the models' global performance, the average RMSE and MAE were respectively 0.54 m and 0.45 m for M1, and 0.58 m and 0.49 m for M2. When assessing the local errors, the best result occurred in Neiva P1, where M1 had an RMSE of 0.18 m while M2 had 0.24 m, values similar to those in Amorosa (0.20 m and 0.26 m, respectively). Hence, it can be concluded that the sediment size data improved the results, reducing the average maximum error by 19 cm in the training dataset and the RMSE and the variance of M1 when compared to M2.

Additionally, the profile slope and retreat/advance were computed to assess the difference between the observed and simulated results. Both models accurately predicted the slope of the profiles, but M1 outperformed M2, with an MAE of 2.1%. M2 had an MAE of 2.4%. The profile retreat/advance results were inaccurate, but both models had a good qualitative agreement. The best retreat/advance prediction results were obtained at an elevation of 2 m, in which the models correctly forecasted the behaviour of nine profiles. Additionally, the predominance of accretion in the profiles during this simulation followed the same pattern observed during the monitoring period. Hence, simulation results achieved satisfactory metrics in predicting the morphological features of the profiles.

The simulation under storm conditions showed that the model is reliable beyond the training data range, but their results must be carefully considered. Data obtained from physics-based numerical models can be included to enhance the robustness of data models. This approach will allow the complementation of the observed field data required to properly feed data models with consistent time series, such as for extreme events, thus increasing its applicability.

In summary, this study reached the objective of combining different data sources to implement a RF model to forecast the evolution of multiple beach profiles. The error metrics achieved by both models were within the error range obtained by previous studies. Therefore, data models can be a valuable tool for monitoring and forecasting the evolution of beach profiles as long as there is enough quality data to implement them. Future studies could assess the interpretability of ML models applied to coastal datasets and test other variables besides those used in this study, contributing to the knowledge advance about complex coastal morphodynamics processes.

3.5. References

Aagaard T, Greenwood B, Hughes M. Sediment transport on dissipative, intermediate and reflective beaches. Earth Sci Rev 2013;124:32–50. https://doi.org/10.1016/j.earscirev.2013.05.002.

Alveirinho Dias JM, Nittrouer CA. Continental shelf sediments of northern Portugal. Cont Shelf Res 1984;3:147–65. https://doi.org/10.1016/0278-4343(84)90004-9.

Antolínez JAA, Méndez FJ, Camus P, Vitousek S, González EM, Ruggiero P, et al. A multiscale climate emulator for long ‐ term morphodynamics (MUSCLE ‐ morpho). J Geophys Res Oceans 2016;121:775 ‐ 91. https://doi.org/10.1002/2015JC011107.

APA, Agência Portuguesa do Ambiente. COaStal MOnitoring Program (COSMO) 2023. https://cosmo.apambiente.pt/data (accessed April 17, 2023).

Araújo M da A. Depósitos continentais e marinhos na plataforma litoral da região do Porto Importância da tectónica na sua organização espacial. Ciências Da Terra 2000;14:111–22.

Benveniste J, Cazenave A, Vignudelli S, Fenoglio-Marc L, Shah R, Almar R, et al. Requirements for a Coastal Hazards Observing System. Front Mar Sci 2019;6. https://doi.org/10.3389/fmars.2019.00348.

Bio A, Gonçalves JA, Iglesias I, Pinho JL, Vieira L, Vieira JM, et al. Coastal erosion: a northern-portuguese case study. In: Martí Ezpeleta A, Lorenzo González N, Royé D, Díaz Poso A, editors., Madrid: Asociación Española de Climatología; 2022.

Bio A, Gonçalves JA, Pinho J, Vieira L, Vieira J, Smirnov G, et al. Indicadores de vulnerabilidade de erosão costeira: Um estudo de caso no Norte de Portugal. Revista de Gestão Costeira Integrada 2020a;20:197–209. https://doi.org/10.5894/rgci-n337.

Bio A, Gonçalves JA, Pinho J, Vieira L, Vieira J, Smirnov G, et al. Indicadores de vulnerabilidade de erosão costeira: Um estudo de caso no Norte de Portugal. Revista de Gestão Costeira Integrada 2020b;20:197–209. https://doi.org/10.5894/rgci-n337.

Biolchi LG, Unguendoli S, Bressan L, Giambastiani BMS, Valentini A. Ensemble technique application to an XBeach-based coastal Early Warning System for the Northwest Adriatic Sea (Emilia-Romagna region, Italy). Coastal Engineering 2022;173. https://doi.org/10.1016/j.coastaleng.2022.104081.

Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification And Regression Trees. Routledge; 2017. https://doi.org/10.1201/9781315139470.

Chen JC, Ning SK, Chen HW, Shu CS. Flooding probability of urban area estimated by decision tree and artificial neural networks. Journal of Hydroinformatics, vol. 10, IWA Publishing; 2008, p. 57–67. https://doi.org/10.2166/hydro.2008.009.

Chen Q, Wang N, Chen Z. Simultaneous mapping of nearshore bathymetry and waves based on physics-informed deep learning. Coastal Engineering 2023;183:104337. https://doi.org/10.1016/j.coastaleng.2023.104337.

Copernicus Marine Service. Atlantic -Iberian Biscay Irish- Ocean Wave Reanalysis 2022. https://doi.org/https://doi.org/10.48670/moi-00030.

Cutler A, Cutler DR, Stevens JR. Random Forests. In: Zhang C, Ma Y, editors. Ensemble Machine Learning, New York, NY: Springer New York; 2012, p. 157–75. https://doi.org/10.1007/978-1-4419-9326-7_5.

Deltares. Delft3D-Flow: Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments - User Manual 2018.

Fontán-Bouzas Á, Andriolo U, Silva PA, Baptista P. Wave Impact Analysis on a Beach-Dune System to Support Coastal Management and Nourishment Works: The Showcase of Mira, Portugal. Front Mar Sci 2022;9. https://doi.org/10.3389/fmars.2022.861569.

Gharagozlou A, Anderson DL, Gorski JF, Dietrich JC. Emulator For Eroded Beach And Dune Profiles Due To Storms. J Geophys Res Earth Surf 2022;127. https://doi.org/10.1029/2022JF006620.

Gomes A, Pinho JLS, Granja HM. Morphodynamic modelling of a gravel beach at the NW Portuguese coast. IX Simpósio sobre a Margem Ibérica Atlântica, Coimbra: 2018, p. 309–10.

Granja H, Pinho JL. Coastal Defense in NW Portugal: The Improbable Victory, 2012, p. 251–66. https://doi.org/10.1007/978-94-007-4123-2_15.

Gunawan PH, Iryanto, Ghozali AL, Ismantohadi E, Baizal ZKA, Satrio A. Data-driven Shoreline Change Forecasting on Eretan Beach Using Random Forest, Institute of Electrical and Electronics Engineers (IEEE); 2023, p. 1–4. https://doi.org/10.1109/icacnis57039.2022.10054972.

Gunning D, Stefik M, Choi J, Miller T, Stumpf S, Yang G-Z. XAI—Explainable artificial intelligence. Sci Robot 2019;4. https://doi.org/10.1126/scirobotics.aay7120.

Iglesias I, Pinho JL, Avilez-Valente P, Melo W, Bio A, Gomes A, et al. Improving Estuarine Hydrodynamic Forecasts Through Numerical Model Ensembles. Front Mar Sci 2022;9. https://doi.org/10.3389/fmars.2022.812255.

Irizarry RA. Introduction to Data Science. 2019. https://doi.org/10.1201/9780429341830.

Jackson D, Short A. Sandy beach morphodynamics. Elsevier; 2020.

Jayson-Quashigah PN, Appeaning Addo K, Amisigo B, Wiafe G. Assessment of short-term beach sediment change in the Volta Delta coast in Ghana using data from Unmanned Aerial Vehicles (Drone). Ocean Coast Manag 2019;182. https://doi.org/10.1016/j.ocecoaman.2019.104952.

Kalligeris N, Smit PB, Ludka BC, Guza RT, Gallien TW. Calibration and assessment of process-based numerical models for beach profile evolution in southern California. Coastal Engineering 2020;158. https://doi.org/10.1016/j.coastaleng.2020.103650.

Karunarathna H, Horrillo-Caraballo JM, Reeve DE. Prediction of cross-shore beach profile evolution using a diffusion type model. Cont Shelf Res 2012;48:157–66. https://doi.org/10.1016/j.csr.2012.08.004.

Kazeminezhad MH, Etemad-Shahidi A. A new method for the prediction of wave runup on vertical piles. Coastal Engineering 2015;98:55–64. https://doi.org/10.1016/j.coastaleng.2015.01.004.

Kubat M. An Introduction to Machine Learning. 2017. https://doi.org/10.1007/978-3-319-63913-0.

Lee S, Lim C, Lee J-L. Prediction of Long-Term Beach Profile Evolution Due to Episodic Wave Incidence Under Tidal Environment. Front Mar Sci 2022;9. https://doi.org/10.3389/fmars.2022.831262.

Luijendijk A, Hagenaars G, Ranasinghe R, Baart F, Donchyts G, Aarninkhof S. The State of the World's Beaches. Sci Rep 2018;8:1–11. https://doi.org/10.1038/s41598-018-24630-6.

Matheen N, Harley MD, Turner IL, Splinter KD, Simmons JA, Thran MC. Bathymetric data requirements for operational coastal erosion forecasting using xbeach. J Mar Sci Eng 2021;9. https://doi.org/10.3390/jmse9101053.

McAllister E, Payo A, Novellino A, Dolphin T, Medina-Lopez E. Multispectral satellite imagery and machine learning for the extraction of shoreline indicators. Coastal Engineering 2022;174:104102. https://doi.org/10.1016/j.coastaleng.2022.104102.

McGranahan G, Balk D, Anderson B. The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. Environ Urban 2007;19:17–37. https://doi.org/10.1177/0956247807076960.

Melo W, Pinho J, Iglesias I, Bio A, Avilez-Valente P, Vieira J, et al. Hydro- and Morphodynamic Impacts of Sea Level Rise: The Minho Estuary Case Study. J Mar Sci Eng 2020;8:441. https://doi.org/10.3390/jmse8060441.

Mil-Homens M, Brito P, Magalhães V, Rosa M, Neres M, Silva Marta, et al. Integrated geophysical and sedimentological datasets for assessment of offshore borrow areas: the CHIMERA project (western Portuguese Coast). Geological Society, London, Special Publications 2022;505:131–53. https://doi.org/10.1144/SP505-2019-100.

Otto P, Piter A, Gijsman R. Statistical analysis of beach profiles – A spatiotemporal functional approach. Coastal Engineering 2021;170. https://doi.org/10.1016/j.coastaleng.2021.103999.

Pedregosa F, Varoquaux G, Gramfort A, Michel V. and Thirion B, Grisel O, Blondel M, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 2011;12:2825–30.

Pedrosa A de S. Risco de erosão no litoral norte de Portugal: uma questão de ordenamento do território. Sociedade & Natureza 2013;25:133–46. https://doi.org/10.1590/S1982-45132013000100011.

Pinho J, Gomes A, Weber de Melo W, Iglesias I, Bio A, Avilez-Valente P, et al. Comparison between traditional and digital image based approaches for beach sediment size characterization. A case study at the NW Portuguese coast. In Preparation 2024.

Pinto CA, Silveira TM, Teixeira SB. Beach nourishment practice in mainland Portugal (1950–2017): Overview and retrospective. Ocean Coast Manag 2020;192:105211. https://doi.org/10.1016/j.ocecoaman.2020.105211.

Ponte Lira C, Nobre Silva A, Taborda R, Freire de Andrade C. Coastline evolution of Portuguese low-lying sandy coast in the last 50 years: an integrated approach. Earth Syst Sci Data 2016;8:265–78. https://doi.org/10.5194/essd-8-265-2016.

Pörtner H-O, Roberts DC, Tignor M, Poloczanska ES, Mintenbeck K, Alegría A, et al. IPCC, 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, Uk and New York, NY, USA: 2022. https://doi.org/10.1017/9781009325844.

Prodger S, Russell P, Davidson M. Grain-size distributions on high-energy sandy beaches and their relation to wave dissipation. Sedimentology 2017;64:1289–302. https://doi.org/10.1111/sed.12353.

Razavi S, Tolson BA, Burn DH. Review of surrogate modeling in water resources. Water Resour Res 2012;48. https://doi.org/10.1029/2011WR011527.

Roelvink D, Reniers A, van Dongeren A, van Thiel de Vries J, McCall R, Lescinski J. Modelling storm impacts on beaches, dunes and barrier islands. Coastal Engineering 2009;56:1133–52. https://doi.org/10.1016/j.coastaleng.2009.08.006.

Sanderson MG, Teixeira M, Fontes N, Silva S, Graça A. The probability of unprecedented high rainfall in wine regions of northern Portugal. Clim Serv 2023;30:100363. https://doi.org/10.1016/j.cliser.2023.100363.

Santos AI, Oliveira A, Pinto JP, Freitas MC. Hydrodynamic and Sediment Transport Patterns in the Minho and Douro Estuaries (NW Portugal) Based on ADCP Monitoring Data: Part 1-Tidal Sediment Exchanges. Coasts 2021;1:31–55. https://doi.org/10.3390/coasts1010003.

de Schipper MA, Ludka BC, Raubenheimer B, Luijendijk AP, Schlacher ThomasA. Beach nourishment has complex implications for the future of sandy shores. Nat Rev Earth Environ 2020;2:70–84. https://doi.org/10.1038/s43017-020-00109-9.

Suzuki T, Cox DT. Evaluating XBeach performance for extreme offshore-directed sediment transport events on a dissipative beach. Coastal Engineering Journal 2021;63:517–31. https://doi.org/10.1080/21664250.2021.1976452.

UN-Habitat UNHSP-. World Cities Report 2022. Nairobi, Kenya: 2022.

Vieira BFV, Pinho JLS, Barros JAO, Antunes do Carmo JS. Hydrodynamics and morphodynamics performance assessment of three coastal protection structures. J Mar Sci Eng 2020;8. https://doi.org/10.3390/jmse8030175.

Viitak M, Avilez-Valente P, Bio A, Bastos L, Iglesias I. Evaluating wind datasets for wave hindcasting in the NW Iberian Peninsula coast. Journal of Operational Oceanography 2021;14:152–65. https://doi.org/10.1080/1755876X.2020.1738121.

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020;17:261–72. https://doi.org/10.1038/s41592-019-0686-2.

Vousdoukas MI, Ranasinghe R, Mentaschi L, Plomaritis TA, Athanasiou P, Luijendijk A, et al. Sandy coastlines under threat of erosion. Nat Clim Chang 2020;in press. https://doi.org/10.1038/s41558-020-0697-0.

Waskom M. seaborn: statistical data visualization. J Open Source Softw 2021;6:3021. https://doi.org/10.21105/joss.03021.

Weber de Melo W, Pinho J, Iglesias I. Coastal morphodynamic emulator for early warning short-term forecasts. Environmental Modelling & Software 2023;165:105729. https://doi.org/10.1016/j.envsoft.2023.105729.

Weber de Melo W, Pinho J, Iglesias I. Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model. Journal of Hydroinformatics 2022. https://doi.org/10.2166/hydro.2022.068.

Ye Z, Shi F, Zhao X, Hu Z, Malej M. A data-driven approach to modeling subgrid-scale shallow marsh hydrodynamics. Coastal Engineering 2021;166. https://doi.org/10.1016/j.coastaleng.2021.103856.

CHAPTER 4.

Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model.

Article published in November 2022 in the Journal of Hydroinformatics.

Weber de Melo, W., Pinho, J., Iglesias, I. (2022). Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model. Journal of Hydroinformatics, 24(6), 1254-1268.

"Nothing in life is to be feared; it is only to be understood."

Marie Curie (1867 – 1934) – Polish physicist and chemist. Physics Nobel Prize winner in 1903.

Chapter 4. Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model.

Abstract

Coastal and estuarine areas present remarkable environmental values, being key zones for the development of many human activities such as tourism, industry, fishing, and other ecosystem services. To promote the sustainable use of these services, effectively managing these areas and their water and sediment resources for present and future conditions is of utmost importance to implement operational forecast platforms using real-time data and numerical models. These platforms are commonly based on numerical modelling suites, which can simulate hydro-morphodynamic patterns with considerable accuracy. However, in many cases, considering the high spatial resolution models that are necessary to develop operational forecast platforms, a high computing capacity is also required, namely for data processing and storage. This work proposes the use of artificial intelligence (AI) models to emulate morphodynamic numerical model results, allowing us to optimize the use of computational resources. A convolutional neural network was implemented, demonstrating its capacity in reproducing the erosion and sedimentation patterns, resembling the numerical model results. The obtained root mean squared error was 0.59 cm, and 74.5 years of morphological evolution was emulated in less than 5 s. The viability of surrogating numerical models by AI techniques to forecast the morphological evolution of estuarine regions was clearly demonstrated.

4.1. Introduction

Understanding the morphological evolution of coastal and estuarine regions is essential to promote the sustainable use of its natural resources and protect populations and ecosystems. Particularly, estuarine areas present a high economic, social, and environmental strategic importance, with valuable ecosystem services such as food resources, leisure, tourism, energy, water, and raw materials. They also present advantages for the development of navigation activities, namely conditions for harbour structure location and a direct connection with oceanic and inland waterways, promoting recreational navigation, ship transport, or the fishing industry, among others (Lewis et al. 2021; Pearson et al. 2021; Taylor & Suthers 2021). However, the anthropogenic pressure in these areas is increasing, as well as the effects of climate change, threatening the availability and quality of these natural system services (Elliott et al. 2019; Pörtner et al. 2022; Xie et al. 2022).

However, it must be noticed that the morphology of alluvial estuarine channels and the hydrodynamic behaviour and sediment transport along them are linked in a feedback loop, resulting in a complex interaction. The morphology is shaped by these two processes and, at the same time, it defines the main sediment transportation patterns (Church & Ferguson 2015; Poelhekke et al. 2016). High current velocities increase the bottom shear stress in the channels, and consequently, the sediment transport, which strongly depends on the sediment size. The sediments are transported downstream and deposited in zones with lower current velocities, creating sediment reservoirs and changing the morphology of the channel. When the hydrodynamic conditions are modified, these deposits can be eroded, serving as a sediment source for other estuarine areas.

These dynamics reinforce the strong necessity to fully understand and forecast the morphology of coastal and estuarine areas to promote navigation security and implement adequate measures to protect the populations and the habitats. For instance, forecasting the sedimentation and erosion patterns of coastal and estuarine regions is necessary to understand the effects of accretion areas during flood events, helping, for example, to optimize dredging operations to maintain the navigation channel depth. The knowledge of sediment transport trends and coastal and fluvial geomorphology is essential to understand the possible impacts of floods, implement measures to avoid the effects of extreme events, and regulating the navigation and use of the sediment resources that are abundant in these ecosystems (Kandolf & Piégay 2003). In addition, there is an increase in the pressure and the vulnerability in the coastal regions due to the intensification of human settlements and commercial activities, which reinforce the need for efficient tools that help to understand the complexity of sediment supply and transport (Church & Ferguson 2015).

In this context, the development of tools to monitor and assess morphological changes is of utmost importance. A common methodology to study the morphological evolution of coastal and estuarine regions is based on the implementation of numerical models, which have already reached the capacity to generate highly accurate solutions (Guerrero et al. 2013; Iglesias et al. 2019; Pinho et al. 2020; Zhou et al. 2021). Sediment transport models, coupled with hydrodynamic numerical models, allow the simulation of erosion and accretion patterns of rivers, estuaries, and coastal areas. However, the natural process that takes place in these environments can take several centuries to reach an equilibrium condition. The complexity of morphological numerical simulations is also increased due to hydrodynamic changes that could take place at different time and spatial scales. The hydrodynamics is affected by tide variations, for example, presenting hourly to daily variations that can vary from centimetres to several metres, with a direct effect on the current velocity and, therefore, on the bottom shear stresses. Longer

time-scale processes like sea level rise (SLR) due to climate change conditions also have a strong effect on the hydrodynamic conditions. However, it must be taken into account that these short- and long-term hydrodynamic variations can result in barely noticed variations in the morphological conditions, depending, among other factors, on the bottom slope and the sediment characteristics (Roering et al. 2001).

Additionally, numerical modelling-based methodologies are usually limited by computational resources, such as memory capacity and processing power. This is even more limited when long-term simulations with high spatial resolution models are required, which is a common practice when the focus of the numerical models is the morphodynamic evolution. The higher the resolution of the model and the amount of data that it generates, the higher the computational resource requirements (processing power and storage capacity), demanding, sometimes, prohibitive computational infrastructures and making the application of this type of platform unfeasible. Nevertheless, long-term solutions of high-resolution models are essential to assess the hydro-morphodynamics on local scales.

A possible solution to overcome that difficulty is to use data-based models, also known as surrogate models (Poelhekke et al. 2016; Chen et al. 2018; Parker et al. 2019). These models are implemented using artificial intelligence (AI) algorithms that can learn relationships between the input and the output data instead of solving the differential equations that describe the fluid and sediment transport. Therefore, the focus of this work was to implement a deep convolutional neural network (CNN) to emulate erosion and sedimentation patterns using the hydrodynamic results of a numerical model as input. The architecture of the network used different branches to generate feature maps of the input variables with different spatial resolutions for posterior concatenation in the last layer of the network (Ronneberger et al. 2015; Santhanam et al. 2016; Li et al. 2021). The depth-averaged current velocity and the bottom shear stress variables were used as inputs, and the erosion and sedimentation results were obtained as outputs. The numerical model was implemented in the Delft3D (D3D, https://oss.deltares.nl/web/delft3d) software by Elmilady et al. (2022), which studied the SLR impact on the long-term morphological evolution of intertidal sandy shoals in a constrained channel-shoal system.

The use of AI in water resources is growing with the increase of data availability and computational processing capacity. For example, genetic programming was used for prediction of the sediment settling velocity (Goldstein & Coco 2014), a deep learning model was applied to forecast the water level in lakes (Barzegar et al. 2021), a long short-term memory (LSTM) model was used to forecast the surface water temperature of large-deep reservoirs (Wang et al. 2022), and a multilayer perceptron model was applied to relate the water discharge with the sediment rate in the Mississippi river (Jain

2008). Most of these problems were simplified to a single dimension, making feasible the application of the mentioned techniques. Since the aim of this study was to emulate the results obtained with a two-dimensional (2DH) numerical model, CNNs were selected due to their capacity to process multidimensional data, allowing the use of 2DH model results as input and output.

CNNs have already been successfully applied to predict the water level and water quality indicators in the Nakdong river basin (Baek et al. 2020), to predict the daily rainfall–runoff at two monitoring stations in Chau Doc and Can Tho (Van et al. 2020), for flood forecast in the Xixian basin (Chen et al. 2021), and to reproduce beach morphology from coastal video monitoring systems in France (Soloy et al. 2021). However, applications of CNN to assess morphodynamic evolution are still rare and mainly focused on sediment load modelling (Nagy et al. 2002; Gupta et al. 2021). It is expected that this study will contribute to fill this gap of knowledge in the application of CNNs to simulate the long-term morphology evolution of coastal systems, demonstrating their capacity and limitations to predict accretion and erosion patterns.

4.2. Methods

The used methodology considers four main steps. The first step consists of the implementation of the hydro-morphodynamic numerical model to generate input data for the AI model. This task was based on the work developed by Elmilady et al. (2022). In their study, a tidal basin with an open seaward boundary and a landward boundary was implemented in the D3D software to investigate SLR impact on the long-term morphodynamics. A 200-year morphological evolution simulation was performed and the results were made publicly available. Using these results, the present study was mainly focused on the AI model development, and details about the hydro-morphodynamic numerical model implementation can be found in the previously referenced work.

Secondly, the results of the depth-averaged current velocity and the bottom shear stress, two hydrodynamic variables obtained with the numerical model, were exported as images to be used as input for the emulator. Additionally, the bathymetric variation results (cumulative erosion and sedimentation) were also exported to be used as an output. In this process, special care was taken to ensure that all the figures were uniform in terms of size, colour scale, and represented area. The difference in the values of the selected variables between consecutive time steps was computed to reduce the upper- and lower-pixel value limits. The limit values of the selected variables were —0.0025 m/s and 0.0025 m/s for the depth-averaged current velocity, -0.04 Pa and 0.04 Pa for the bottom shear stress, and -0.15 m and 0.15 m for the variation in the bathymetry (negative values represent erosion and positive accretion).

In the third step, the numerical model result datasets were organized. The images were loaded and split into training and testing datasets, selecting 292 images for training and 73 for testing.

In the last step, the architecture of the model was set, the training optimizer RMSprop was selected, and the metrics to evaluate the model performance were defined. All these tasks are explained in detail in the following paper subsections.

4.2.1. Numerical model setup and data pre-processing

The D3D is a fully integrated software suite for coastal, riverine, and estuarine areas, simulating non-steady flow and transport phenomena (Deltares 2018). The D3D solves the three-dimensional (3D) time-averaged Navier–Stokes equations for incompressible fluids in the x, y, and z directions, assuming the shallow water equations and the Boussinesq assumptions (Agoshkov et al. 1993; Akan 2006; Deltares 2018). For sediment transport, the software solves a 3D advection-diffusion equation for suspended sediments (Deltares 2018). It considers the Van Rijn approach for sand transport, including bed and suspended loads, and the formulation also considers the combined effects of currents and waves (Elmilady et al. 2022). The implementation of these kinds of models is based on the definition of a numerical grid where the transport equations are solved, depending on the boundary conditions. Finer grids produce more detailed solutions, although they increase the simulation computational costs. On the contrary, coarse grids are faster to process, although they can oversimplify the information, resulting in less accurate models.

The numerical model used in this paper was implemented by Elmilady et al. (2022), whose work investigated the SLR effects on the evolution of long-term morphodynamics of a sandy channel-shoal system dominated by intertidal sandy shoals, considering the effects of wind and waves in similar conditions to those that occur in the Wadden Sea. The numerical model had an outer coarse grid with 100 m × 200 m of spatial resolution and an inner fine grid with 33 m × 66 m resolution. Two sandy sediment fractions of 100 μm and 250 μm were considered to represent fine and coarse fractions, respectively. Two hydrodynamic years were simulated, including a morphological acceleration factor of 100, resulting in 200 years of morphological evolution.

The numerical model had one open sea boundary where a semidiurnal tidal component with an amplitude of 1.5 m was adopted. All the other boundaries were closed. Also, a spectral wave model implemented in Simulating Waves Nearshore (SWAN) was used to simulate wind-generated waves. Additional details of the model can be found in Elmilady et al. (2022).

The D3D model results from the high-resolution domain were selected as input and output to the AI model. Images of bed shear stress and square depth-averaged velocity magnitude were generated with

a 6-month frequency, resulting in 365 images for each one of the variables. Images of erosion and accretion were also processed with the same frequency to be used by the CNN as output. Each image showed the variation in the cumulative erosion/sedimentation considering the current time step (t) and the previous time step (t-1). All images were processed, ensuring that the same colour scale, image dimensions, and spatial resolution (DPI) were adopted. The D3D output file was initially loaded using the D3D MATLAB interface, the numerical model results were read, and a correspondent image was created and exported to a proper image format file. This procedure is repeated for all the simulation time steps. The images were exported at 300 DPI, resulting in images with 1,360 pixels × 1,060 pixels. A grey colour map was also considered.

The values of the bottom shear stress varied between 0 Pa and 0.34 Pa, the depth-averaged velocity magnitude varied between 0 m/s and 0.35 m/s, and the bathymetry varied between -11.08 m and 13.59 m. These limits were determined by computing the maximum and minimum values observed in the domain, considering all the time steps of the numerical model results. It must be stressed that the input and output values of the images are the differences between two consecutive time steps but not the value of the variable at single time steps.

After processing the images, the TensorFlow function *image_dataset_from_directory* was used, creating a matrix containing each variable data (Abadi et al. 2015). 80% (292 images) of the data generated by the numerical model was used to train the model and define the values of the CNN parameters. The remaining 20% (73 images) of the data generated was used to assess the performance of the model. These division rates were selected considering the size of the available dataset. It is recommended to use 10% – 40% of the data for testing and the remaining for training (Palani et al. 2008; Barzegar et al. 2016; Agrawal & Mittal 2020). Furthermore, the training dataset must provide sufficient and varied samples that allow the model to learn the data pattern. The testing dataset must include values that are not necessarily considered in the training dataset to ensure accurate forecasting for broad conditions, preventing overfitted models.

4.2.2. Input Variables

The number of input variables of the CNN model should be defined before generating the artificial dataset that includes the hydrodynamic results obtained by the numerical model. Four locations (Figure 4-1) were selected within the domain considering the following criteria: (i) far from the boundaries, (ii) over the axis of the main formed channels, and (iii) erosion/ accretion phenomena prone (two locations for each) at the end of the numerical simulation. At these locations, the water level, water depth, bed shear stress, depth-averaged velocity, and square depth-averaged velocity were correlated with cumulative

erosion/sedimentation using a Pearson correlation analysis to determine the most adequate variables to be used as input to the AI model (Asuero et al. 2006). It should be noted that the Pearson correlation coefficient indicates the relation between two linear variables. Considering that the hydro-morphodynamic variables do not present linear behaviours, the correlation results were not used to exclude variables with no correlation but to guide the selection of variables to be considered.

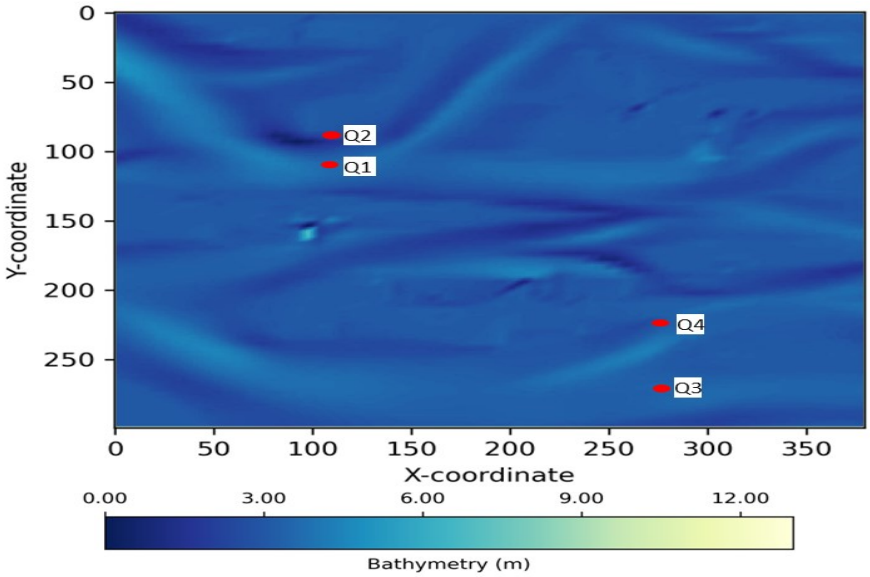


Figure 4-1. Initial bathymetry of the estuary and location of selected points for erosion/sedimentation correlation analysis.

4.2.3. Convolutional Neural Networks

CNNs were designed to work with grid-structured input, like images, which is an adequate characteristic for emulation of 2DH numerical models' results datasets. Each layer of a CNN is a 3D grid with height, width, and depth (or channels) (Aggarwal 2018). Coloured images have three RGB channels, each one representing the intensity of each pixel in the red, green, and blue colour scale, while greyscale images have the same value for all three RGB channels. This way, when using greyscale images as a data source for AI models, it is common to reduce the number of channels to one, decreasing the number of elements in the dataset and optimizing the performance of the model.

Each convolution operation consists of sliding a filter in each position of the image and performing a dot product between the pixel's values and the filter. CNNs also present the property of equivariance to translation, which means that the computed pixel value will be processed in the same way regarding its spatial location (Aggarwal 2018).

The CNN model was implemented in a Spyder environment version 5.2.2 using Python version 3.9.7 and the machine learning system TensorFlow version 2.7.0 (Abadi et al. 2015). This framework was selected because it is an open-source platform that allows the implementation of different machine

learning models, such as artificial neural networks and deep neural networks. The TensorFlow Keras Functional Application Programming Interface (API) was selected due to the requirement of sharing and combining layers in the network architecture, providing more flexibility during the network implementation. The CNN model used two sets of images as input and one set of images as output. This will require more flexibility during the implementation of the model, like, for example, the use of concatenating operations to merge the convolutional layer outputs for posterior processing.

The definition of the hyperparameters of the convolutional layer is essential to achieve satisfactory results with a CNN. Hyperparameters are model parameters of the CNN layers whose values are not estimated from data. For instance, the number of filters is one hyperparameter of each convolution layer that is defined before training the model, while the values of the weights of each filter are defined during the network training (Noriega 2005; Agrawal & Mittal 2020). Finetuning the hyperparameter values is necessary to optimize the network computational performance and improve the accuracy of the model. In this work, five hyperparameters were analysed:

(i) Filters: Each filter captures different information. In each convolution operation, the filter passes along each position of the image and performs a dot product between the data and the filter value (Aggarwal 2018), resulting in a feature map.

(ii) Kernel size: It is the size, in width and height dimensions, of each filter.

(iii) Strides: they are used to control the number of pixels shifts over the input data and to reduce the dimensions of the output. If strides are equal to (2,2), the output size will be half of the input size in the width and height directions. This parameter can be replaced by the use of pooling layers, which also reduces the input dimensions.

(iv) Dilation rate: It is used to increase the range of the filter preserving its size (Yu & Koltun 2015). It is the area of influence of each pixel.

(v) Activation function: It is an activated neuron that can be retained and mapped by nonlinear functions. It is used to increase the performance of the neural network model (Wang et al. 2020). Different activation functions result in completely different outputs. Common activation functions are the hyperbolic tangent (tanh), sigmoid, or Rectified Linear Unit (ReLU). In the CNN model implemented in this study, the tanh and the ReLU were selected as activation functions.

The number of filters and the filter size are the most important hyperparameters of a CNN. Variations in these parameters can strongly affect the performance of the algorithm (Agrawal & Mittal 2020; Ahmed & Karim 2020). Particularly, the filter size effect on the result accuracy is strongly dependent on the characteristics of the data that forces the CNN. It should also be noticed that a higher

number of filters will produce an increase in the model's accuracy but also an increase in the processing power requirements. However, a small number of filters can result in non-representative feature maps, decreasing the performance of the model.

The activation function is also one of the main decisions that must be taken into account in the architecture of a CNN. The selected function is responsible for the addition of nonlinearity to the network. This function must be carefully chosen since it affects the computational cost of the network. It is necessary to select functions that demand low computational resources (Aggarwal 2018). It is worth noting that changes in the activation function can completely modify the output of the model. Wang et al. (2020) studied different activation functions in a facial expression recognition model, demonstrating that the accuracy of the model can vary more than 30% according to the selected function.

AI techniques with convolutional layers were already selected to be used with numerical model solutions. Tu et al. (2021) used the weather research and forecast (WRF) model to generate inputs for a CNN that downscaled hybrid precipitation forecasts, and Kim & Song (2020) implemented a discharge estimator using a CNN with hydrological images as inputs. However, the application of these techniques to solve morphodynamic problems is still scarce and more related to classification problems, such as, for example, the work of Ellenson et al. (2020), who used a CNN to classify the nearshore morphological pattern of beaches using images from cameras. This type of network has the potential to develop complex models using images as inputs, making it possible to model the variables of interest in a whole domain, instead of restricting the results to specific points. Therefore, the implementation of a CNN model to emulate a numerical model becomes a promising choice.

The CNN model implemented in this work has five layers, including one input layer, three hidden layers (branch 1, branch 2, and branch 3), and one output layer (Figure 4-2). The architecture was based on the application of a recursively branched deconvolutional network with three branches, each one used to extract features at different scales (Santhanam et al. 2016). The activation function of the convolutional layers was the ReLU (Fukushima 1975).

The input layer read the data, converted it into images with 380 pixels × 300 pixels, rescaled the values of the pixels between -1 and 1, and performed three convolutional operations. Two of these operations are considered strides equal to two for down-sampling the data, resulting in images with 190 pixels × 150 pixels in the first step and finally with 95 pixels × 75 pixels. The image size reduction was necessary to prevent memory exhaustion during the network training and to amplify the features of the images. The use of strides allows one to rapidly increase the receptive field of each feature and reduce the dimensions of the entire layer (Aggarwal 2018).

The branch 1 and branch 2 layers applied convolution and up-sampling operations in the images with a quarter and a half of the resolution, respectively. The branch 1 inputs were the 95 pixel × 75 pixel images. The branch 2 inputs were the 190 pixel ×150 pixel images and the branch's 1 output. Lastly, the branch 3 input was the branch 2 output and the original images (380 pixels × 300 pixels).

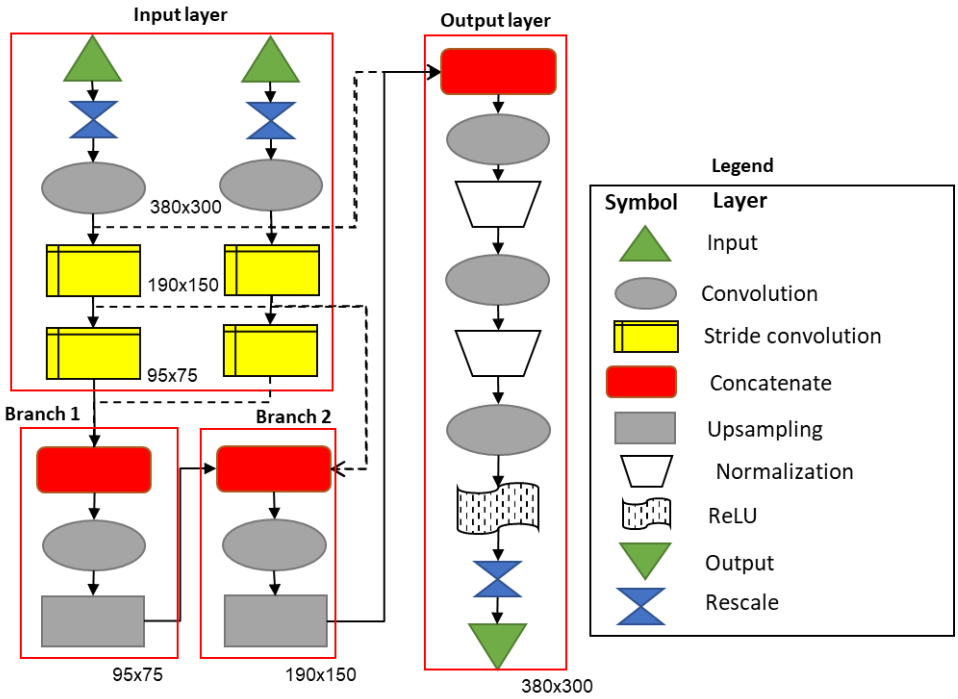


Figure 4-2. Architecture of the CNN.

Finally, the output layer performed three convolution operations, in which the two first convolutions were followed by normalization layers, and the last convolution activation function was the hyperbolic tangent. Also, a ReLU layer, with the maximum output value set to 1, was selected to limit the pixel values between 0 and 1. A rescale operation was performed to set the pixel values between 0 and 255, the same scale presented by the input images. These last two layers become necessary because, in each epoch of the training phase, the algorithm determines the error by comparing the output of the AI model with the images of the training datasets. Therefore, to limit errors, both figures must be on the same scale. The addition of the normalization layers standardized the inputs of the convolution layers, improving the metrics in the training dataset and reducing the error between the emulator and the numerical model.

Different image sizes were tested to avoid memory exhaustion and assess the impact of the input resolution on the performance of the model. Theoretically, larger images present more detailed results, however, they also demand more computational resources during the training phase. Since each cell of the numerical model represents a set of pixels, if the input size is reduced, this will not produce a reduction in the details of the images.

4.2.4. Training and testing

This task consists in defining the values of the weights of each layer through an iterative process. This is performed according to an optimization algorithm responsible for reducing the losses and increasing the accuracy of the output (Vijendra Babu et al. 2020). The optimizer used in this task was the RMSprop, which restricts the vertical oscillations during the search for the values of the weights, reducing the time to converge (Vijendra Babu et al. 2020). A learning rate of 0.9 was used. The mean squared error (MSE) was chosen as the loss function. The root mean squared error (RMSE) was also computed to assess the performance of the CNN in the testing dataset. Also, it was necessary to convert the RMSE from the pixel scale to centimetres according to the following equation:

$$\frac{RMSE_{pix}}{scale_{pix}} = \frac{RMSE_{cm}}{scale_{cm}} \tag{4-1}$$

where the *RMSEpix* is the RMSE in the pixel scale, *scalepix* is the difference between the upper and lower limits of the pixel scale, *RMSEcm* is the RMSE in the centimetres scale, and *scalecm* is the difference between the upper and lower limits of the centimetres scale.

Each epoch during the training took approximately 25 s to be processed, and the maximum number of epochs in the training was 100. The training was developed using parallel computing and the CUDA toolkit v.11.2 (Harish & Narayanan 2007) in a computer with a graphics card Nvidia GeForce GTX 1060 with 6 Gb, a CPU Intel Core i7-8750H 2.20 GHz, 16 GB random access memory (RAM), and the Windows operating system.

4.3. Results and Discussion

4.3.1. Numerical model results

Figure 4-3 presents an example of the images of each variable used as input and output at different time steps of the simulation. The images were exported in greyscale to reduce the number of channels in the CNN and optimize its performance. Each image represents the difference of the considered variables between the current and the previous time step. In the squared depth averaged velocity images, the information of possible erosion (black) and accretion (white) areas during that time instant can be identified associated with higher and lower velocities, respectively. The darker colour in the shear stress represents the areas with the higher values, indicating also where the main erosion channels can be found. In the sedimentation and erosion results, it can be observed the evolution of the channels according to the time step. Grey tones indicate neither erosion or sedimentation in that time step, black colour indicate sedimentation and white colour erosion.

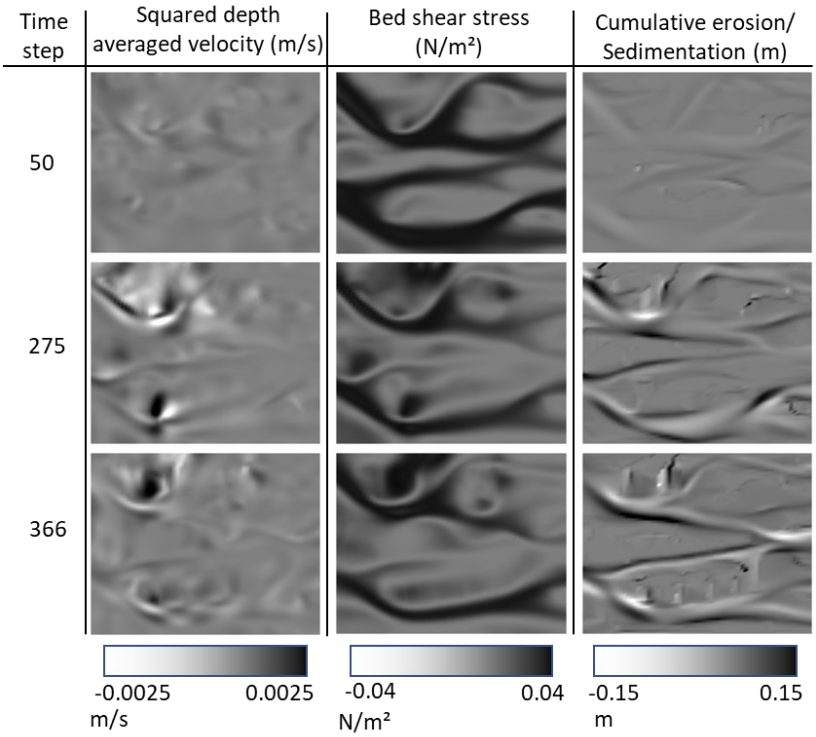


Figure 4-3. Examples of the D3D results images used to train the CNN.

4.3.2. Input variables correlation

Table 1 presents the correlation results between the hydrodynamic variables and the cumulative erosion and sedimentation. These results indicate that a simple correlation is not sufficient to affirm the relevance of the input variables due to the non-linear relation between them. While at point Q1 the best correlation (-0.95) was obtained with water depth, the correlation at point Q2 with the same variable was 0.67. The correlation with the water level for the selected locations that presented accretion processes presented values of 0.57 at Q1 and 0.96 at Q4. But this lack of clear erosion/sedimentation patterns correlation with hydrodynamic variables can be observed not only for water depth or water level but for all the analysed variables.

Table 4-1. Hydrodynamic variables' correlation with cumulative erosion/sedimentation.

Obs. point	Pattern	Water level	Water depth	Magnitude of bed shear stress	Magnitude of depth-averaged velocity	Squared depth-averaged velocity	Cumulative erosion/ sedimentation
Q1	Sedimentation	0.57	-0.95	0.62	0.62	0.62	1
Q2	Erosion	-0.81	-0.99	0.41	0.43	0.41	1
Q3	Erosion	-0.93	0.71	0.72	0.71	0.71	1
Q4	Sedimentation	0.96	0.67	0.95	0.97	0.95	1

Concerning the water depth, the proximity between the points resulted in a similar correlation, regardless of whether the point is eroding or accreting. The other variables presented a pattern with a higher similarity considering the same point. All the correlation values obtained were closer for the velocities and the bed shear stress. This behaviour is clearly explained because the squared depth-averaged velocity and the bottom shear stress are dependent on the depth-averaged velocity. The squared depth-averaged velocity is a linear transformation of the depth-averaged velocity, and the bottom shear stress is proportional to this variable.

After analysing the correlation results presented in Table 1 and considering the physical formulations that describe the sediment transport, the magnitude of bed shear stress and the squared depth-averaged velocity were selected to be tested as input variables.

4.3.3. Training results

To simplify the reference to the convolution layers used in the CNN, they were named according to the layer to which each convolution belongs. Conv n is referred to the input layer, Conv b.n to the convolutions of the branches and Conv 4.n to the convolutions in the output layer, being n the number of the convolution in that layer and b the number of the branch.

The values of the hyperparameters used to implement the CNN were tested to find the combination that produces the best performance and accuracy. It must be stressed that the activation function of the layer Conv 4.3 was the hyperbolical tangent due to the requirement of its results to be in the range of -1 to 1. This will simplify the output values and ensure that, after the ReLU and rescaling layers, all the pixel values are in the interval between 0 and 255.

Furthermore, pooling layers were tested after the convolutional operations to decrease the size of the images in the branches of the network, improving the processing capacity of the model. However, the performance of the model decreased. As the CNN was defined to forecast both points of erosion and sedimentation, the use of a maximum number of pooling layers filtered the information of only one pattern. For this reason, the layers Conv 2 and Conv 3 used 2x2 strides to reduce the input footprint. Also, dropout layers were tested in the place of the normalization layers. However, the results were overestimated when comparing to the performance of the selected architecture.

Different combinations of hyperparameters values were tested. The number of filters in each layer varied from 1 to 16, and the kernel size from (3,3) to (7,7). The combination that resulted in the best model performance and accuracy is presented in Table 4-2.

Table 4-2. Hyperparameters of the convolutional layers.

Layer	Layer Name	Filters	Kernel Size	Strides	Dilation rate	Activation function

Input	Conv 1	10	(5,5)	(1,1)	(3,3)	ReLU
	Conv 2	10	(5,5)	(2,2)	(1,1)	ReLU
	Conv 3	16	(5,5)	(2,2)	(1,1)	ReLU
Branch 1	Conv 1.1	16	(5,5)	(1,1)	(3,3)	ReLU
Branch 2	Conv 2.1	16	(5,5)	(1,1)	(3,3)	ReLU
Branch 3	Conv 3.1	10	(5,5)	(1,1)	(3,3)	ReLU
Output	Conv 4.1	10	(5,5)	(1,1)	(3,3)	ReLU
	Conv 4.2	10	(5,5)	(1,1)	(3,3)	ReLU
	Conv 4.3	1	(5,5)	(1,1)	(3,3)	tanh

Furthermore, it was needed to reduce the resolution of the input images from 1360x1060 pixels to 380x300 pixels to prevent memory exhaustion during the training. It was observed that this reduction slightly decreased the accuracy of the model, meanwhile, higher resolution images than the one selected presented performance problems in the training or testing. The reduction from 1360x1060 pixels to 380x300 pixels represents 1.3276 million fewer pixels per image to be analysed by the network, being the highest resolution value that preserved the proportion between the width and the height of the images without exhausting the memory of the processor. It must be also stressed that each cell in the numerical model grid is represented by a set of pixels. Hence, the reduction in the number of pixels does not affect the resolution of the results when comparing to the numerical model results.

4.3.4. Testing results

A visual comparison between the simulated and the emulated results shows that the main channels are similar in both images (Figure 4-4). The erosion is represented in red colours as positive values and the sedimentation in the blue colours as negative values. However, the CNN tended to predict more sedimentation spots than the numerical model. Despite these differences, the MSE between the numerical model and the surrogate model in the testing dataset (73 samples) was 26.41 cm, and the RMSE was 5.01 cm considering the 0 to 255-pixel scale. Both metrics indicate a reasonable agreement between both models. Considering that a difference of one pixel represents an error of 0.12 cm, the obtained mean RMSE for all time steps is 0.59 cm.

The RMSE for each test sample is presented in Figure 4-5 and the mean RMSE is the mean value of the RMSE considering all the samples. The performance of the emulator is comparable to the results obtained by other authors that apply a similar approach to other water resource problems. Barzegar et al. (2021) achieved a RMSE of 0.04 m with their CNN-LSTM model to forecast water levels in the Michigan Lake, and Wang et al (2022) achieved a 0.25 °C RMSE with a model that forecasted the surface water temperature of a reservoir. However, the output of both models is 1D, while, in this work, the outputs are

2D, being necessary to determine the mean RMSE to compare, in each time step, the differences between the numerical model and the emulator.

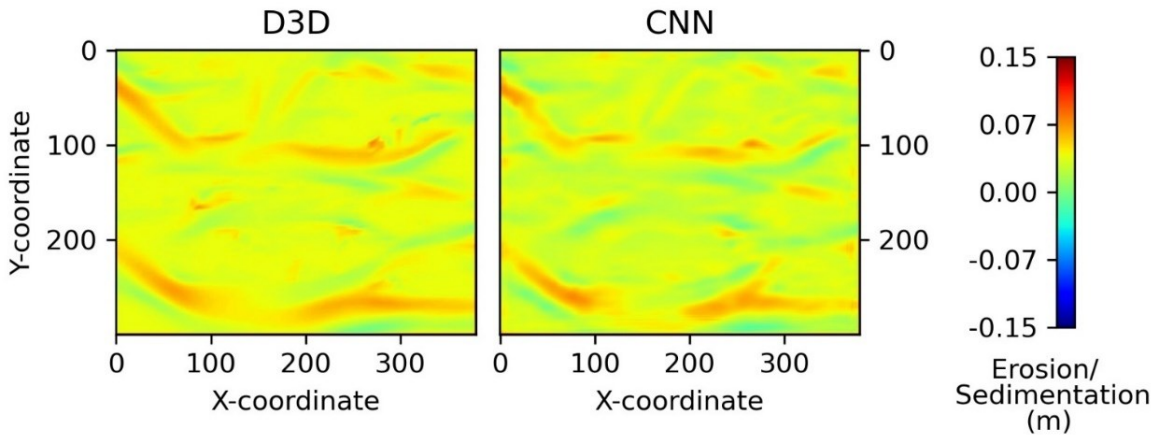


Figure 4-4. Cumulative sedimentation and erosion results of the numerical model (D3D) and the data model (CNN).

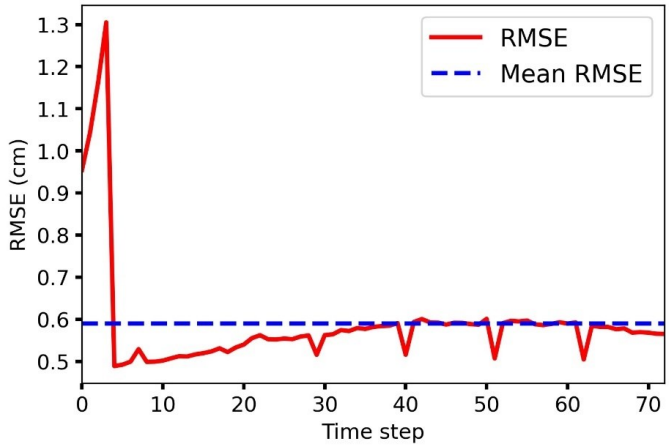


Figure 4-5. RMSE variation in the testing dataset.

The CNN here presented was able to reproduce the erosion not only in the main channels but also in secondary channels, which demonstrates the capacity of the surrogate model in reproducing the erosion/accretion patterns of complex riverine, coastal, and estuarine areas. Though, the numerical model output is smoother than the CNN, which could mean that the numerical model is able to better represent the sediment transport.

Regarding the error of the data model, the computation of the RMSE considers all the pixels of the figures. The maximum RMSE occurred in the third time step, reaching 1.3 cm. It was also observed that the error oscillated between 0.5 cm and 0.6 cm, demonstrating the robustness of the CNN model.

The comparison of the numerical model and the CNN was also performed at the observation points inside the domain during the simulation (Figure 4-6). Q1, Q2 and Q3 presented a good agreement between the numerical model and the CNN results. The higher oscillation in the initial time steps is

associated with a high variation in the input values. In the following time steps, the oscillations are relatively small and the values are closer to the numerical model values. In Q4 the performance of the CNN is almost identical to the one obtained with the numerical model, however, a systematic error that tends to underestimate the numerical model results can be observed.

Regarding the behaviour of the AI model, both Figure 4-4 and Figure 4-6 predictions are much noisier than the results of the numerical model. The emulator forecasted more areas with no sediment transport than the numerical model, represented by the green colours in Figure 4-4. This indicates that some adjustments are necessary to improve its performance. This problem can be probably solved by increasing the training dataset size, which will allow the AI model to better fit its weights and increase its accuracy, or even adding more input variables.

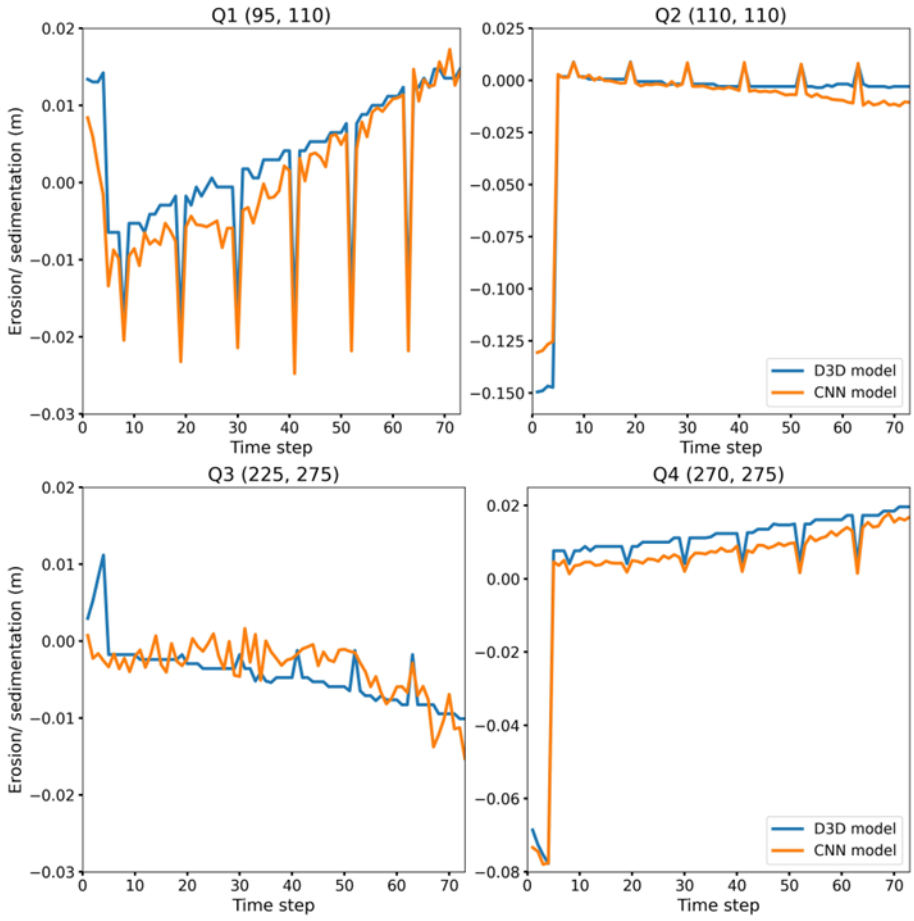


Figure 4-6. Differences in the erosion/sedimentation prediction between models at selected points.

The results of the numerical model presented a noisier behaviour than the CNN results, which represents one of the challenges in the implementation of the emulator. The data used for training and testing presented an oscillation between time steps resulting from the hydro-morphodynamic variables. Hence, it is already expected that the results will not be totally linear. In Q2 it is demonstrated that, when

the oscillation is smaller, the agreement between the models is better. But despite these differences, it is important to highlight that, on average, the results of both models are close.

Despite the good results obtained as revealed by the computed metrics, indicating a satisfactory performance of the emulator, it is important to note that certain aspects must be improved in future works to optimise the application of CNNs to morphodynamic studies.

Firstly, the D3D exported images were represented using the same grey colour scale for all time steps. The scale was defined using the highest and lowest values for each considered variable along the whole numerical simulation, ensuring that the values closer to the limits will be considered in the CNN. However, this large colour scale could be hiding small oscillations. Among the strong changes that the sedimentation patterns have during extreme events, the erosion/accretion processes take place in longer timescales with lower velocities and weaker sediment transport. The consideration of larger colour scales can underestimate sediment transport during long-term simulations. Therefore, it is possible that a change in the limits of the colour bar of the input images, without changing any other variable, can improve the network results. However, it must be highlighted that if the colour scale varies for each time step, the AI model will have difficulties interpreting what an oscillation means in the values of the pixels.

Secondly, the size of the images in the input layer must be carefully considered. If the resolution of the input images is low, the time needed by the network to process all the data is shorter, but the details of the images will be also simplified. Hence, it is important to make a balance between the resolution and the computational performance during the implementation of the CNN. The theoretically minimum resolution would be restricting one pixel per numerical model cell, considering a square image with all the cells aligned. However, this reduction would oversimplify the model, not bringing any advantage in the processing capacity. On the other hand, increasing the resolution of the images can result in memory exhaustion because it will increase the computational costs of the model.

Regarding the performance of the emulator, it was able to simulate 74.5 years in less than 5 seconds (test dataset), while for a numerical model, it took approximately 15 hours to simulate the same period with the same computer. So, it is possible to increase the complexity of the AI model to improve its accuracy without being concerned about its computational performance. It was demonstrated that the use of AI algorithms can be an interesting choice to surrogate numerical models when the processing capacity is limited.

4.3.5. Numerical models versus data models

Before comparing the D3D and CNN models results, it is of utmost importance to understand how each model works. Numerical models solve the conservation and transport equations that describe

the motion of the water and the transport of sediments for a specific domain. They present the advantage of considering the hydrodynamic and morphodynamic physical-based formulations, and when the models are properly calibrated/validated, they can forecast events for a broader range of boundary conditions. However, the calibration is completely dependent on observational field data to ensure the model's accuracy. At the same time, the calibration must be performed for similar conditions than the ones to be forecasted. For example, if the focus is to forecast the effect of floods, the model must be calibrated with data from previous floods. Otherwise, the uncertainties of the numerical model solutions will be higher. Among the data needed for calibration processes, the numerical model also demands information about the boundary conditions to force the model, namely tides, waves, and river flow. In addition, up-to-date bathymetry is essential to build the numerical grid to accurately represent the hydro-morphodynamic patterns and accurately forecast future conditions.

The considerable time reduction obtained in the simulations performed by the emulator can be explained by how each model computes the variables in each time step. The numerical models integrate a set of 3D transport equations both in time and space (grid cells), allowing the interaction between the different hydro-morphodynamic variables and the adjacent cells. The emulator doesn't solve these equations. It multiplies the input matrix by the trained weights of each layer. Due to the complexity of the transport equations, solving these differential equations will require more computational resources than multiplying matrixes. However, the most time-consuming activity to implement an AI model is the training because the determination of the hyperparameters and the necessary epochs to achieve a stable result is a laborious iterative process.

Combining the use of both models can be an interesting choice to diminish the computational costs and maintaining the accuracy of the solutions. The use of data models is computationally more efficient than the numerical models for most of the conditions and, once the model has learned how to interpret new input data, its accuracy will improve, being possible to make more robust data models over time. The numerical models can be useful by generating artificial data to feed the data models with events without measured field data, such as extreme floods, storms or droughts. Surrogate models fed with numerical model data can bring some insight into the unknown conditions, reducing its error when tested in unprecedented data. This idea becomes even more interesting when the data models can learn from images, as remote sensing data and its training dataset are complemented by numerical models' outputs, increasing the diversity of applications of numerical and data models.

4.3.6. Future discussions

While AI algorithms are a powerful tool to study complex phenomena such as long-term morphodynamics, understanding the problems of AI-based models and the limit in their applications is not an easy task. Future works to improve the CNN performance must consider the utilization of a higher number of input images in the training process, the addition of more variables to the input data sets, the addition of more filters, the inclusion of changes in other hyperparameters values, or even the modification of the proposed network architecture, including the selection of long short-term memory layers, which are ideal for models based on time-series datasets.

Particularly, the assessment of the CNN performance considering different input and output resolutions is an interesting focus for future research. The emulator was much faster than the numerical model, and the main variable that affects the emulator's computational time is the resolution of the images. Hence, to assess the potential limits of the proposed methodology, the sensitivity of the technique to the image resolution should be assessed.

AI models' performance can present extremely high accuracy if the expected forecast results are in the range of the data for which the model was trained. However, the reliability of the AI model is reduced when it is extrapolated for other environmental conditions not included in the training phase. In these situations, the data models do not know which output must be associated with the input, since they were not trained for that set of conditions. Despite these limitations, data models are a promising choice for operational forecast platforms because they can be updated as more data of extreme events become available in the forecast system.

4.4. Conclusion

In this work, a deep CNN data model was successfully implemented to emulate the morphology evolution of an estuary. The results of a hydro-morphodynamic numerical model were selected as input of the CNN, effectively combining the capacity of real-time processing of data models with the extrapolation capacity of the numerical models to improve the computational efficiency of the morphology evolution forecast. The CNN was implemented to emulate the bathymetric evolution simulated by the D3D numerical model. The hydrodynamic results of squared averaged depth velocity and bottom shear stress obtained with the numerical model were used as input for the CNN, as well as the erosion and sedimentation variations for output. The images to train and test the surrogate model were carefully generated to maintain the same pattern in size and colour scale.

Obtained results demonstrated the emulator's capacity in reproducing the accretion/erosion patterns in the estuarine channels. However, the emulator was not able to represent all the sediment

transport patterns reproduced by the numerical model. When the performance of the data model was analysed at four specific locations, it can be observed that the data model forecast presented a noisier behaviour than the numerical model. However, the mean RMSE was 0.59 cm, which demonstrates an acceptable performance of the CNN.

The performance of the emulator was widely superior to the numerical model. While the numerical model needed 15 hours to simulate 74.5 years, the CNN needed only 5 seconds. This demonstrates that the use of AI models will allow optimizing computational resources, reducing the operation cost of forecasting platforms. Data models are a possible solution to surrogate high-resolution numerical models, strongly reducing the computational resources needed to process and store the data, making it a possible solution for real-time monitoring systems and operational forecasting.

4.5. References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., ... Zheng, X. (2015). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.

Aggarwal, C. C. (2018). Neural Networks and Deep Learning. In Artificial Intelligence. Springer International Publishing. https://doi.org/10.1007/978-3-319-94463-0

Agoshkov, V. I., Ambrosi, D., Pennati, V., Quarteroni, A., & Saleri, F. (1993). Mathematical and numerical modelling of shallow water flow. Computational Mechanics, 11(5–6), 280–299. https://doi.org/10.1007/BF00350088

Agrawal, A., & Mittal, N. (2020). Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy. The Visual Computer, 36(2), 405–412. https://doi.org/10.1007/s00371-019-01630-9

Ahmed, W. S., & Karim, A. A. A. (2020). The Impact of Filter Size and Number of Filters on Classification Accuracy in CNN. 2020 International Conference on Computer Science and Software Engineering (CSASE), 88–93. https://doi.org/10.1109/CSASE48920.2020.9142089

Akan, A. O. (2006). Introduction to unsteady open-channel flow. In Open Channel Hydraulics (Vol. 1, pp. 315–359). Elsevier. https://doi.org/10.1016/B978-075066857-6/50009-6

Asuero, A. G., Sayago, A., & González, A. G. (2006). The Correlation Coefficient: An Overview. Critical Reviews in Analytical Chemistry, 36(1), 41–59. https://doi.org/10.1080/10408340500526766

Baek, S.-S., Pyo, J., & Chun, J. A. (2020). Prediction of Water Level and Water Quality Using a CNN-LSTM Combined Deep Learning Approach. Water, 12(12), 3399. https://doi.org/10.3390/w12123399

Barzegar, R., Aalami, M. T., & Adamowski, J. (2021). Coupling a hybrid CNN-LSTM deep learning model with a Boundary Corrected Maximal Overlap Discrete Wavelet Transform for multiscale Lake water level forecasting. Journal of Hydrology, 598(March), 126196. https://doi.org/10.1016/j.jhydrol.2021.126196

Barzegar, R., Moghaddam, A. A., & Baghban, H. (2016). A supervised committee machine artificial intelligent for improving DRASTIC method to assess groundwater contamination risk: a case study from Tabriz plain aquifer, Iran. Stochastic Environmental Research and Risk Assessment, 30(3), 883–899. https://doi.org/10.1007/s00477-015-1088-3

Chen, C., Hui, Q., Xie, W., Wan, S., Zhou, Y., & Pei, Q. (2021). Convolutional Neural Networks for forecasting flood process in Internet-of-Things enabled smart city. Computer Networks, 186(June 2020), 107744. https://doi.org/10.1016/j.comnet.2020.107744

Chen, Roy, S. B., & Hutton, P. H. (2018). Emulation of a process-based estuarine hydrodynamic model. Hydrological Sciences Journal, 63(5), 783–802. https://doi.org/10.1080/02626667.2018.1447112

Church, M., & Ferguson, R. I. (2015). Morphodynamics: Rivers beyond steady state. Water Resources Research, 51(4), 1883–1897. https://doi.org/10.1002/2014WR016862

Deltares. (2018). Delft3D-Flow: Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments - User Manual.

Ellenson, A. N., Simmons, J. A., Wilson, G. W., Hesser, T. J., & Splinter, K. D. (2020). Beach state recognition using argus imagery and convolutional neural networks. Remote Sensing, 12(23), 1–20. https://doi.org/10.3390/rs12233953

Elliott, M., Day, J. W., Ramachandran, R., & Wolanski, E. (2019). A Synthesis: What Is the Future for Coasts, Estuaries, Deltas and Other Transitional Habitats in 2050 and Beyond? In Coasts and Estuaries (pp. 1–28). Elsevier. https://doi.org/10.1016/B978-0-12-814003-1.00001-0

Elmilady, H., van der Wegen, M., Roelvink, D., & van der SpModelling2022). Modeling the Morphodynamic Response of Estuarine Intertidal Shoals to Sea-Level Rise. Journal of Geophysical Research: Earth Surface, 127(1), 1–26. https://doi.org/10.1029/2021JF006152

Fukushima, K. (1975). Cognitron: A self-organizing multilayered neural network. Biological Cybernetics, 20(3–4), 121–136. https://doi.org/10.1007/BF00342633

Goldstein, E. B., & Coco, G. (2014). A machine learning approach for the prediction of settling velocity. Water Resources Research, 50(4), 3595–3601. https://doi.org/10.1002/2013WR015116

Guerrero, M., Di Federico, V., & Lamberti, A. (2013). Calibration of a 2-D morphodynamic model using water–sediment flux maps derived from an ADCP recording. Journal of Hydroinformatics, 15(3), 813–828. https://doi.org/10.2166/hydro.2012.126

Gupta, D., Hazarika, B. B., Berlin, M., Sharma, U. M., & Mishra, K. (2021). Artificial intelligence for suspended sediment load prediction: a review. Environmental Earth Sciences, 80(9), 346. https://doi.org/10.1007/s12665-021-09625-3

Harish, P., & Narayanan, P. J. (2007). Accelerating Large Graph Algorithms on the GPU Using CUDA. In High Performance Computing – HiPC 2007: Vol. 4873 LNCS (pp. 197–208). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77220-0_21

Iglesias, I., Venâncio, S., Pinho, J. L., Avilez-Valente, P., & Vieira, J. M. P. (2019). Two Models Solutions for the Douro Estuary: Flood Risk Assessment and Breakwater Effects. Estuaries and Coasts, 42(2), 348–364. https://doi.org/10.1007/s12237-018-0477-5

Jain, S. K. (2008). Development of Integrated Discharge and Sediment Rating Relation Using a Compound Neural Network. Journal of Hydrologic Engineering, 13(3), 124–131. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:3(124)

Kandolf, G. M., & Piégay, H. (2003). Tools in Fluvial Geomorphology (G. M. Kondolf & H. Piégay, Eds.; Second). Wiley. https://doi.org/10.1002/0470868333

Kim, D. Y., & Song, C. M. (2020). Developing a Discharge Estimation Model for Ungauged Watershed Using CNN and Hydrological Image. Water, 12(12), 3534. https://doi.org/10.3390/w12123534

Lewis, N. S., Young, D. R., Folger, C. L., & DeWitt, T. H. (2021). Assessing the Relative Importance of Estuarine Nursery Habitats—a Dungeness Crab (Cancer magister) Case Study. Estuaries and Coasts, 44(4), 1062–1073. https://doi.org/10.1007/s12237-020-00821-1

Li, J., Bare, B., Zhou, S., Yan, B., & Li, K. (2021). Organ-Branched CNN for Robust Face Super-Resolution. 2021 IEEE International Conference on Multimedia and Expo (ICME), 1–6. https://doi.org/10.1109/ICME51207.2021.9428152

Nagy, H. M., Watanabe, K., & Hirano, M. (2002). Prediction of Sediment Load Concentration in Rivers using Artificial Neural Network Model. Journal of Hydraulic Engineering, 128(6), 588–595. https://doi.org/10.1061/(ASCE)0733-9429(2002)128:6(588)

Noriega, L. (2005). Multilayer perceptron tutorial. School of Computing. Staffordshire University, 1–12.

Palani, S., Liong, S. Y., & Tkalich, P. (2008). An ANN application for water quality forecasting. Marine Pollution Bulletin, 56(9), 1586–1597. https://doi.org/10.1016/j.marpolbul.2008.05.021

Parker, K., Ruggiero, P., Serafin, K. A., & Hill, D. F. (2019). Emulation as an approach for rmodellinguarine modeling. Coastal Engineering, 150(March), 79–93. https://doi.org/10.1016/j.coastaleng.2019.03.004

Pearson, R. G., Connolly, N. M., Davis, A. M., & Brodie, J. E. (2021). Fresh waters and estuaries of the Great Barrier Reef catchment: Effects and management of anthropogenic disturbance on biodiversity, ecology and connectivity. Marine Pollution Bulletin, 166(February), 112194. https://doi.org/10.1016/j.marpolbul.2021.112194

Pinho, J. L. S., Vieira, L., Vieira, J. M. P., Venâncio, S., Simões, N. E., Sá Marques, J. A., & Santos, F. S. (2020). Assessing causes and associated water levels for an urban flood using hydroinformatic tools. Journal of Hydroinformatics, 22(1), 61–76. https://doi.org/10.2166/hydro.2019.019

Poelhekke, L., Jäger, W. S., van Dongeren, A., Plomaritis, T. A., McCall, R., & Ferreira, Ó. (2016). Predicting coastal hazards for sandy coasts with a Bayesian Network. Coastal Engineering, 118, 21–34. https://doi.org/10.1016/j.coastaleng.2016.08.011

Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., K. Mintenbeck, A., Alegría, M., Craig, S., Langsdorf, S., Löschke, V., Möller, Okem, A., & Rama, B. (2022). Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. Cambridge, UK and New York, NY, USA. https://doi.org/10.1017/9781009325844.

Roering, J. J., Kirchner, J. W., & Dietrich, W. E. (2001). Hillslope evolution by nonlinear, slope-dependent transport: Steady state morphology and equilibrium adjustment timescales. Journal of Geophysical Research: Solid Earth, 106(B8), 16499–16513. https://doi.org/10.1029/2001JB000323

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 9351, Issue Cvd, pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Santhanam, V., Morariu, V. I., & Davis, L. S. (2016). Generalized Deep Image to Image Regression. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua, 5395–5405. https://doi.org/10.1109/CVPR.2017.573

Soloy, A., Turki, I., Lecoq, N., Gutiérrez Barceló, Á. D., Costa, S., Laignel, B., Bazin, B., Soufflet, Y., Le Louargant, L., & Maquaire, O. (2021). A fully automated method for monitoring the intertidal topography using Video Monitoring Systems. Coastal Engineering, 167(May 2020). https://doi.org/10.1016/j.coastaleng.2021.103894

Taylor, M. D., & Suthers, I. M. (2021). The Socio-ecological System of Urban Fisheries in Estuaries. Estuaries and Coasts, 44(7), 1744–1751. https://doi.org/10.1007/s12237-021-00916-3

Tu, T., Ishida, K., Ercan, A., Kiyama, M., Amagasaki, M., & Zhao, T. (2021). Hybrid precipitation downscaling over coastal watersheds in Japan using WRF and CNN. Journal of Hydrology: Regional Studies, 37(May), 100921. https://doi.org/10.1016/j.ejrh.2021.100921

Van, S. P., Le, H. M., Thanh, D. V., Dang, T. D., Loc, H. H., & Anh, D. T. (2020). Deep learning convolutional neural network in rainfall-runoff modelling. Journal of Hydroinformatics, 22(3), 541–561. https://doi.org/10.2166/hydro.2020.095

Vijendra Babu, D., Karthikeyan, C., Shreya, & Kumar, A. (2020). Performance Analysis of Cost and Accuracy for Whale Swarm and RMSprop Optimizer. IOP Conference Series: Materials Science and Engineering, 993(1), 012080. https://doi.org/10.1088/1757-899X/993/1/012080

Wang, L., Xu, B., Zhang, C., Fu, G., Chen, X., Zheng, Y., & Zhang, J. (2022). Surface water temperature prediction in large-deep reservoirs using a long short-term memory model. Ecological Indicators, 134(October 2021), 108491. https://doi.org/10.1016/j.ecolind.2021.108491

Wang, Y., Li, Y., Song, Y., & Rong, X. (2020). The Influence of the Activation Function in a Convolution Neural Network Model of Facial Expression Recognition. Applied Sciences, 10(5), 1897. https://doi.org/10.3390/app10051897

Xie, D., Schwarz, C., Kleinhans, M. G., Zhou, Z., & Maanen, B. (2022). Implications of Coastal Conditions and Sea-Level Rise on Mangrove Vulnerability: A Bio-Morphodynamic Modeling Study. Journal of Geophysical Research: Earth Surface, 127(3), 1–28. https://doi.org/10.1029/2021jf006301

Yu, F., & Koltun, V. (2015). Multi-Scale Context Aggregation by Dilated Convolutions. 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings.

Zhou, X., Ma, W., Echizenya, W., & Yamazaki, D. (2021). The uncertainty of flood frequency analyses in hydrodynamic model simulations. Natural Hazards and Earth System Sciences, 21(3), 1071–1085. https://doi.org/10.5194/nhess-21-1071-2021

CHAPTER 5.

Coastal morphodynamic emulator for early warning short-term forecasts.

Article published in May of 2023 in the Journal of Environmental Modelling & Software.

Weber de Melo, W., Pinho, J., and Iglesias, I. (2023). Coastal morphodynamic emulator for early warning short-term forecasts. Environmental Modelling & Software, 165, 105729.

"Esse mundo que aí está foi feito por nós, portanto, pode ser por nós reinventado."

Mário Sérgio Cortella (1954 –) – Brazilian philosopher and writer.

Chapter 5. Coastal morphodynamic emulator for early warning short-term forecasts.

Abstract

The use of numerical models to anticipate the effects of floods and storms in coastal regions is essential to mitigate the damages of these natural disasters. However, local studies require high spatial and temporal resolution numerical models, limiting their use due to the involved high computational costs. This constraint becomes even more critical when these models are used for real-time monitoring and warning systems. Therefore, the objective of this paper was to reduce the computational time of coastal morphodynamic models' simulations by implementing a deep learning emulator. The emulator performance was evaluated using different scenarios run with the XBeach software, which considered different grid resolutions and the effects of a storm event in the morphodynamic patterns around a breakwater and a groin. The morphodynamic simulation time was reduced by 23%, and it was identified that the major restriction to reducing the computational cost was the hydrodynamic numerical model simulation.

5.1. Introduction

Forecasting the impacts of extreme events, such as droughts, floods, and storms, is fundamental to identify threatened areas and to allow early actions to increase resilience and reduce the eventual related damages. Anticipating extreme events becomes even more relevant in low-lying coastal areas, which present high risks due to the combination of meteo-oceanic events, climate trends and human activities. In open ocean coastal areas, cities and other settlements are more vulnerable to coastal flooding, whilst pluvial floods are one of the major threats in deltas and estuaries (Pörtner et al., 2022). Moreover, these areas shelter approximately 11% of the global population, which lives within 10 m above mean sea level (Haasnoot et al., 2021; Pörtner et al., 2022; Vousdoukas et al., 2020).

To anticipate and mitigate extreme events impacts, the implementation of forecasting systems must be addressed. These systems are usually implemented based on numerical models, such as XBeach and D3D (Deltares, 2018; Roelvink et al., 2009), which had demonstrated to be accurate in predicting HMD variables under extreme events scenarios (Ferreira et al., 2019; Iglesias et al., 2022; Simmons & Splinter, 2022; Vousdoukas et al., 2012).

However, these numerical models demand considerable simulation computational time, mainly when applied to high spatial and/or temporal resolution domains and when different modules are used in a forecast simulation (e.g., hydrodynamics, waves, winds, sediments, and morphology). However, the use of high-resolution models is crucial to correctly simulate hydrodynamic and morphodynamic patterns at coastal stretches. Moreover, real-time monitoring and warning systems require computationally

efficient numerical models to optimize the prediction of future trends and to allow a faster response by competent authorities. To improve the computing efficiency of numerical models like XBeach, Rautenbach et al (2022) compared the performance of the software in hydrodynamics and wave dynamics simulations using a CPU and a GPU, finding that even a desktop grade GPU can compete with the computational efficiency of high-performance computing CPU facilities. However, there are still few studies about the computational efficiency of HMD numerical models and ways to improve it.

An alternative to reduce the computational costs of HMD numerical models is the use of emulators, which are based on the application of statistical techniques that surrogate the numerical model. These are artificial intelligence-derived methods, namely ML and DL, which implementation for surrogating morphodynamic numerical models had been approached by other authors. Poelhekke et al. (2016) developed a XBeach emulator to predict hazards at Faro beach, Anderson et al. (2021) used several numerical models and a Gaussian Process Regression (GPR) emulator to predict wave runup and overwash depths, and Gharagozlou et al. (2022) developed a GPR-based emulator for post-storm subaerial beach and dune profile shapes. However, such works developed emulators for specific punctual locations within the domain. Hence, this work intends to fill this gap solving the emulators constraint in morphodynamic variables forecasting at full 2DH domains.

In a previous work of the authors, a DL model with convolutional layers, was implemented using simulation results obtained with D3D software (Weber de Melo et al (2022). The developed technique used images of a D3D numerical model hydrodynamic variables as input and was able to forecast the long-term morphodynamic evolution of an intertidal shoal estuary, demonstrating that it was possible to reduce the total simulation time by surrogating the morphodynamic module of the numerical model for an emulator. However, short-term morphodynamics in beaches are more complex than long-term estuarine morphodynamics. Estuarine morphodynamics main drivers are the water currents, which varies according to tides and freshwater flows. On the contrary, beach morphodynamics results from the complex interactions between winds, non-linear waves processes, tides, water currents and return flows, and the processes involved in coastal morphodynamics are dominated by complex onshore and offshore transport. Thus, the numerical model selected in Weber de Melo et al (2022) is not well-suited to model sand beaches bar behaviour because the processes of dune erosion and offshore sediment transport by the return flow are not included in the numerical code. On the other hand, XBeach software has all necessary processes to model bar behaviour (Trouw et al., 2012). These authors also note that the default values for wave-related bed load and suspended load factors in D3D are too high and, this way, the physics of the model are likely flawed.

In this context, this paper has the objective to improve the DL model architecture by Weber de Melo et al. (2022) and test if this improved architecture could reproduce the morphodynamic results of an experiment run with the XBeach model considering an extreme event scenario. The objective is to demonstrate if the new emulator, that successfully surrogated the D3D morphodynamic module for an estuarine environment, could also be applied to XBeach simulations of non-linear wave dominated coastal stretches. Two simulation sets were used to evaluate the XBeach emulator. The first set used a simplified linear bathymetry to assess if a change in the up-sampling operation in the neural network could reduce the error in the validation dataset and to determine the total simulation time reduction. The second simulation set tested if the emulator could reproduce the morphological evolution of a more complex coastal stretch involving a sandy beach with different coastal defence structures, including a submerged breakwater and a groin.

5.2. Methodology

The methodology applied in this study was adapted and improved from the previous work of Weber de Melo et al. (2022), which designed and implemented an HMD numerical model emulator to forecast the long-term morphological evolution of an estuarine area. However, as an already mentioned in the introduction, the processes that affect the morphodynamics in estuaries are different from those that occur in coastal areas. In the first case, the main morphological evolution driver is the water currents generated by river flows and tides. In the second case, the waves, tides and winds are the main drivers that affect the coastal morphological evolution. Due to the complexity of the morphological processes in coastal regions and the need to introduce other non-linear variables besides the water current, this work developed a new methodology to reduce the computational costs of a XBeach model to emulate the short-term morphological evolution of coastal stretches.

5.2.1. XBeach model

XBeach is an open-source numerical model that simulates HMD processes in sandy coastal areas by solving 2DH equations for wave propagation, flow, sediment transport and bottom changes (Roelvink et al., 2009). This software solves the shallow water equation for low-frequency waves and average flows considering the depth-averaged generalized lagrangian mean formulation (Roelvink et al., 2009). The sediment transport equation was solved accordingly to the Van-Thiel-Van Rijn formulation (van Rijn, 2007; van Thiel de Vries, 2009).

XBeach-based methodologies had already been demonstrated to be capable of achieving reliable results, having been applied to study the uncertainties in run-up predictions on natural beaches (e.g., Rutten et al., 2021), to predict extreme offshore directed sediment transport (Suzuki & Cox, 2021) and to model the morphological response of a sandy barrier island during hurricane conditions (Smallegan et

al., 2016). However, XBeach simulations require intensive computational resources, mainly when the domain has a high spatial resolution, restraining its application in real-time early warning forecasting and decision support systems (Ferreira et al., 2019; Gharagozlou et al., 2022; Poelhekke et al., 2016).

In this study, two different simulation sets were defined to evaluate and optimize the performance of the emulator. The spatial domain of the first set was a simplified beach with a linear bathymetry (Figure 5-1a). The focus of this first set was to assess the up-sampling operation in the architecture of the emulator to reduce the error in the forecasts. In the second set, the capacity of the emulator in reproducing the morphodynamics was evaluated in a more complex geometrical domain involving two different coastal defence structures, a submerged breakwater (Figure 5-1b) and a groin (Figure 5-1c).

5.2.1.1. Coastal domain characteristics

Three different coastal domains were defined to analyse the performance of the emulator (Figure 5-1). The first domain consists of a beach with a 2 km cross shore width and 4 km alongshore length discretized with a 100x200 grid with 20 m cells size resolution. The bathymetry varied from 0 m to -20 m and the topography from 0 m to 10 m. The topo-bathymetry is homogeneous in the north-south direction. In the west-east direction, it varied linearly (Figure 5-1 - a).

The second and third domains used a 335x375 grid with 5 m cells size resolution, resulting in a coastal domain of 1.68 km cross-shore width and 1.88 km alongshore length. The bathymetry varied between 0 m to -12 m at the offshore boundary, and the topography from 0 m to 2 m at the inland boundary (Figure 5-1 - b and c).

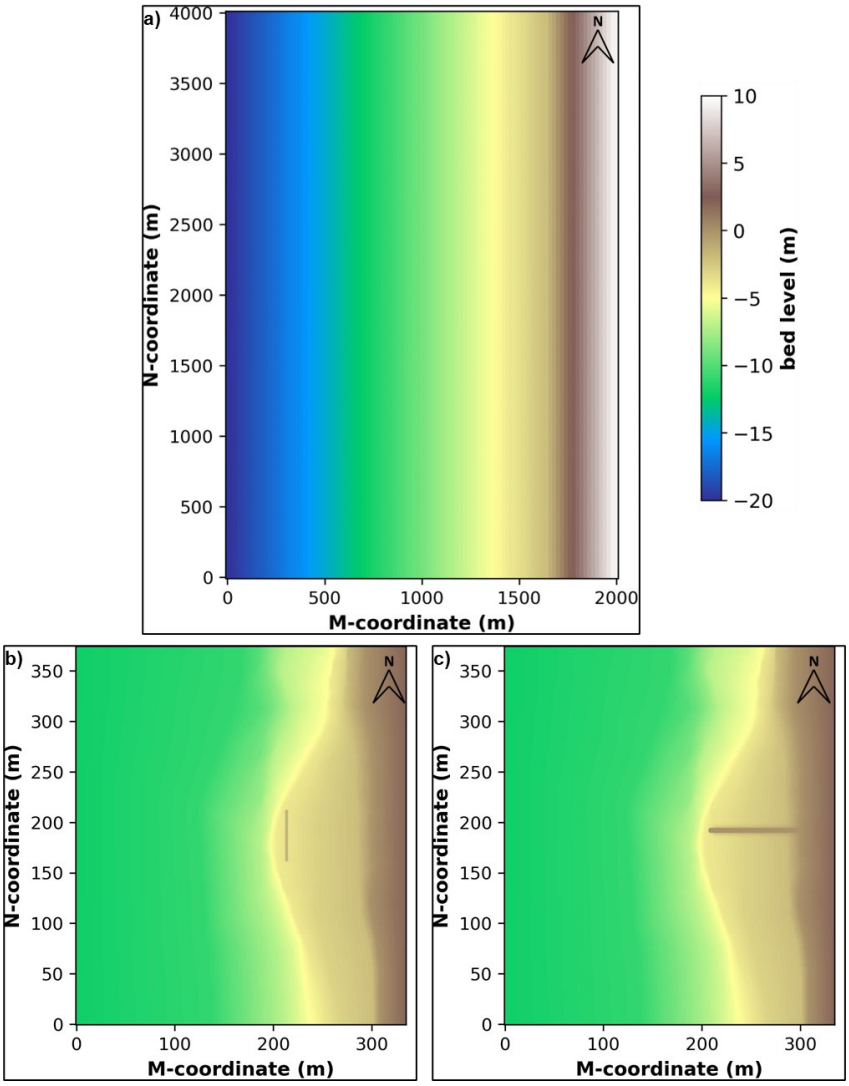


Figure 5-1. Topo-bathymetry considered for each XBeach domains: a) Linear topo-bathymetry domain; b) Breakwater domain; c) Groin domain.

5.2.1.2. Numerical model initial and boundary conditions

The XBeach simulations run in the surfbeat mode, in which the short-wave motion is solved using the wave action equation (Roelvink et al., 2009). For the first set of simulations that used the linear bathymetry, the initial water level was assumed equal to 1.6 m at the offshore boundary and 1.3 m at the landward boundary, while in the second set of simulations, which included domains with coastal defence structures, the initial water level was assumed equal to the mean sea level (0 m MSL).

The Joint North Sea Wave Project (JONSWAP) spectral wave model was selected to force the models to represent the occurrence of an extreme event. The significant wave height used in simulation set 1 was 6 m, intending to simulate a highly energetic storm to force intense morphological changes during the simulation. The significant wave height (Hs) was 3 m for the simulation set 2, which is a representative value, based on the historical wave measurements, of a typical storm in Portugal (Oliveira et al., 2020). For this geographical location, the average peak periods (Tp) vary between 5 s to 12 s and

the mean Hs usually varies between 1.5 m to 2 m, but values between 3 m to 6 m are recurrent (Vieira et al., 2020; Viitak et al., 2021).

The hydrodynamic simulation period was 6 h for both simulation sets, although the sediment transport and morphology modules started with half hour lag in the simulation set 1 and with one-day lag in simulation set 2 to avoid numerical inconsistencies. Additionally, a morphological acceleration factor (MORFAC) of 20 was adopted in simulation set 2 to represent the morphological evolution along 5 days. Furthermore, the adopted critical avalanching slope above water was 0.15 and underwater was 1. The summary of the boundary conditions used for each simulation set are presented in Table 5-1, where the *mainang* parameter is the main wave angle in the nautical convention, s is the directional spreading coefficient, *gammajsp* is the peak enhancement factor and fnyq is the highest frequency of the JONSWAP spectrum.

Table 5-1. Domain characteristics and boundary conditions of simulation sets.

Simulation set	1	2
Objective	Evaluate up-sampling operation and simulation computational time	Evaluate the emulator performance with different scenarios
Domain	2 km x 4 km grid with 20 m resolution	1.68 km x 1.88 km grid with 5 m resolution
Bathymetry	Linear bathymetry	Simplified typical northern Portugal coastal bathymetry
Morfac	1	20
Boundary conditions (JONSWAP wave model)		
Hs (m)	6	3
Tp (s)	15	15
mainang (°)	300	315
s	20	20
gammajsp	3.3	3.3
fnyq (Hz)	1	1

With these conditions, the total simulation time was 28 min for the first domain, 65 min for the groin domain and 67 min for the breakwater domain, using a 6 cores processor with a 2.2 GHz base frequency.

5.2.1.3. Simulation Scenarios

As already briefly explained in the introduction, each simulation set was elaborated to accomplish different objectives. The first set had the goal to study up sampling operation to improve the performance of the emulator. The second set assessed the capacity of the DL model to emulate more complex coastal environments under the influence of coastal defence structures and assessed different combinations of inputs and outputs during the training of the network. Scenarios 1 to 4 in simulation set 2 assessed the

capacity of the emulator in reproducing different morphodynamic variables, the current time step bed level change (CTBLC) or the cumulative erosion and sedimentation (CES). Scenarios 5 to 8 were defined to evaluate the possibility of implementing a single emulator for different domains. The models were trained using the data of both domains, but its performance was assessed considering the results of each domain individually.

The characteristics of the scenarios are synthesized in Table 2. The simulation set 1 presents the up-sampling functions that were tested. ConvT refers to the transpose convolution layer, followed by the activation function that was used in that layer. At the simulation set 2, the acronyms refer to the training and testing conditions. The first acronym refers to the domains of the data used to train the network. BW is for breakwater, G is for groin, and BWG for both domains. The second acronym refers to the XBeach variable used as output by the emulator (CES or CTBLC). The third acronym, presented in scenarios 5 to 8, refers to the domain wherein results were used to assess the performance of the emulator. This last acronym was only necessary to differentiate the validating conditions of the models that were trained with the data of both domains; hence, it was not necessary in scenarios 1 to 4.

Table 5-2. Testing conditions of the DL models. In simulation set 1, the up-sampling layer was evaluated whilst simulation set 2 assessed different combinations of inputs and outputs.

Simulation set		**1**	**2**
Scenarios	1	Up-sampling layer	BW - CES
	2	ConvT, ReLU	BW - CTBLC
	3	ConvT, tanh	G - CES
	4	-	G - CTBLC
	5	-	BWG - CES - BW
	6	-	BWG - CTBLC - BW
	7	-	BWG - CES -G
	8	-	BWG - CTBLC - G

Additionally, the sensitivity of the emulator to the numerical model grid resolution was assessed. For this test, the set 1 was selected using the same model domain and the same topo-bathymetry but with two different grid resolutions: 20 m x 20 m grid resolution (original domain) and 40 m x 40 m grid resolution (a quarter of the original resolution). These models, which will receive the acronyms OR, for the highest resolution domain, and LR, for the lowest resolution domain, were used to assess if the emulator could learn using the coarser resolution hydrodynamic input to predict the morphodynamic output of the finer domain without losing accuracy.

5.2.2. Implementation of the morphodynamic emulator

The DL model was implemented in the Tensor Flow framework (Abadi et al., 2016), in Python programming language version 3.9.7 using a Spyder environment version 5.2.2. Details about the implementation of the emulator are presented in the following sub-sections.

5.2.2.1. Input and output images processing

The DL network used the Generalized Lagrangian Mean velocity (GLMV) magnitude and the bottom shear stress (BSS) as inputs. The output in simulation set 1 was CTBLC. In simulation set 2, the CES result was additionally used. To generate the image datasets with these variables, a Python script was programmed to ensure that all the images presented the same resolution and colour scale characteristics. The axes of the images were turned off and the margins were set to 0 to ensure that all the pixels of the images coincided with the numerical model domain.

The netcdf4 library (Whitaker et al., 2020) was used to read the XBeach results and the Matplotlib library was used to plot the results (Hunter, 2007). The images were created with the imshow function and were exported with 300 dpi, resulting in a resolution of 452x900 pixels for simulation set 1 and 540x600 pixels for simulation set 2. The images were plotted in grayscale, and the limit values of the colour scales were determined using a programmed routine that iterated through all the variables' datasets to find the maximum and minimum values for each variable at all run time steps. Brighter tones indicate positive values (sedimentation) while darker tones indicate negative values (erosion) or zero, in the case of GLMV and BSS. Examples of these images are presented in Figure 5-2.

The exporting loop was divided into 3 parts: firstly, it read the results of the ith time step and created an image with the pre-defined configurations. Secondly, it updated the image name according to the respective time step, to avoid overwriting the results. Thirdly, the image was exported to a folder in which all the images were stored according to each variable. It resulted in 54 and 150 images per variable for the simulation set 1 and 2, respectively. The scenario that evaluated the emulator with different input and output grid resolutions used 109 images per variable, maintaining the same training/test ratio. This increase in the number of images was necessary to reduce the emulator mean error.

For the simulation set 1, only the results of the CTBLC were used to evaluate the performance of the model. For the set 2, the CES results were also used to train the network and to evaluate if the emulator could properly forecast the morphological evolution of the domains. The inputs were the same in this second case, the unique difference was the output variable used for training the network.

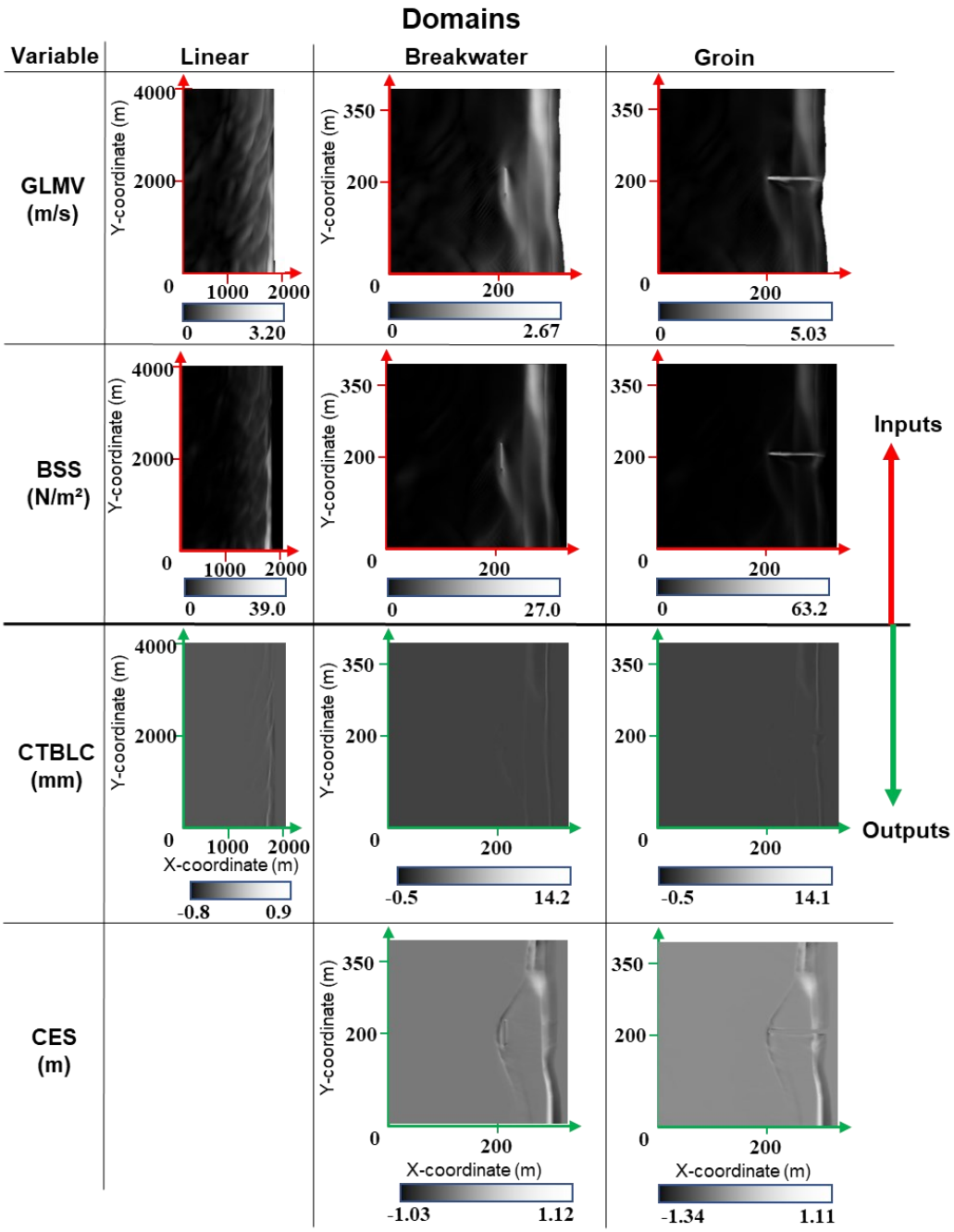


Figure 5-2. Examples of images of the hydrodynamic model results used as input (GLMV and BSS) and of the morphodynamic model results used as output (CTBLC and CES) by the deep learning model. The bound values used to plot the images are in the grayscale colour bar above each image. The DL models used the first two variables as input and at least one of the last two variables as output.

5.2.2.2. Network architecture

A U-net and a recursively deconvolutional branched network are the basis for the emulator architecture, in which each branch mapped the image features of one specific resolution (Ronneberger et al., 2015; Santhanam et al., 2016). This configuration, which was also used (by Weber de Melo et al.,

(2022), is presented in Figure 5-3. The hyperparameters, namely the activation functions, number of filters, kernel size, strides and dilation rate of the convolutional layers remained the same.

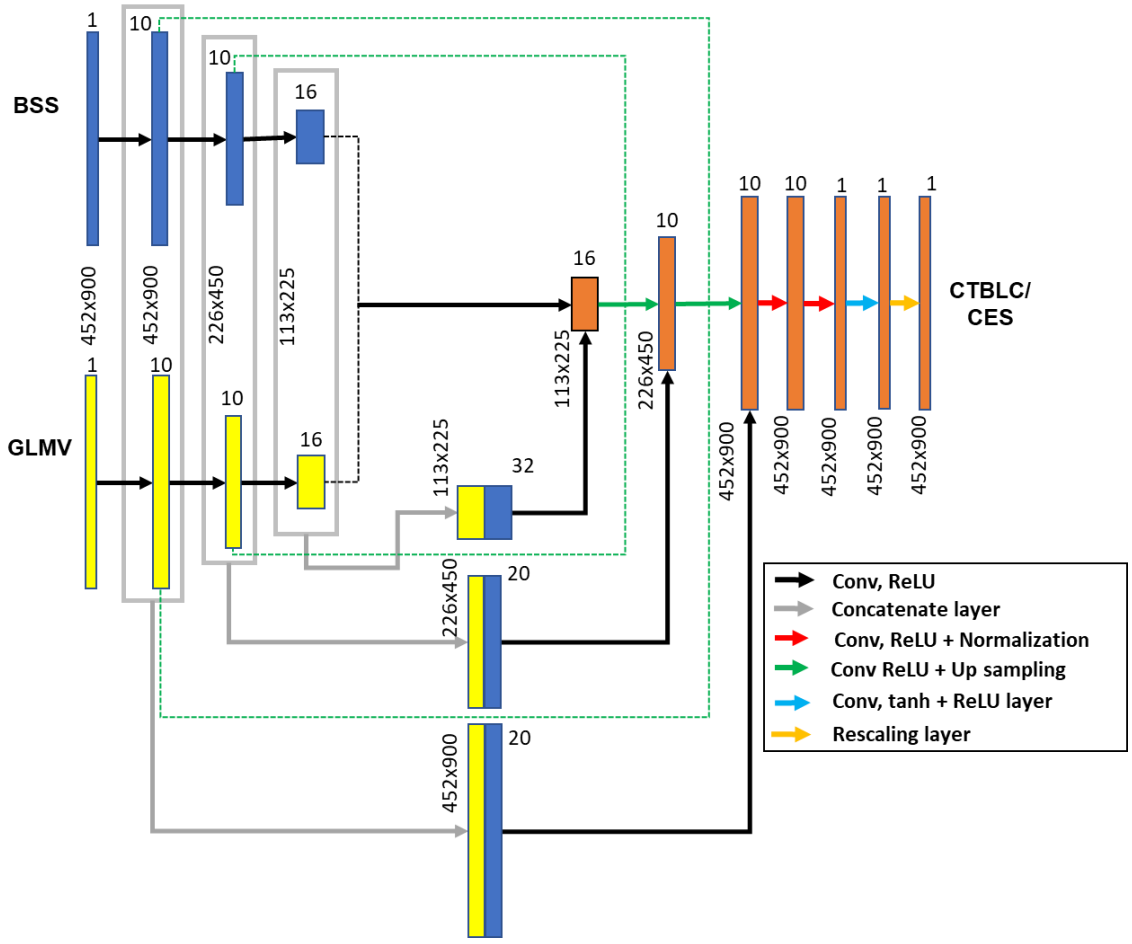


Figure 5-3. DL model architecture. The rectangles represent the resolution of the image in each network layer. The numbers above each layer are the number of filters, and each coloured arrow represents one different operation performed in the network.

The DL model was implemented using the Functional API of the TensorFlow framework, due to the complexity of the chosen architecture. The network initially mapped the features of the input data and reduced the footprint of the data using convolutional operations with strides set to 2, to half the resolution of the layer inputs. After each stride convolution, a concatenation layer merged the layers' outputs. After that, up-sampling layers restored the original resolution of the input data, and 3 other convolutions were performed. The activation function was the ReLU in all convolutional layers except for the last, which used a hyperbolic tangent function (tanh).

Besides, a deconvolution layer, also known as transposed convolution (Dumoulin & Visin, 2016), replaced the up-sampling operation (green arrows in Figure 3). This change increases the number of trainable parameters in the architecture of the network, allowing the improvement of the emulator performance. The ReLU and tanh activation functions were assessed during this step.

The main difference between these two operations is how they increase the resolution of the images. The up-sampling layer increases the number and/or columns of the input data. The values of the new elements will depend on the interpolation method of the layer. The deconvolutions, on the other hand, consists of a convolution with a filter derived from the transposition and inversion of the tensor resulting from the original convolution filter (Aggarwal, 2018). This type of layer presents the same properties, parameters and hyperparameters of a convolution layer, however, the output has more dimensions than the input.

In the last two layers, a ReLU operation with a maximum value equal to 1 limited the values of the output image between 0 and 1, and a rescaling operation set the output values in a 0-255 scale, which is the same scale as the output images.

5.2.2.3. Training and validation

The networks were trained using the RMSprop optimiser with a learning rate of 0.005 and a limit of 300 epochs. Early stopping was configured to interrupt the training if the performance of the emulator stabilised after 15 consecutive epochs. The loss function was the MSE, and the RMSE was used to monitor the training progress. 80% of the images were used for training and 20% for validating the model.

Each training epoch lasted approximately 20 seconds for simulation set 1 and 35 seconds for set 2. The training was performed in a CPU Intel Core i7-8750H 2.20 GHz, 16 Gb RAM and a graphics card Nvidia GeForce GTX 1060 with 6 Gb, using parallel computing aided by the CUDA toolkit v. 11.2 (Harish & Narayanan, 2007).

The model architecture with the best performance obtained with the simulation set 1 was used in the simulation set 2 scenarios. For scenarios 1 to 4 the models were trained for each domain and each output variable (CES and CTBLC), resulting in 4 emulators. In scenarios 5 to 8, the models were trained using the data of both domains at once, resulting in two emulators. These last scenarios verified the possibility of using only one model to predict the morphological evolution of different domains. Additionally, in this second case, the models would be capable of identifying the domain according to the input data.

5.2.2.4. Performance analysis

To compare the results of the numerical models and the results of the emulator, the RMSE of the model in the validation dataset was computed for each time step:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Em - Nm)^2} \quad (5\text{-}1)$$

where n is the number of grid points, Em is the result of the emulator and Nm is the result of the numerical model. Also. the mean error was determined to identify the location of the highest errors in the domain:

$$ME = \sqrt{\frac{1}{m}\sum_{i=1}^{n}(Em - Nm)} \quad (5\text{-}2)$$

where m is the number of time steps of the validation dataset.

Besides, the kernel-density estimation (KDE) of the errors in the domains was computed, considering all the time steps. The KDE method estimates the probability density function of a non-parametric dataset (Chen, 2017). This last metric is relevant to successfully analyse the error distribution in all the domains to identify areas in which the emulator performance is better.

5.3. Results and discussion

5.3.1. Simulation set 1

Simulation set 1 had the objective of evaluating the influence of the up-sampling operation in the architecture of the emulator. The RMSE in scenario 1 was 0.0241 mm, 0.0247 mm, and 0.0183 mm for scenarios 1, 2 and 3, respectively.

All the models could achieve reasonable results, however, the use of the tanh activation function in the transposed convolution layer (Scenario 3) resulted in the emulator with the best performance. It probably occurred due to the increase in the number of trainable parameters in the network, granting a better approximation to the numerical model results. Additionally, it did not meaningfully affect the computational time required by the emulator to process the validation dataset. Therefore, the architecture used in scenario 3 was used to evaluate the computational performance in simulation set 2 scenarios.

5.3.2. Computational performance

The DL models needed 5 seconds to generate the morphodynamics results referred to the whole dataset (training and validation). However, it is also necessary to run the hydrodynamic module of the numerical model to generate the input data required by the emulator.

Therefore, to correctly estimate the computational time required for the DL model, it was necessary to run the numerical model with the sediment transport and morphology modules deactivated. This comparison was conducted using the simulation set 1 model, due to its simplified bathymetry and, consequently, hydrodynamics patterns. This model run a 6 h simulation, with morfac set to 0 and with the sediment transport and morphology modules deactivated.

With both morphodynamics modules activated, 28 minutes were necessary to run the simulation. With only the hydrodynamics module activated, the simulation took 23 minutes. Additionally, 15 seconds

were necessary to generate 108 input images (54 per variable), and 5 seconds to run the DL model with this data, representing a reduction of 23% in the total time necessary to forecast the morphodynamic evolution of the domain, considering the predictions for the whole simulation period. Therefore, the main limiter to reduce the morphodynamics forecast computational time is the hydrodynamic simulation.

The emulator that used different input and output resolutions had the best performance gain. The hydrodynamic simulation lasted 3.3 minutes, resulting in a time reduction of 87% to generate the morphodynamics results. As the major restriction to reduce the computational time was the hydrodynamic simulation, reducing the grid resolution to train the emulator could represent a significant performance gain.

Additionally, it must be highlighted that it would be necessary to update the numerical model bathymetry before generating the hydrodynamics results of the next forecast window. The bathymetry variation directly affects the hydrodynamics variables, consequently, update it during this type of simulation is of utmost importance to correctly forecast the HMD variables and reduce the uncertainties of the results.

Looking forward to optimizing the generation of the hydrodynamics results, three approaches could be considered. Firstly, the XBeach simulation time presents a high dependence on the wave velocities that must satisfy the CFL condition (Courant et al., 1928; Roelvink et al., 2009; Schneider et al., 2013). The simplest method to ensure this condition and maintain the grid resolution is to set the smallest time step for the entire simulation, which is determined according to the highest wave speed (Gnedin et al., 2018). Although, this approach could increase the total simulation time, especially if the objective of the model is to simulate extreme events, which would result in higher wave velocities and demand a reduction in the simulation time step to satisfy the CFL condition.

Alternatively, the reduction of the grid resolution would allow higher time steps, reducing the total simulation time and satisfying the CFL condition. However, the focus of this methodology was to reduce the computational costs of high-resolution numerical models during short-term events, hence, this hypothesis was not considered in this work.

The last option would be to optimise the input image generation loop. In this study, a loop was used to generate the images of each variable. In each step of the loop, the image referring to the ith time step was created and exported. To reduce this time, a single loop for all variables was performed, which partially reduced the redundancy in the code, instead of using one loop for each variable, optimising this process. In despite of that, the total time required to generate the hydrodynamics images was 15 seconds. Hence, an optimisation in this process would not significantly reduce the total time required.

Additionally, an important result achieved, if compared with previous studies, is the reduction in the number of images necessary to train the emulator. In Weber de Melo et al (Weber de Melo et al., 2022), 1095 images were generated to train and test the emulator, whereas herein only 162 images were needed in the simulation set 1 to achieve good results. This demonstrate that using AI models do not necessarily require a huge database for training and testing, which agrees with data-centric AI concepts that affirms that the quality of the data is far more important than its quantity (Sambasivan et al., 2021). The XBeach database required 20.61 Mb of disk space, while the D3D needed 137.1 Mb, representing a reduction of 85% in the disk space usage. The emulator with different input/output resolutions needed 40.8 Mb, representing 70% less disk space.

5.3.3. Influence of the grid resolution

The difference between the emulators trained with inputs of different resolution and the numerical model output for the test dataset are presented in Figure 5-4. The mean error of both emulators was negligible, but the mean error of the LR emulator was one order higher than the OR emulator. The maximum error occurred in the OR emulator results was 21 cm and in the LR emulator was 26 cm. These values are considerably higher than the mean error, however they are restricted to specific areas of the domain, particularly near the domain boundaries, which normally fall out of the region of interest when using numerical modelling approaches.

Figure 5-5 shows the RMSE of the OR and LR emulators in the test dataset. It can be observed that the performance of both models is remarkably similar in all points and the maximum mean error was below 5 cm in the OR model and below 8 cm in the LR. Besides, both error curves follow a similar pattern, having the same behaviour at timesteps 14 and 17, for example.

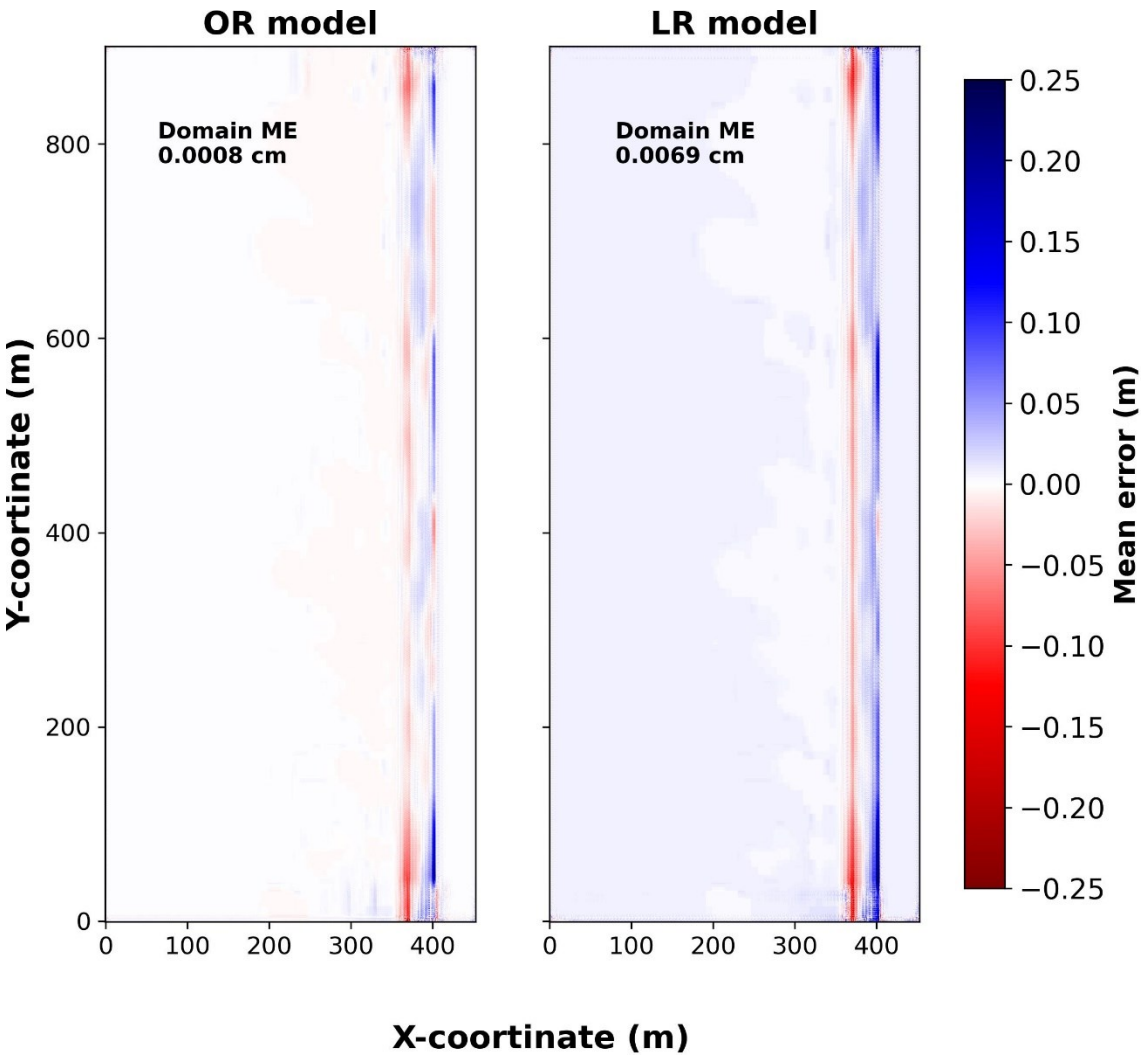


Figure 5-4. Comparison between emulators using inputs with different grid resolution.

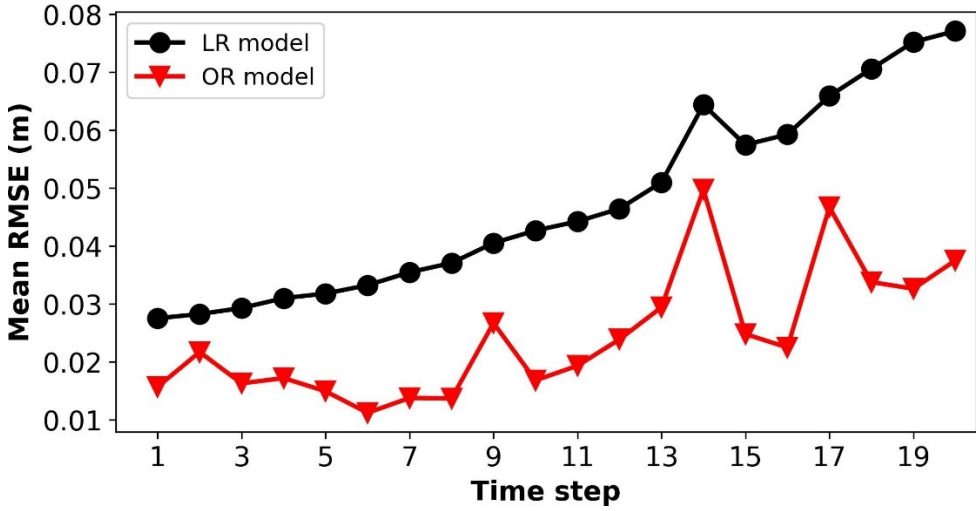


Figure 5-5. Emulators RMSE in the test dataset.

5.3.4. Simulation set 2

Figure 5-6 presents a comparison between the output in the last time step of the XBeach model with the results obtained with the emulator using the CES images (scenarios 1 and 3). The emulator

could reproduce the main patterns of the morphological evolution of the domains with a remarkable resemblance. The erosion and accretion bars can be observed in the southern area of the beach in both images, as well as a larger accretion area around the coordinate (300, 300). This pattern results from the 315° wave angle of the wave spectrum applied at the boundaries of the XBeach model, which created this accretion spot upstream of the structures. However, the emulator underestimated the accretion area near this coordinate, which is evident due to the thinner blue bar in the breakwater domain and the reduced quantity of purple pixels in the groin domain.

This result demonstrates that the use of CES as output instead of CTBLC did not significantly affect the performance of the emulator. The emulator accurately simulated the main morphodynamic patterns of both study areas, independently of the selected domain.

Comparing the results of both selected coastal structures, the output of the breakwater emulator is more similar to the numerical model result than the output of the groin emulator. This could be related with the uniformity of the breakwater influence along the beach when compared to the groin. The groin breaks the linearity of the erosion/accretion bars, creating an area with a more complex morphodynamic pattern. Hence, more training images would be necessary to improve the precision of the groin emulator.

Additionally, the emulator underestimated the erosion that occurred near the groin, in the centre of the domain. It can be observed that the numerical model simulated a larger erosion area in the north-south direction, whilst this pattern was not accurately forecasted by the emulator, mainly at the north of the breakwater. In the case of the breakwater emulator, it underestimated the accretion that occurred in the east face.

The performance of the emulator could be improved by increasing the size and variability of the training dataset variables. The use of more images would give more information about the output pattern expected in that area of the domain, increasing its performance. Moreover, the error in the areas wherein the morphodynamic variation is more uniform was lower than the error in areas with higher variability, namely near the breakwater and the groin. However, it must be stressed that the emulator could still identify that the morphological change in the groin was zero, probably because this pattern was more common in all training time steps.

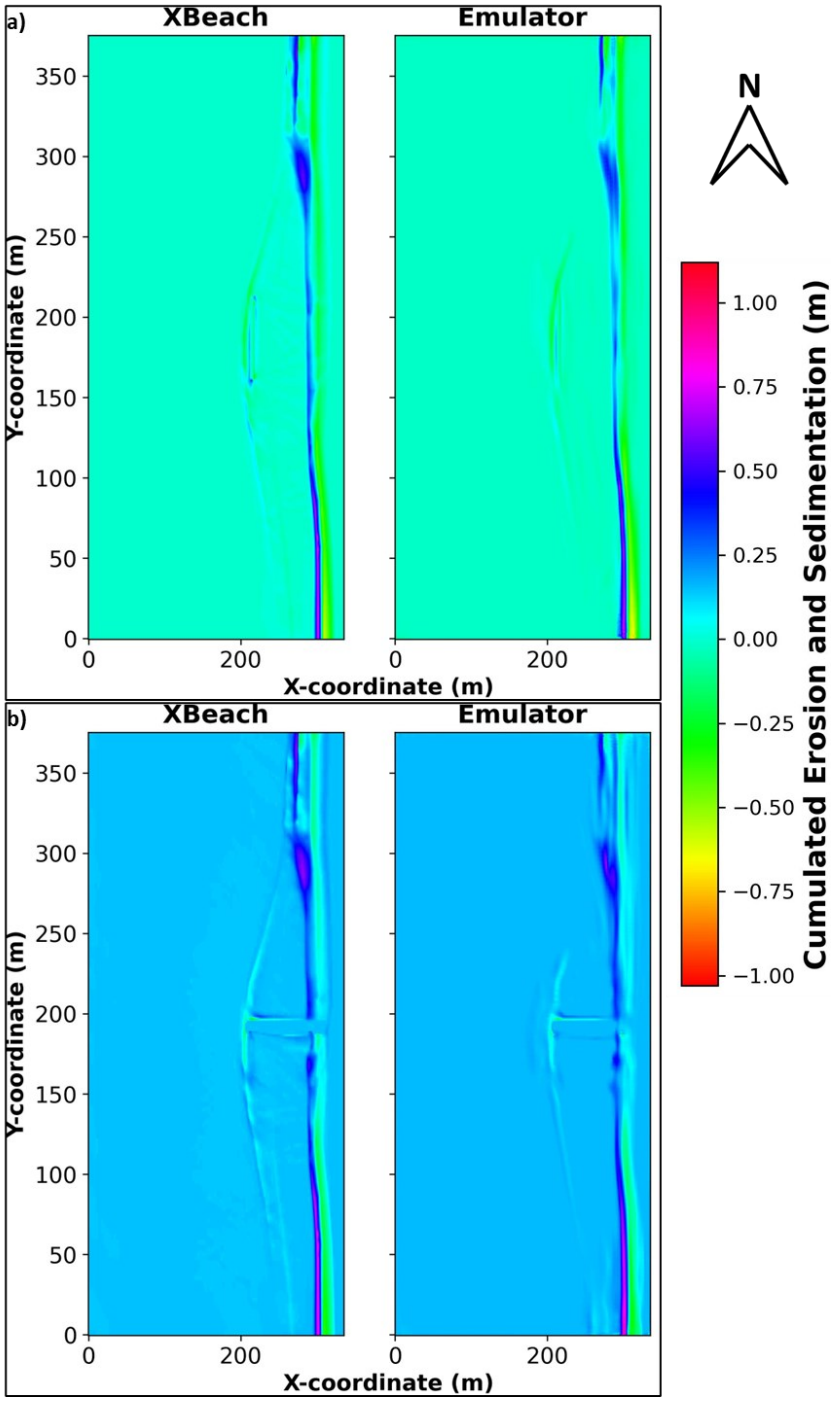


Figure 5-6. Comparison between the CES simulated by the XBeach model and the emulator in the last simulation time step. a) Breakwater domain; b) Groin domain.

To assess the error variance of the emulators in all the validation timesteps, the RMSE was computed between the DL model output and the numerical model result for each pixel of the image. The mean value of the RMSE was then determined by dividing the RMSE for the validation time steps. Figure 5-7 presents these results, separated by the output variable used to train the DL models.

The results demonstrated that the models trained in only one domain had a better performance than the models trained in both domains. It probably befell because the input variables values were

affected by the groin and the breakwater only in the area around these interventions. In the other parts of the domain, the input values were more resemblant due to the boundary conditions being the same. Hence, the network could learn the CES patterns for both domains, having similar errors as demonstrated in Figure 5-7– a.

In Figure 5-7 – b, the error of scenarios 6 and 8 was quite different when compared to the models trained with CES data. It probably happened because the CTBLC patterns are more similar between the domains, given that these results show only the areas in which the bed level was altered in the last time step. This fact could explain why the performance of scenario 8 was closer to scenarios 2 and 4, a different pattern when analysing the CES results. Despite that, the performance of the models was better when the training conditions were from only one domain. It probably occurred owing to a possible ambiguity in the input and output data since most of the pixel values in the input images were not affected by the domain.

Besides, the BWG emulator had to decide which domain to prioritise during the training, resulting in a better adjustment to the groin domain instead of the breakwater. This could explain why the error in scenario 8 was lower than the error in scenario 6. However, the performance in scenario 8 overcame scenario 4 between time steps 15 and 23, indicating that the DL model parameters were more suitable for those inputs. This would allow the use of ensemble techniques (Biolchi et al., 2022; Iglesias et al., 2022) to create a unique and more robust model that could merge individual models trained in specific conditions, allowing the model to select a determined input-output path according to the available data.

Figure 5-8 presents the mean error obtained in the simulation set 2 scenarios. This result allows to identify the average location of the major errors in the numerical model domains. The whiter colours represent areas where the error was near 0, red colours represent the emulator underestimation and blue colours represent overestimation. Scenarios 2, 3 and 5 had an average overestimation of the numerical model results, whilst the other scenarios presented an average underestimation.

Figure 5-8 – a shows that, in all scenarios, the DL models underestimated the values of an erosion area that occurred near the coordinates (280, 350). This probably occurred owing to the proximity to the boundary of the numerical model, resulting in a higher uncertainty for that area. Moreover, the model also underestimated the accretion area at coordinates (300, 300). Despite that, all the models had a reasonable agreement between the numerical model results and the data model outputs.

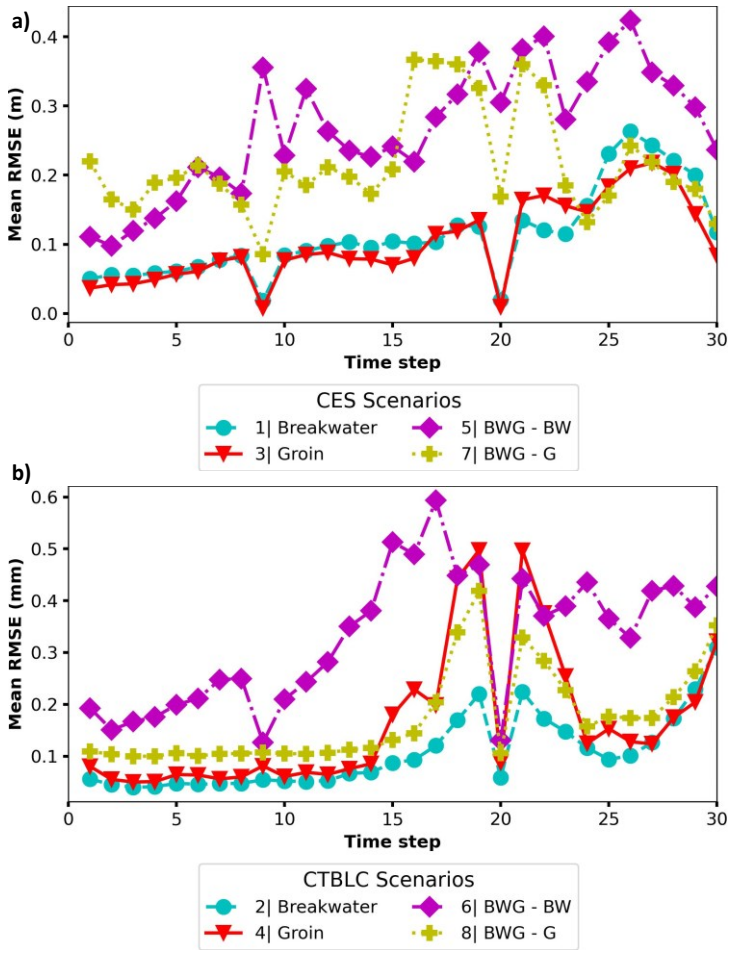


Figure 5-7. Models mean RMSE at each validating time step. a) Error (m) of the models trained in CES data; b) Errors (mm) of the models trained in CTBLC data.

It can be observed in Figure 5-8 – b that the highest errors in scenarios 2 and 6 occurred near the beach, where the sediment transport is more intense, although is in this same area where the lowest errors in scenarios 4 and 8 are found. This demonstrates that the DL model trained in both domains (scenarios 6 and 8) adjusted its parameters to better represent only one domain, which was the one with lower errors. This same behaviour can be observed in scenarios 5 and 7, in which the error in scenario 7 was much lower than the error in scenario 5.

Lastly, the error distribution was plotted (Figure 5-9), which allows for assessing the variability of the errors, considering all the time steps of the validation dataset. The peak values in the distributions are approximately the mean errors presented in Figure 5-8. The models trained in single domain data had a lower variability in the errors, regardless of the output variable. Additionally, the emulators presented acceptable results for all the domain areas, which is indicated by the error low variability.

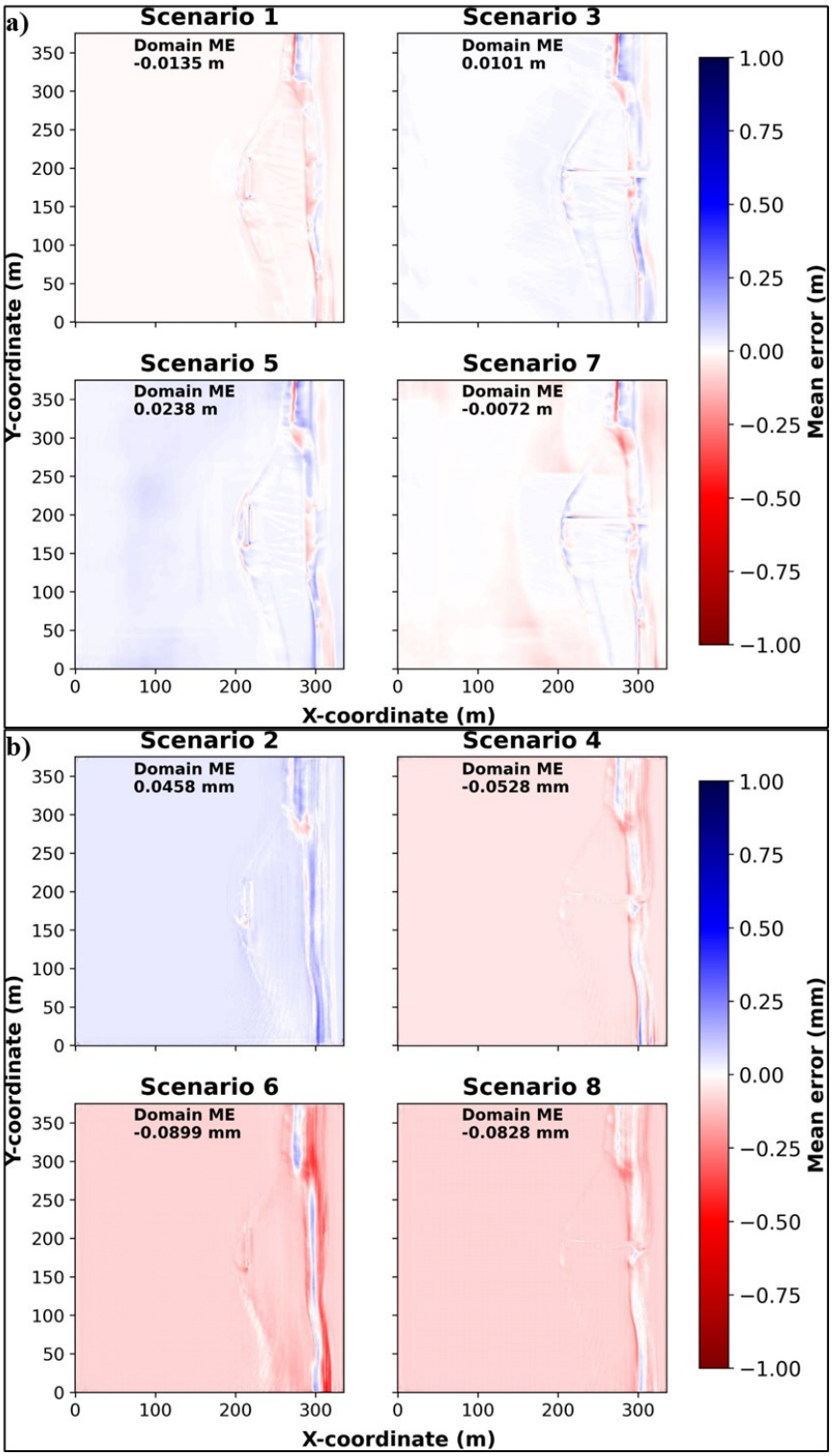


Figure 5-8. Mean error of the DL models in the domains. Red colours represent emulator underestimation and blue colours overestimation. White colours indicate a perfect agreement between the numerical model and the emulator. a) Error of the models (m) in forecasting the CES; b) Error of the models (mm) in forecasting CTBLC.

Therefore, DL emulators could successfully surrogate morphological modules of numerical modelling tools as the XBeach without significant changes in the result. The errors achieved by the DL

models represented 1.1% of the maximum CES values and 0.6% of the maximum CTBLC values, reinforcing the similarity between the results of the emulator and the XBeach models.

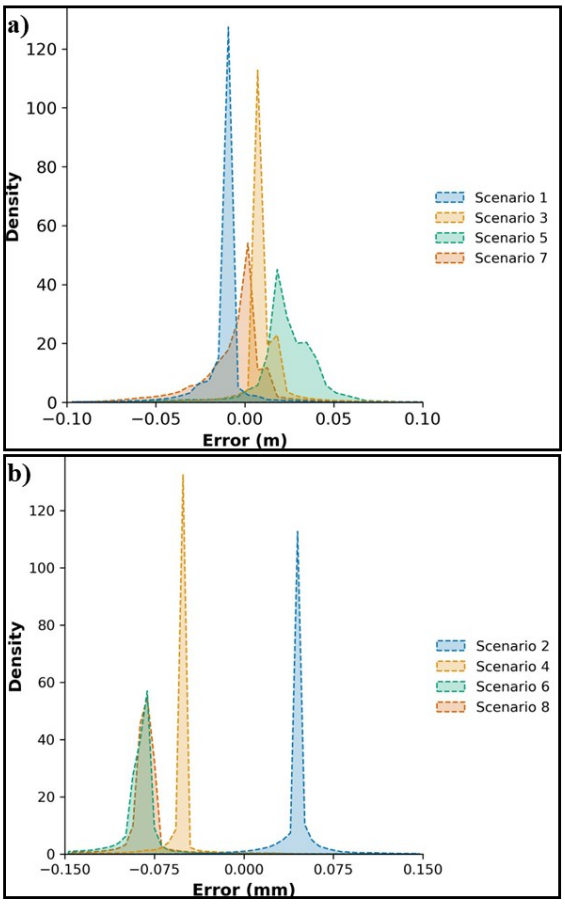


Figure 5-9. DL model error distribution. a) Distribution of the CEM errors; b) Distribution of the CTBLC errors.

5.3.5. Future of HMD modelling

The capacity of the emulator in reproducing the main morphodynamic patterns of coastal beaches and the reduction of computational resources requirements demonstrate the potential of data-driven models, staring to modify the logic behind environmental modelling. Numerical models can achieve reliable results by solving deterministic equations, although these equations simplify some aspects of the reality. For instance, the bottom friction coefficient is usually simplified to a uniform value for all the computational domain, although it is known that this coefficient varies according to the characteristics of the sediments, bottom structure, depth and vegetation. These details can be inserted into numerical models, but the efforts in their implementation would sharply rose.

Data-driven models, conversely, intrinsically consider all the domain features, as it will be reflected in the training data. It is also important to stress that the data characteristics determines the validation interval in which data models can be applied. For example, for the implementation of a model to predict extreme events effects it is not expected that a data model trained in average conditions will

have satisfactory performance, as well as a numerical model calibrated in average conditions will not either. In this context, a change of paradigm may be interesting for future modelling methodologies. Instead of creating more complex models that can consider all sources of uncertainty, efforts could be directed to improve the quality, variability, and frequency of data, thereby reducing the uncertainties of forecasts to the characteristics of the training data and to the errors in the measurements.

Regarding the computational resources required for modelling, the use of parallel computing and high-performance computers would reduce the simulation computational time. However, these types of machines, and their necessary infrastructure, have high associated costs, limiting the application of real time forecasting platforms to those who can afford them. Therefore, the development of methodologies that does not depend on super computers is a promising path to guarantee a wider application of early warning systems.

5.4. Conclusions

This study applied an emulator to a hydrodynamic coastal numerical model using a deep learning approach implemented in Python and the Tensor Flow framework. Images of the hydrodynamic variables, bed shear stress and GLM velocity, simulated with a XBeach model, were used as input of the DL models that forecasted the morphological evolution for three different domains.

In simulation set 1, a DL architecture based on the previous work of Weber de Melo et al. (2022) had a satisfactory performance for the XBeach models, demonstrating the suitability of the purposed methodology to reduce the computational time of morphodynamic numerical models. There is no reason to think that the developed emulator cannot be used with other HMD software appropriate to simulate coastal stretches.

Additionally, the performance of the emulator was improved by the replacement of the up-sampling operation in the neural network architecture by transposed convolutions. This procedure increased the number of trainable parameters in the network, allowing the reduction of the error between the numerical model results and the DL model outputs without significantly affecting the computational time required by the emulator to forecast the validation dataset. Furthermore, the tanh activation function resulted in a lower error when compared to the ReLU activation function.

Regarding simulation set 2, it was demonstrated that the emulator could reproduce the numerical model morphodynamic results in coastal domains under the influence of a submerged breakwater or a groin. The errors obtained were two orders lower than the maximum and minimum values of the output variables.

Furthermore, the emulator accurately reproduced both CES and CTBLC results, indicating that the DL model can be easily adapted to forecast any variable of interest if there is enough data to train the network. To determine the size of the training dataset, it is necessary to assess the results in the validation dataset. The use of few data can result in models that can barely forecast the main patterns of the output variable, whilst the use of excessive data can drastically increase the time required for training the network. It is recommended that the training data includes enough variability to correctly represent the application range of the emulator. Moreover, the image database used in this paper requires 85% less disk space than in Weber de Melo et al. (2022), demonstrating that the methodology can achieve reliable results with fewer data.

It can be also concluded that it is possible to train one emulator with data from different domains, however, some precautions are necessary, namely identifying the possibility of a domain prioritization by the DL model. The errors increased in three of the scenarios in which the emulators were trained with both datasets (scenarios 5, 6 and 7) when compared to the emulators trained with a single one. Scenario 8 model results overcame scenario 4 in several time steps, although its error was slightly worse in the others time steps.

Regarding the computational costs of the emulator, it was achieved a time reduction of 23% in simulating the morphodynamic evolution using numerical models with the same input and output grid resolution, whilst a time reduction of 87% was obtained using the emulator with low grid resolution inputs. The main time-consuming task in this methodology was the generation of hydrodynamics results by the numerical model, which limited the improvement of the proposed methodology. Hence, the performance gain with the emulator application would be higher for long-term hydrodynamic simulations due to the lower wave velocities magnitudes, which allows simulations with higher time steps. Although the error in the LR emulator had slightly increased, its results were still similar to the numerical model. Additionally, the use of images as input and output allows the potential application of this methodology to any other numerical model, as mentioned before.

Reducing the total computational time required for simulating the short-term morphological evolution of a coastal area can be important to facilitate the application of XBeach models to real-time early warning and forecasting systems. This would allow to optimise the response of the authorities during extreme events and to increase the resilience of coastal communities.

Finally, this study demonstrated that the use of DL techniques can reduce the computational costs of numerical models, allowing their implementation in real-time forecasting and warning systems. The emulators accurately reproduced the CES and CTBLC results obtained with the XBeach software,

which is one of the most advanced coastal morphodynamics simulation model available, achieving an error two orders lower than the morphodynamics variables range values.

5.5. References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., ... Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. http://arxiv.org/abs/1603.04467

Aggarwal, C. C. (2018). Neural Networks and Deep Learning. In Artificial Intelligence. Springer International Publishing. https://doi.org/10.1007/978-3-319-94463-0

Anderson, D. L., Ruggiero, P., Mendez, F. J., Barnard, P. L., Erikson, L. H., O'Neill, A. C., Merrifield, M., Rueda, A., Cagigal, L., & Marra, J. (2021). Projecting Climate Dependent Coastal Flood Risk With a Hybrid Statistical Dynamical Model. Earth's Future, 9(12). https://doi.org/10.1029/2021EF002285

Biolchi, L. G., Unguendoli, S., Bressan, L., Giambastiani, B. M. S., & Valentini, A. (2022). Ensemble technique application to an XBeach-based coastal Early Warning System for the Northwest Adriatic Sea (Emilia-Romagna region, Italy). Coastal Engineering, 173. https://doi.org/10.1016/j.coastaleng.2022.104081

C.A. Oliveira, T., Cagnin, E., & A. Silva, P. (2020). Wind-waves in the coast of mainland Portugal induced by post-tropical storms. Ocean Engineering, 217. https://doi.org/10.1016/j.oceaneng.2020.108020

Chen, Y.-C. (2017). A tutorial on kernel density estimation and recent advances. Biostatistics & Epidemiology, 1(1), 161–187. https://doi.org/10.1080/24709360.2017.1396742

Courant, R., Friedrichs, K., & Lewy, H. (1928). Über die partiellen Differenzengleichungen der mathematischen Physik. Mathematische Annalen, 100(1), 32–74. https://doi.org/10.1007/BF01448839

Deltares. (2018). Delft3D-Flow: Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments - User Manual.

Dumoulin, V., & Visin, F. (2016). A guide to convolution arithmetic for deep learning. http://arxiv.org/abs/1603.07285

Ferreira, Ó., Plomaritis, T. A., & Costas, S. (2019). Effectiveness assessment of risk reduction measures at coastal areas using a decision support system: Findings from Emma storm. Science of the Total Environment, 657, 124–135. https://doi.org/10.1016/j.scitotenv.2018.11.478

Gharagozlou, A., Anderson, D. L., Gorski, J. F., & Dietrich, J. C. (2022). Emulator For Eroded Beach And Dune Profiles Due To Storms. Journal of Geophysical Research: Earth Surface, 127(8). https://doi.org/10.1029/2022JF006620

Gnedin, N. Y., Semenov, V. A., & Kravtsov, A. v. (2018). Enforcing the Courant–Friedrichs–Lewy condition in explicitly conservative local time stepping schemes. Journal of Computational Physics, 359, 93–105. https://doi.org/10.1016/j.jcp.2018.01.008

Haasnoot, M., Winter, G., Brown, S., Dawson, R. J., Ward, P. J., & Eilander, D. (2021). Long-term sea-level rise necessitates a commitment to adaptation: A first order assessment. Climate Risk Management, 34. https://doi.org/10.1016/j.crm.2021.100355

Harish, P., & Narayanan, P. J. (2007). Accelerating Large Graph Algorithms on the GPU Using CUDA. In High Performance Computing – HiPC 2007: Vol. 4873 LNCS (pp. 197–208). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77220-0_21

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in Science \& Engineering, 9(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Iglesias, I., Pinho, J. L., Avilez-Valente, P., Melo, W., Bio, A., Gomes, A., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2022). Improving Estuarine Hydrodynamic Forecasts Through Numerical Model Ensembles. Frontiers in Marine Science, 9. https://doi.org/10.3389/fmars.2022.812255

Poelhekke, L., Jäger, W. S., van Dongeren, A., Plomaritis, T. A., McCall, R., & Ferreira, Ó. (2016). Predicting coastal hazards for sandy coasts with a Bayesian Network. Coastal Engineering, 118, 21–34. https://doi.org/10.1016/j.coastaleng.2016.08.011

Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., & Rama, B. (2022). IPCC, 2022: Climate Change 2022: Impacts,

Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. https://doi.org/10.1017/9781009325844

Rautenbach, C., Trenham, C., Benn, D., Hoeke, R., & Bosserelle, C. (2022). Computing efficiency of XBeach hydro- and wave dynamics on Graphics Processing Units (GPUs). Environmental Modelling & Software, 157, 105532. https://doi.org/10.1016/j.envsoft.2022.105532

Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R., & Lescinski, J. (2009). Modelling storm impacts on beaches, dunes and barrier islands. Coastal Engineering, 56(11–12), 1133–1152. https://doi.org/10.1016/j.coastaleng.2009.08.006

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 9351, Issue Cvd, pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Rutten, J., Torres-Freyermuth, A., & Puleo, J. A. (2021). Uncertainty in runup predictions on natural beaches using XBeach nonhydrostatic. Coastal Engineering, 166(February), 103869. https://doi.org/10.1016/j.coastaleng.2021.103869

Sambasivan, N., Kapania, S., & Highfll, H. (2021, May 6). Everyone wants to do the model work, not the data work: Data cascades in high-stakes ai. Conference on Human Factors in Computing Systems - Proceedings. https://doi.org/10.1145/3411764.3445518

Santhanam, V., Morariu, V. I., & Davis, L. S. (2016). Generalized Deep Image to Image Regression. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua, 5395–5405. https://doi.org/10.1109/CVPR.2017.573

Schneider, K., Kolomenskiy, D., & Deriaz, E. (2013). Is the CFL Condition Sufficient? Some Remarks. In The Courant–Friedrichs–Lewy (CFL) Condition (pp. 139–146). Birkhäuser Boston. https://doi.org/10.1007/978-0-8176-8394-8_9

Simmons, J. A., & Splinter, K. D. (2022). A multi-model ensemble approach to coastal storm erosion prediction. Environmental Modelling and Software, 150. https://doi.org/10.1016/j.envsoft.2022.105356

Smallegan, S. M., Irish, J. L., Van Dongeren, A. R., & Den Bieman, J. P. (2016). Morphological response of a sandy barrier island with a buried seawall during Hurricane Sandy. Coastal Engineering, 110, 102–110. https://doi.org/10.1016/j.coastaleng.2016.01.005

Suzuki, T., & Cox, D. T. (2021). Evaluating XBeach performance for extreme offshore-directed sediment transport events on a dissipative beach. Coastal Engineering Journal, 63(4), 517–531. https://doi.org/10.1080/21664250.2021.1976452

Trouw, K. J. M., Zimmermann, N., Mathys, M., Delgado, R., & Roelvink, D. (2012). Numerical modelling of hydrodynamics and sediment transport in the surf zone: a sensitivity study with different types of numerical models. Coastal Engineering Proceedings, 1(33), 23. https://doi.org/10.9753/icce.v33.sediment.23

van Rijn, L. C. (2007). Unified View of Sediment Transport by Currents and Waves. I: Initiation of Motion, Bed Roughness, and Bed-Load Transport. Journal of Hydraulic Engineering, 133(6), 649–667. https://doi.org/10.1061/(ASCE)0733-9429(2007)133:6(649)

van Thiel de Vries, J. S. M. (2009). Dune erosion during storm surges. Faculty of Civil Engineering and Geosciences - TU Delft.

Vieira, B. F. V., Pinho, J. L. S., Barros, J. A. O., & Antunes do Carmo, J. S. (2020). Hydrodynamics and morphodynamics performance assessment of three coastal protection structures. Journal of Marine Science and Engineering, 8(3). https://doi.org/10.3390/jmse8030175

Viitak, M., Avilez-Valente, P., Bio, A., Bastos, L., & Iglesias, I. (2021). Evaluating wind datasets for wave hindcasting in the NW Iberian Peninsula coast. Journal of Operational Oceanography, 14(2), 152–165. https://doi.org/10.1080/1755876X.2020.1738121

Vousdoukas, M. I., Ferreira, Ó., Almeida, L. P., & Pacheco, A. (2012). Toward reliable storm-hazard forecasts: XBeach calibration and its potential application in an operational early-warning system. Ocean Dynamics, 62(7), 1001–1015. https://doi.org/10.1007/s10236-012-0544-6

Vousdoukas, M. I., Ranasinghe, R., Mentaschi, L., Plomaritis, T. A., Athanasiou, P., Luijendijk, A., & Feyen, L. (2020). Sandy coastlines are under threat of erosion. Nature Climate Change, in press (March). https://doi.org/10.1038/s41558-020-0697-0

Weber de Melo, W., Pinho, J., & Iglesias, I. (2022). Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model. Journal of Hydroinformatics. https://doi.org/10.2166/hydro.2022.068

Whitaker, J., Khrulev, C., Huard, D., Paulik, C., Hoyer, S., Pastewaka, L., Mohr, A., Marquardt, C., Counwenberg, B., Taves, M., Cuntz, M., Roet, S., Whitaker, J., Brett, M., Bohnet, M., Hetland, R., Korenčiak, M., Barnam Andrew, Hamman, J., ... Kluyver, T. (2020). Unidata/ netcdf4-python: version 1.5.5 release. https://doi.org/https://doi.org/10.5281/zenodo.4308773

CHAPTER 6.

Early warning system for floods at estuarine areas: combining artificial intelligence with numerical models.

Manuscript number NHAZ-D-24-00244 submitted in February 2024 to the Natural Hazards Journal.

“The universe doesn't allow perfection.”

Stephen Hawking (1942 - 2018). Theoretical physicist. Quote from the book “A brief history of time”.

Chapter 6. Early warning system for flood in estuarine areas: combining artificial intelligence with numerical models

Abstract

Floods are among the most common natural disasters, causing countless losses every year worldwide and demanding urgent measures to mitigate their impacts. This study proposes a novel combination of artificial intelligence and process-based models to construct a flood early warning system (FEWS) for estuarine regions. Using streamflow and rainfall data, a deep learning model with long short-term memory layers was used to forecast the river discharge at the fluvial boundary of an estuary. Afterwards, a hydrodynamic numerical model was used to simulate water levels in the estuary. The river discharge predictors were trained using different forecasting windows varying from 3 h to 36 h to assess the relationship between the time window and accuracy. The insertion of attention layers into the network architecture was evaluated in order to enhance its forecasting capacity. The FEWS was implemented in the Douro River estuary, a densely urbanised flood-prone area in northern Portugal. The results demonstrated that the Douro Estuary FEWS is reliable for discharges of 6000 m^3/s, with predictions made 36 h in advance. For values higher than this, the uncertainties in the model predictions increased; however, they were still capable of detecting flood occurrences.

6.1. Introduction

The increase in the intensity and frequency of extreme events is one of the most severe impacts of climate change. Floods, storms, and droughts, which usually present longer return periods, are becoming common problems worldwide, threatening vulnerable populations and causing several losses. In particular, floods are among the most severe natural disasters, mainly affecting lower-income countries (Rentschler et al., 2022), and the cost of floods was estimated at US$ 10.2 trillion per year considering the climate scenario of 1.5 °C above pre-industrial levels (Jevrejeva et al., 2018).

Therefore, implementing measures to mitigate the impacts of extreme events is of utmost importance. Non-structural measures, such as flood early warning systems (FEWS), are feasible and effective climate adaptation measures (WMO, 2022), and are fostered as one of the main disaster risk reduction measures by the Sendai Framework for Disaster Risk Reduction (UNDRR, 2023). An early warning system is composed of five pillars (WMO, 2022): 1) disaster risk and management that focuses on the identification and mapping of hazards and vulnerabilities on a local scale; 2) observations and forecasting that monitors risk-related parameters and configures forecasting and warning systems; 3) warning dissemination and communication that guarantees that the warnings reach everyone in the risk area with clear information; 4) preparedness and response capabilities that ensure that the population

knows how to react during a risk scenario; and 5) monitoring and evaluation: frequent updates and assessments of the early warning systems.

In this study, we propose a novel combination of data and process-based models to implement the forecasting component of a FEWS for estuarine regions. We fed DL models with LSTM layers with streamflow and rainfall data to forecast the river discharge at the fluvial boundary. Subsequently, a hydrodynamic numerical model was configured to simulate water levels in the study area. The DL models were trained using different forecasting windows varying from 3 h to 36 h to assess the relationship between the time window and the accuracy. Furthermore, an attention layer was tested in the architecture of the network to enhance its forecasting capacity.

Recent studies have demonstrated the suitability of LSTM networks for time-series forecasting. Kim & Kim (2020) combined a LSTM network with a random forest algorithm to forecast the flood hazard rates in South Korea, Liu et al. (2023) implemented a LSTM-based model to forecast the floods of the Xiangjiang River in China, and Seshan et al. (2024) used a LSTM-based autoencoder to predict the water quality in a wastewater treatment plant. Moreover, the combination of LSTM with attention mechanisms can improve the forecasting capacity of the predictive models (Wang et al.2023; Fan et al. 2019).

Our framework was implemented in the Douro River estuary, a flood-prone area in northern Portugal (Venâncio et al., 2019; Weber de Melo et al., 2022), where inundations are among the most threatening natural disasters in the region (Monteiro et al.2017) and are caused mainly by intense rainfall peaks and low infiltration rates (Alcoforado et al.2021). Measurements of past water level peaks can still be found at historical sites, such as port wine cellars in Vila Nova de Gaia. Additionally, the region presents economic and social significance, being the second-largest Portuguese urban area with more than 1.7 million inhabitants.

The Douro River estuary flows into the North Atlantic Ocean, and is limited upstream by the Crestuma-Lever Dam, which controls the river discharges into the estuary. One of the challenges of this study is forecasting the discharges in the Crestuma Lever Dam because their values vary according to the dam hydropower generation schemes and storage capacity. An accurate estimate of fluvial discharges is essential for precise downstream flood simulations, which are critical for implementing a FEWS in the region. Therefore, we aimed to implement discharge predictors using LSTM models and combine them with a process-based model to establish a real-time FEWS for the Douro River estuary.

6.2. Study area

The Douro River rises in Serra de Urbión, Spain, and is 927 km long until its mouth in Portugal, between Porto and Vila Nova de Gaia cities. The basin covers an area of 97603 km^2 (APA, 2014). The

estuarine region is highly urbanised, with tourism, navigation, and Porto wine cellars being among the most important economic activities.

The Douro region experiences seasonal changes in precipitation, with a wetter period in winter and a drier period in summer. The annual average precipitation is 1030 mm, varying between 2500 mm and 400 mm (APA, 2014). The average river flow is 1000 m^3/s during winter, whereas it is only 200 m^3/s during summer (Iglesias et al., 2019). Figure 6-1 shows the location of the study area, the Douro River sub-basins considered in this study, and the domain of the numerical model, which will be detailed in the following sections.

The hydrodynamic behaviour in the Douro Estuary is dominated by the tide and the river discharges, which is explained by its narrow and deep bathymetry (Iglesias et al., 2019; Weber de Melo et al., 2022). River discharges are controlled by the Crestuma-Lever dam, which is located 21 km upstream of the estuary mouth. The operation of a hydropower plant determines the freshwater inflow into the estuary according to the energy generation demand (Iglesias et al., 2021).

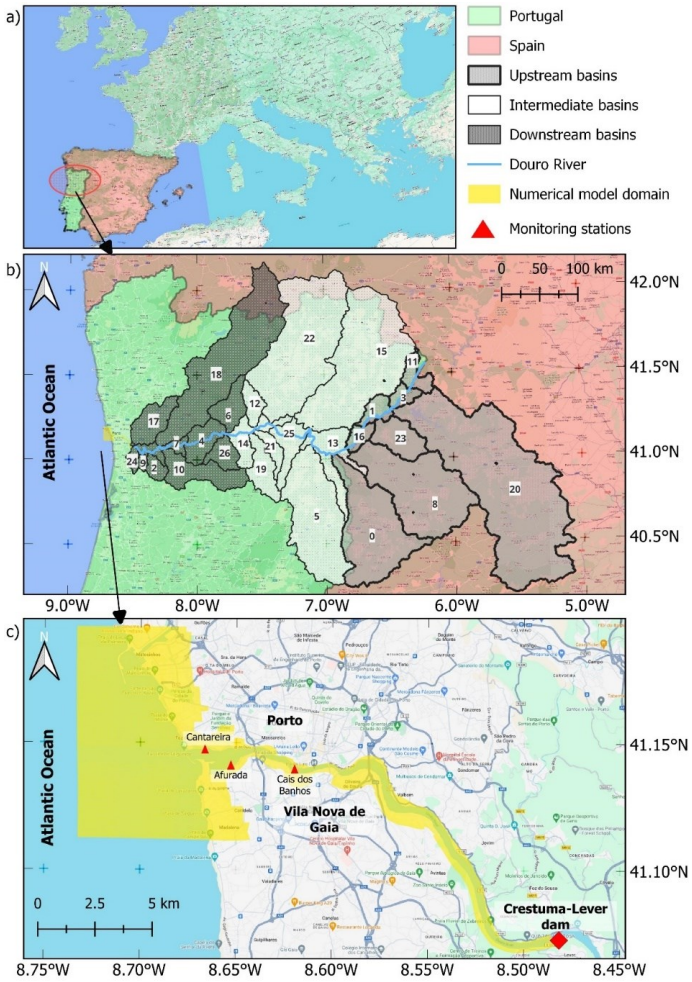


Figure 6-1. a) Douro River localisation in the Iberian Peninsula; b) Douro River sub-basins and estuarine area; c) Numerical model domain and fluvial boundary in the Crestuma-Lever dam.

6.3. Methodology

Four main tasks were required to successfully implement the FEWS in the Douro River estuary. Firstly, the field data necessary to implement both modelling solutions was collected and analysed. Secondly, an estuarine numerical model was implemented and calibrated. Thirdly, a data model was applied to forecast the river discharge at the fluvial boundary of the estuary. Finally, the performance of the proposed forecasting system was evaluated. Figure 6-2 shows the forecasting framework designed and adopted in this study.

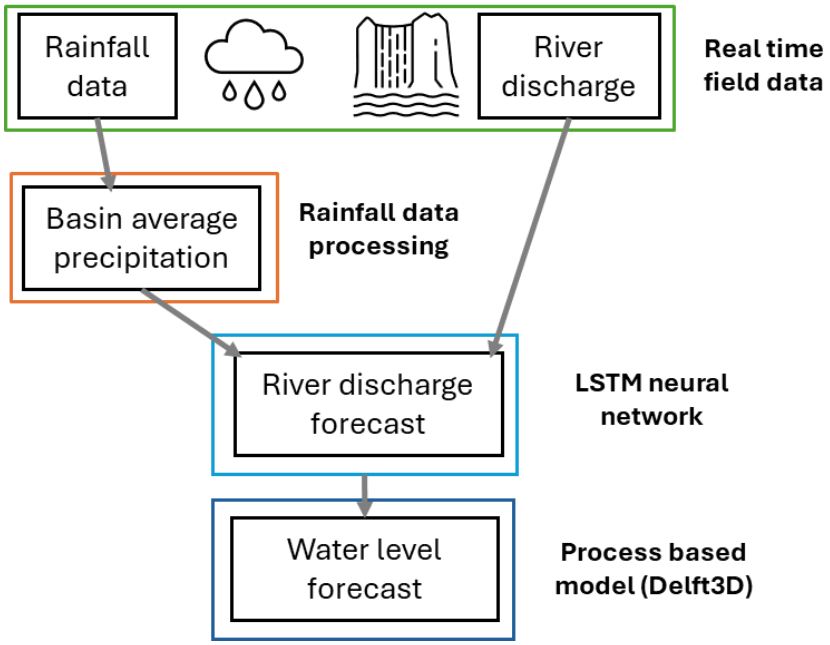


Figure 6-2. FEWS framework. Main steps to forecast water levels.

Hence, the Douro FEWS was configured to use near real-time precipitation and river discharge measured data, which are processed by an LSTM neural network that forecasts the river discharge in the fluvial boundary of the Douro River estuary. Then, water levels are simulated by the hydrodynamic numerical model of the estuarine area.

6.3.1. Field data collecting and processing

The models used in this study required rainfall and river discharge data. Historical precipitation data was obtained from the ERA5-Land database provided by the Copernicus Climate Change Service (Muñoz Sabater, 2019) implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF). This dataset is updated twice a day at 06:00 and 18:00, providing hourly outputs for the next 18 h (C3S, 2024). We selected the total precipitation from 2001 to 2021 with an hourly time-scale resolution and 0.25° x 0.25° horizontal resolution. The ERA5-Land data was interpolated to the 27 sub-basins upstream of Crestuma-Lever (Fig.1b), and the hourly precipitation volume in each basin was computed. Before passing these data to the LSTM model, the 5-days cumulated volume was calculated

and divided by the distance between the basin centroid and outlet in the Crestuma-Lever dam to provide information about the spatial influence of each basin on discharge. Therefore, the basins closest to Crestuma significantly influenced the discharge prediction.

The selected sub-basins were delimited using Shuttle Radar Topography Mission (SRTM) data and divided into three groups, upstream, intermediate, and downstream basins, to assess the influence of each group on river discharge. A total amount of 27 basins were identified and numbered from 0 to 26.

The precipitation patterns in the basins (Figure 6-3) indicate that the basins closest to the ocean have higher precipitation values than those away from the ocean. The Atlantic Ocean is the main source of moisture in this region. The most prominent situation is where the Atlantic westerly flows enter the Iberian Peninsula, mostly affecting the north-western sector during winter. Precipitation intensity is highest in north-western areas because of local orography and decreases inland (Ramos et al., 2014). Consequently, heavier rainfall occurs in the downstream sub-basin areas, whereas comparatively milder precipitation occurs in the upstream regions.

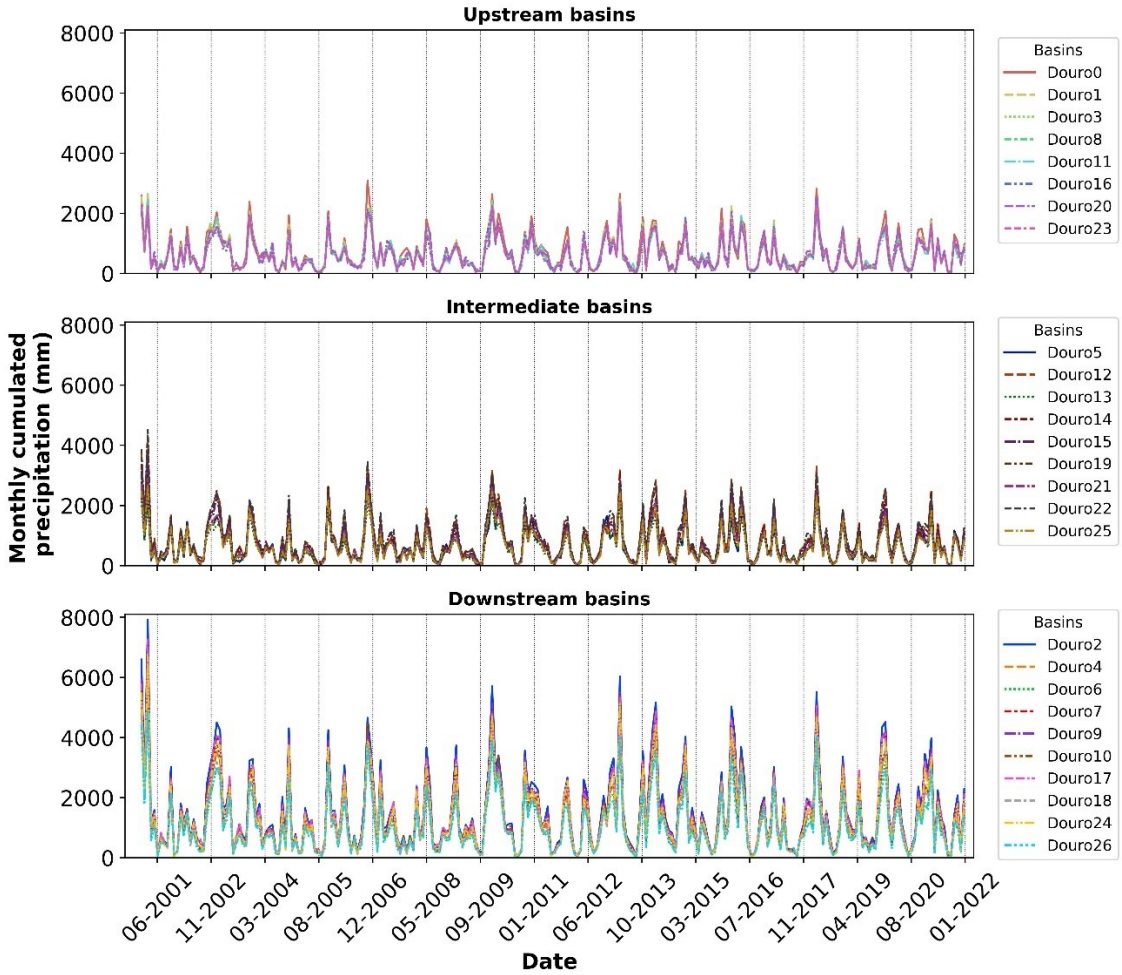


Figure 6-3. Monthly cumulated precipitation in the Douro River sub-basins.

River discharge in Crestuma was provided by the Portuguese Environmental Agency (APA) and is available at the National Information System on Water Resources (SNIRH) at station 07G/01A (Figure 6-4). We used all the available fluvial discharge information that overlapped on the same dates as the precipitation data. Furthermore, the data fed into the LSTM models were filtered to remove part of the data unrelated to floods, thereby reducing the training costs of the model. To accomplish this task, a threshold of 800 m^3/s was defined. The null values in the time series are not "natural", being generated due to the operation of the Crestuma-Lever dam. This means that normal flows in the estuarine region will be difficult to forecast owing to the protocols of the dam, which depend on electricity demand. However, flood events are conditioned by natural processes, even considering the dam operation, and thus, there is a dependence between precipitation and river flow. This means that the constructed model will focus on forecasting flood events. Uncertainties will be higher if the models are applied under normal conditions.

When crossing the precipitation data with the discharge data, we can see that the discharge peaks occur on the same dates as the intense rainfall measurements, which demonstrates that precipitation is one of the main flood drivers in the Douro Estuary. Precipitation in the downstream basins is of higher significance in terms of flood events when compared with the precipitation in other regions of the Douro basins.

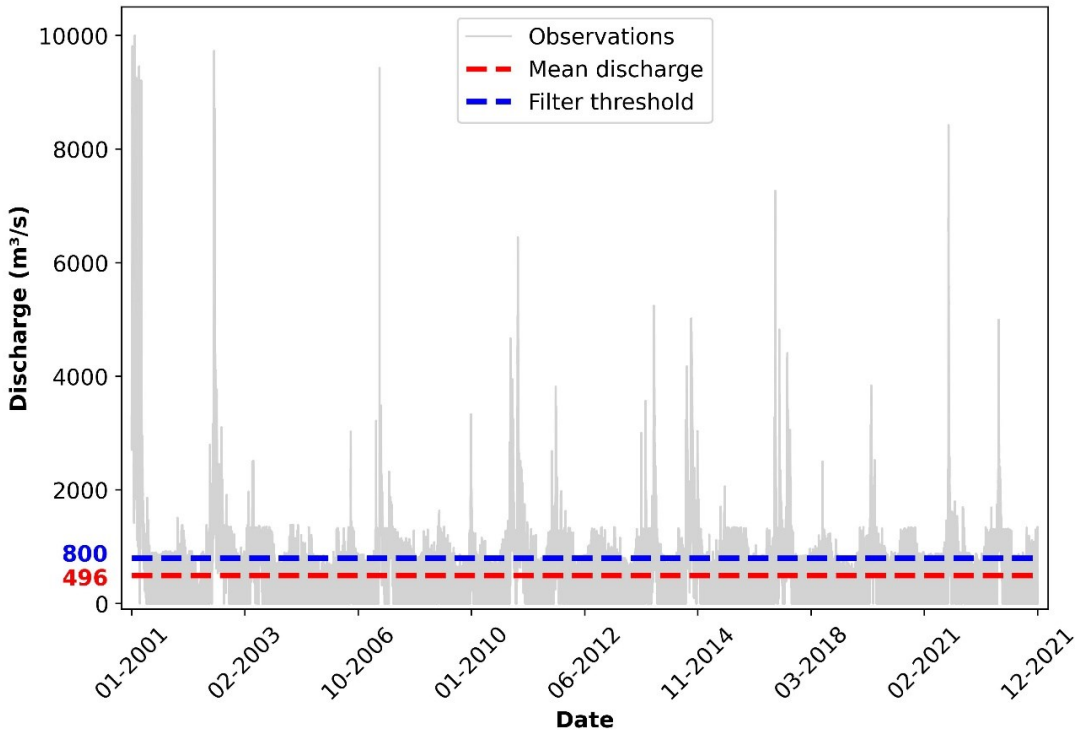


Figure 6-4. Historical data of the Douro River discharge at the Crestuma-Lever dam.

The hydrodynamic numerical model also used bathymetric data from previous field campaigns (Bastos et al., 2012; Weber de Melo et al., 2022). Additionally, water level data from three monitoring stations in Douro Estuary were used to calibrate the numerical model. These stations were located in Cais dos Banhos (SNIRH station), Cantareira, and Afurada (Portuguese Hydrographic Institute (IH) stations).

6.3.2. Implementation of the data model

LSTM networks were designed to solve the problem of recurrent neural networks that present vanishing gradients during backpropagation and inferior performance in remembering long sequences (Aggarwal, 2018). Each LSTM cell has three gates: the input gate (i_t) determines the values of the input that are used to update the state cell (6-1), the forget gate (f_t) determines the values of the previous cell state that should be forgotten (6-2), and the output gate (o_t) determines the values of the cell state that will be used as output (6-3). During the execution of an LSTM layer, a candidate cell state ($\bar{C}_t$) is determined, which are the candidate values that can be added to the cell state(6-4). The cell state (C_t) is then computed by combining the previous cell state with the candidate cell state and the forget gate (6-5). Lastly, the output of the LSTM cell (also known as the hidden state h_t) is determined by the element-wise multiplication of the output gate by the cell state (6-6).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (6\text{-}1)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (6\text{-}2)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6\text{-}3)$$

$$\bar{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (6\text{-}4)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \bar{C}_t \quad (6\text{-}5)$$

$$h_t = o_t \odot \tanh(C_t) \quad (6\text{-}6)$$

where W_i, W_f, W_o, and W_c are the weight vectors of the input gate, forget gate, output gate, and candidate cell state, respectively, b_i, b_f, b_o, and b_c are the bias vectors of the input gate, forget gate, output gate, and candidate cell state, respectively, h_{t-1} is the previous hidden state, x_t is the current input, σ is the activation function, and $\odot$ represents the element-wise multiplication.

In this study, LSTM neural networks were used to forecast fluvial discharge in the Crestuma-Lever dam. The networks were implemented in Python using the TensorFlow library (Abadi et al.2016). This library provides a broad spectrum of DL solutions, allowing full customisation of the models, and it was selected owing to its versatility and modelling capacities. The models were implemented using the Functional API, which permits the model to process each input using different operations and layers with different hyperparameters.

Two architectures were evaluated: one using only LSTM layers, and the other using attention layers after the LSTM layer (Figure 6-5). Hereafter, the data model with only LSTM layers will be referred to as LSTM and the model with attention layers will be referred to as ATT. Attention layers have already been demonstrated to improve the long-term forecasting capacity of LSTM networks, which justifies the choice of this type of architecture (Granata et al., 2022; Vaswani et al.2017).

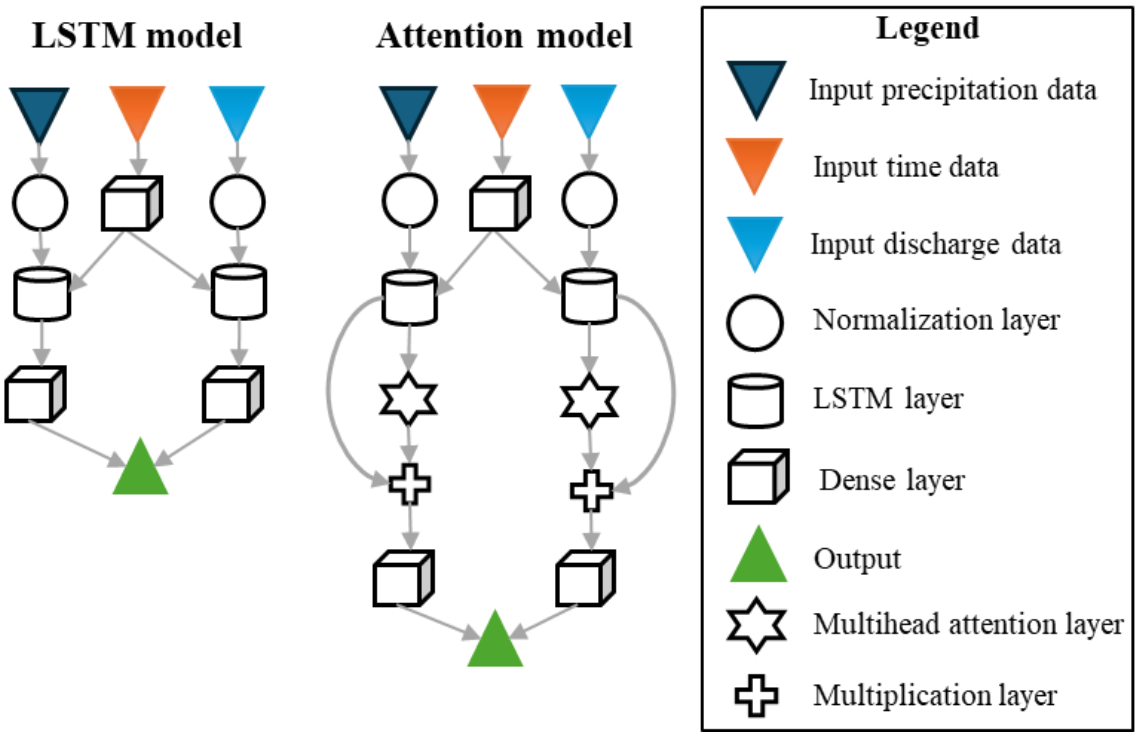


Figure 6-5. Architecture of the implemented data models.

Each input was processed differently inside the network architecture, resulting in two main branches according to each input variable. The precipitation input was calculated by dividing the cumulative precipitation over the last five days by the distance between the basins' centroid and the Crestuma Lever dam. The input time data were the month, day, and hour of each observation, and the values were divided by 12, 31, and 24, respectively, to normalise the data. The input discharge was used without transformation. Hence, the precipitation, time, and river discharge input vectors have lengths of 27, 3, and 1, respectively.

The time data were concatenated with the other variables to provide the LSTM layers with the time dimension. The LSTM layers used the ReLU function (Hara et al., 2015) for activation and the hard sigmoid function for recurrent activation. The number of units in the precipitation and discharge branches was 40 and 10, respectively. The dense layers of both branches also used the ReLU function for activation and 20 and 15 units, respectively. Finally, the output layer was a single-unit dense layer with ReLU activation connected to both branches.

The multi-head attention layers used 20 and 54 heads for the discharge and precipitation branches, respectively. The key and value dimensions were 15 for both branches. These values were obtained iteratively during the training phase of the model.

The models were trained using 80% of the data and tested at 20%. The Nadam optimiser (Dozat, 2016) was used with a learning rate of 0.0005. The choice of the loss function was the MSE due to the emphasis on predicting extreme events, as it is more sensitive to outliers. The model was trained using 100 epochs, a batch size of 48, and an early stop criterion after three sequential epochs, with no improvement in the loss of the test dataset. Furthermore, six different data models were implemented, each trained to forecast the river discharge with time windows of 3 h, 6 h, 12 h, 18 h, 24 h, and 36 h in advance, resulting in six models for each architecture. The models were trained sequentially, starting with those with lower prediction time horizons and finalising with the highest.

6.3.3. Implementation of the numerical model

The hydrodynamic process-based model was implemented using the D3D software. This software has been widely used to simulate hydrodynamic and morphodynamic patterns in rivers, estuaries, and coastal areas (Des et al., 2019; Elmilady et al., 2022; Iglesias et al., 2019; Paprotny et al., 2020; Weber de Melo et al., 2022). Using the finite-volume method, the software solves the time-averaged Navier-Stokes equations for incompressible fluids, adopting shallow water equations and Bussinesq assumptions (Agoshkov et al., 1993; Deltares, 2018).

The model's domain was bounded by an elevation of 4 m above the mean sea level. It was configured with two open boundaries, one in the Atlantic Ocean forced by tidal harmonic constituents and the other in the Crestuma-Lever dam, forced by hourly river discharges. The model had 213 grid points in the M direction and 108 points in the N direction. Most grid cells near the boundaries were elongated in the longitudinal direction. The model resolution varied from 20 m × 20 m in the estuarine area near the breakwater, 100 m × 400 m in the Atlantic Ocean boundary, and 7 m × 120 m in the fluvial boundary. The differences in the grid cell dimensions optimised the simulation to ensure that the areas of interest inside the estuary had an adequate resolution, but, at the same time, reducing the storage and processing requirements of the simulation.

This version of the D3D model is an improvement of that implemented by Weber de Melo et al. (2022). The grid was redesigned and a new bathymetry was added according to the new grid. The high-resolution zone in the estuary was expanded owing to the availability of new monitoring stations at Cantareira and Afurada, increasing the reliability of the model in this area.

The model was calibrated using OpenDA, which automatically determines the values of the model parameters (Garcia et al., 2015; Melo et al., 2020). It consists of running the model several times until it finds the parameter values that result in minimum error. A cost function is computed at each iteration by considering the model's error. The optimal result combines the parameter values that minimise the cost function. The calibration parameters for this study were the Manning coefficients in the u and v directions and the harmonic constituents A0, M2, and S2. The calibration procedure used data between 2 and 17 of March2022 because of the availability of water level data at the three monitoring stations located in Afurada, Cantareira, and Cais dos Banhos, considering an entire spring-neap tidal cycle.

Additionally, for the calibration procedure, it is advisable to define the bounds in which the values of the parameters can vary. The upper and lower limits may vary according to the user experience. For instance, it is convenient to set a lower limit that ensures that the Manning coefficient will always be higher than zero to guarantee that the numerical model is only simulated using values based on reality. Otherwise, OpenDA may use inadequate values, resulting in numerical errors during model calibration or in finding solutions that are only applicable to the calibration data. In the case of tidal harmonic constituents, the bounds may vary in intervals that maintain the hierarchy between the most and least important constituents in the study area.

6.3.4. FEWS simulation and error metrics

A flood period that occurred in January 2016 was selected to assess the performance of Douro FEWS. This period was selected due to the data availability at the monitoring station of Cais dos Banhos. This experiment consisted of forecasting the discharge at the fluvial boundary of the numerical model with the AI models, and then performing a simulation with the predicted values to forecast the water levels at the estuary. In this scenario, all fluvial discharge values were considered, including those lower than 800 m^3/s, to verify the performance of the model under normal conditions and to process the data flow from the source to the model without any additional pre-processing requirements. Thus, it was possible to evaluate the generalisation capacity of the data models. Subsequently, the error at Cais dos Banhos was determined.

To assess the performance of the AI models and system, RMSE, (6-7), mean absolute percentage error (MAPE, (6-8)), and Kling Gupta (KGE, (6-9) efficiency (Gupta et al., 2009) were computed using the HydroEval library (Hallouin, 2021). In addition to the classic coefficients such as RMSE and MAPE, KGE is also a metric suitable for comparing observations with simulations. It has the advantage of simultaneously considering correlation, variability, and bias. KGE equal to one indicate a perfect match.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_o - y_s)^2} \quad (6\text{-}7)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_o - y_s|}{y_o} \quad (6\text{-}8)$$

$$KGE = 1 - \sqrt{(\rho - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (6\text{-}9)$$

$$\alpha = \frac{\sigma_s}{\sigma_o} \quad (6\text{-}10)$$

$$\beta = \frac{\mu_s}{\mu_o} \quad (6\text{-}11)$$

where y_o is the observed value, y_s is the simulated value, ρ is the Pearson correlation coefficient, α (6-10) is the variability term, β (6-11) is the bias term, σ_s is the variance of simulated values, σ_o is the variance of the observed values, μ_s is the mean of the simulated values, and μ_o is the mean of the observed values.

6.4. Results

This section presents the main outcomes of the study. First, the performance of the AI model is discussed, followed by the calibration of the numerical model results, and a simulation of the Douro FEWS for the January 2016 flood.

6.4.1. Performance of the data models

The error metrics of the data models for the testing dataset are presented in Table 6-1. It was possible to verify that the models with a lower time window performed better, as expected. The higher the time window, the higher the uncertainty between the forecasting instant and the occurrence of an event, which justifies this performance pattern. Additionally, the attention layer did not significantly improve the prediction performance. Although the ATT models presented a lower RMSE considering time windows of 3 h to 12h, the LSTM models outperformed the ATT models considering the KGE criteria. This probably occurred because the LSTM models had lower variability when comparing the α term. Both models exhibited similar performance for the other time windows. The obtained results indicated that the use of attention layers did not significantly improve the performance of the LSTM models, resulting in error metrics with minor differences.

Table 6-1. Performance metrics of the data models.

Model	RMSE (m^3/s)	MAPE	KGE	ρ	α	β
LSTM - 3h	237.26	0.143	0.95	0.950	1.007	1.016
ATT - 3h	218.84	0.144	0.92	0.957	0.939	1.029

LSTM - 6h	281.37	0.160	0.92	0.928	0.990	1.025
ATT - 6h	269.82	0.161	0.89	0.933	0.920	1.032
LSTM - 12h	344.54	0.173	0.87	0.888	0.938	1.028
ATT - 12h	340.58	0.174	0.83	0.890	0.870	1.028
LSTM - 18h	395.40	0.186	0.82	0.850	0.906	1.032
ATT - 18h	394.55	0.190	0.81	0.852	0.897	1.049
LSTM - 24h	444.58	0.200	0.76	0.808	0.873	1.051
ATT - 24h	453.64	0.210	0.77	0.806	0.893	1.075
LSTM - 36h	534.28	0.239	0.66	0.719	0.836	1.087
ATT - 36h	534.42	0.222	0.65	0.711	0.823	1.067

The density plot of the differences between the observed data and simulated results was computed (Figure 6-6). This graph illustrates the behaviour of the models that were explained by analysing the error metrics. The LSTM models outperformed the ATT models for the time windows of 3 h, 6 h, and 12 h, as the peaks of their density plots were closer to 0. When comparing models with the same forecasting horizons, the density plots of the ATT models were higher than those of the LSTM density plots. This behaviour indicated that the error variability in the ATT models was lower than that in the LSTM models. Furthermore, this graph shows that the models tended to underestimate the observed values, which was expected because all the β parameters were higher than 1, indicating that the means of the simulated values were slightly higher than the observed values.

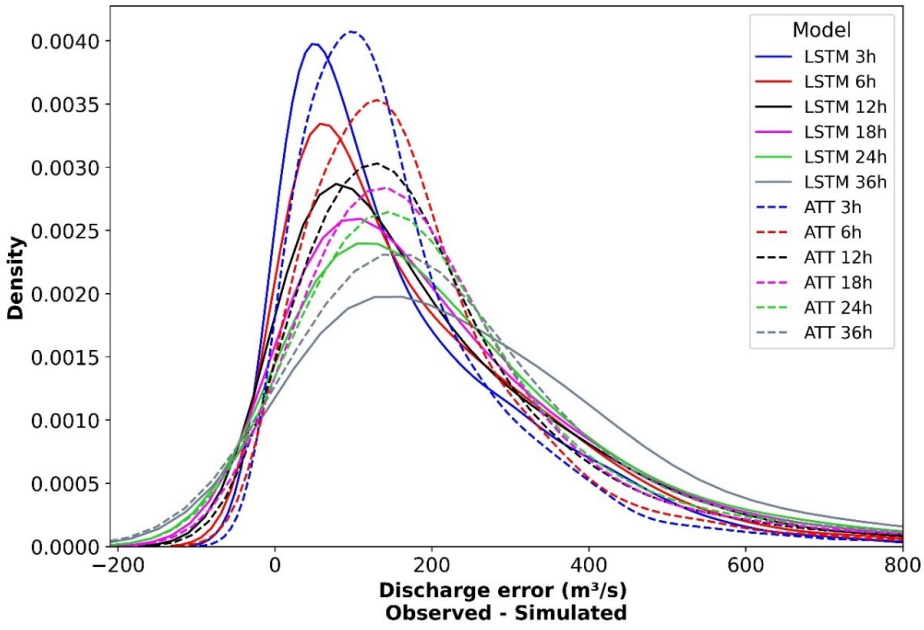


Figure 6-6. Error distribution of data models.

In addition, the comparison of the error distribution between the LSTM and ATT models shows that the lower the time window, the narrower the density plot, indicating a lower variability in the predicted

values and, consequently, lower errors. This behaviour is expected due to the higher uncertainties when the forecasting window increases.

Finally, the data models were used to forecast the flow discharge in the Crestuma-Lever dam during a flood event that occurred in January 2016. Figure 6-7 shows that all LSTM models accurately simulated flows of up to 3500 m^3/s, with slight deviations from the observed values. Below this threshold, the model results diverged, with the lower time window models exhibiting superior performance. For the ATT models, the same behaviour was observed for discharges higher than 6000 m^3/s. During the peak period, the performance of the models with a shorter time window was better, following the expected behaviour. This result confirmed that the ATT models presented lower variability, as already represented in the density plot of the errors.

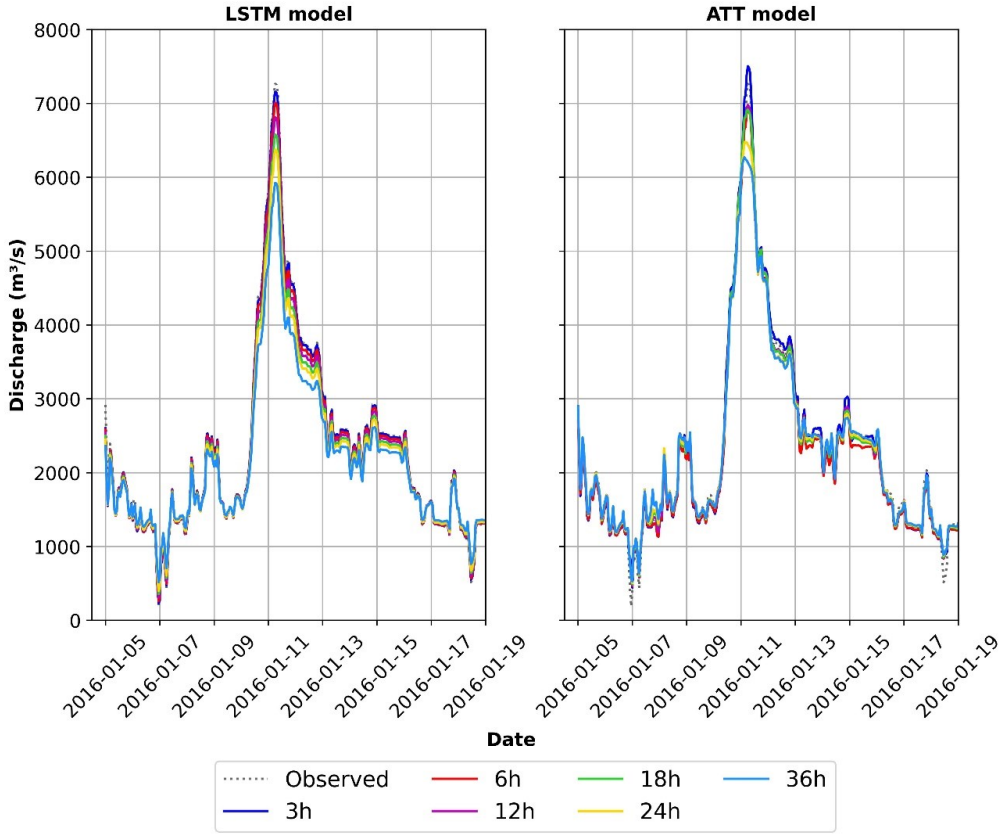


Figure 6-7. Flow forecast at the Crestuma-Lever dam during the January 2016 flood.

When assessing the performance of the models to forecast the peak discharge, the LSTM models showed inferior performance with time windows higher than 18 h, whilst the ATT model's performance worsened for 24 h and 36 h, indicating that the attention layer improved the forecasting window. The results showed that the LSTM models demonstrated superior performance compared to the ATT models in predicting values below 800 m^3/s. This suggests that the LSTM models are well suited for forecasting peak discharges and average values. In summary, the attention layers reduced the variability of the

model's predictions and enhanced their capacity to forecast extreme flows, mainly for models with higher forecasting windows. However, it also worsened the performance of the model for values below the discharge threshold used during training.

6.4.2. Calibration of the process-based model

The calibration of the numerical model with OpenDA (Figure 6-8) showed good agreement between the observed data and simulated results, achieving RMSE values closer to 0.15 m. The water levels in the stations varied between 0.5 m and 4.1 m. This improved the last version of the Douro model (Weber de Melo et al., 2022), which achieved an RMSE of 0.199 m using only the monitoring station in Cais dos Banhos. This version not only reduced the RMSE in Cais dos Banhos to 0.161 m but was also calibrated with two more monitoring stations, thereby increasing the reliability of the estuarine model. The final values of the parameters are presented in Table 6-2.

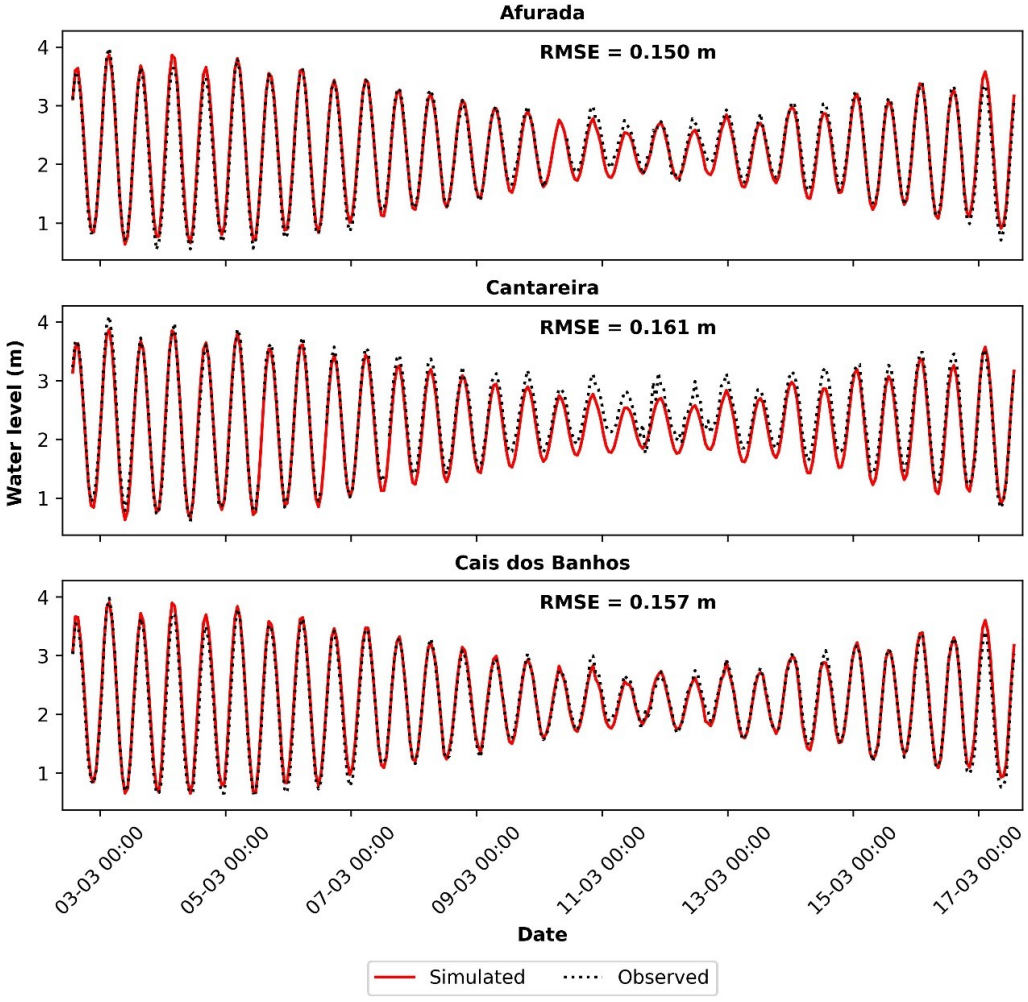


Figure 6-8. Calibration results of the hydrodynamic numerical model.

Table 6-2. Numerical model calibrated parameters.

Parameter	Manning - U direction ($m^{-1/3}/s$)	Manning - V direction ($m^{-1/3}/s$)	A0 (m)	M2 amplitude (m)	M2 phase angle (degree)	S2 amplitude (m)	S2 phase angle (degree)
Value	0.0439	0.506	2.20	1.024	64.555	0.376	91.626

6.4.3. Performance of the FEWS platform

The FEWS platform was assessed by computing the error metrics at the monitoring station in Cais dos Banhos. Figure 6-9 shows the water level resulting from applying the different models results at the fluvial boundary during the peak period. The acronym GD represents the simulation performed using ground data at the fluvial boundary for comparison with the results obtained with the data models. It can be observed that all simulations underestimated the water level before the peak, which indicates a systematic error in the hydrodynamic numerical model. This occurred because the model was calibrated for average conditions, and in this application, it was considered a flood situation. A previous study suggested that different values of the roughness coefficient should be considered to properly represent the conditions in the Douro Estuary for different scenarios (Iglesias et al., 2019). Despite this, all models forecasted similar values compared to the GD results.

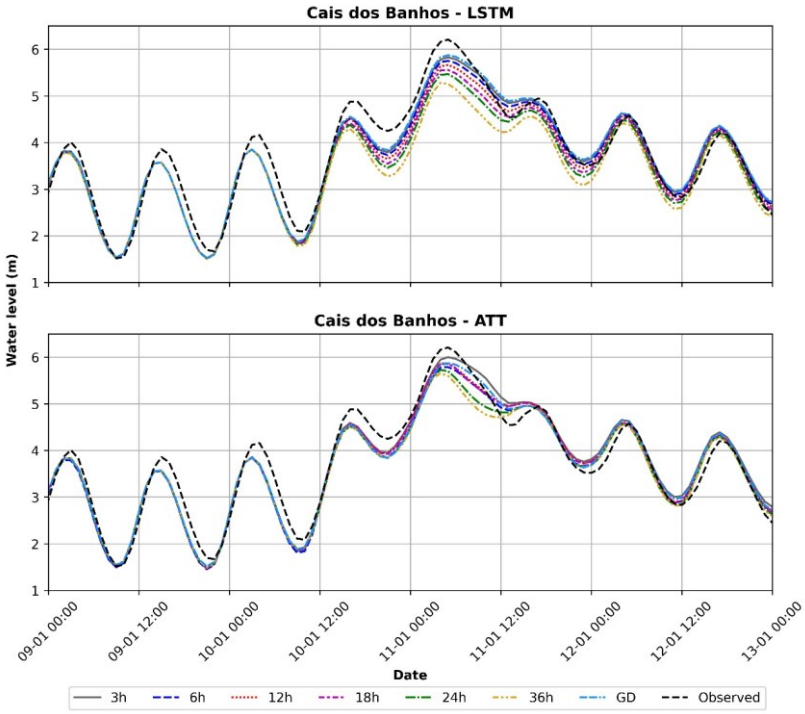


Figure 6-9. Water level simulation results at Cais dos Banhos using the LSTM and ATT models to generate discharge values at the fluvial boundary.

Moreover, higher variability can be observed in the results of the LSTM models compared to the ATT models, reflecting the behaviour of the discharge predictions and indicating that the attention layers increased the forecasting horizon of the models. Indeed, all simulations using the ATT model results as boundary conditions had similar metrics, even 36 hours in advance. The error metrics of the Douro FEWS simulation are presented in Table 6-3. All ATT model results overlap for water levels lower than 4 m, indicating an almost perfect agreement. The errors increase during the peak period, in which a higher variability can be detected.

Table 6-3. Performance comparison between LSTM and ATT models for forecasting water levels in Cais dos Banhos.

Boundary model	RMSE (m^3/s)	MAPE	KGE	ρ	α	β
LSTM - 3h	0.233	0.072	0.919	0.973	0.925	1.014
ATT - 3h	0.232	0.072	0.952	0.972	0.965	1.018
LSTM - 6h	0.233	0.071	0.907	0.973	0.912	1.010
ATT - 6h	0.222	0.067	0.926	0.974	0.931	1.006
LSTM - 12h	0.238	0.071	0.890	0.973	0.894	1.005
ATT - 12h	0.223	0.069	0.930	0.975	0.936	1.015
LSTM - 18h	0.248	0.072	0.870	0.971	0.873	0.999
ATT - 18h	0.221	0.068	0.929	0.975	0.935	1.014
LSTM - 24h	0.262	0.073	0.852	0.970	0.856	0.993
ATT - 24h	0.230	0.069	0.901	0.974	0.905	1.011
LSTM - 36h	0.300	0.077	0.814	0.963	0.819	0.983
ATT - 36h	0.238	0.070	0.891	0.973	0.895	1.010

6.5. Discussion

The DL models yielded satisfactory results, achieving low-error metrics when forecasting fluvial discharges. The LSTM models performed better in the validation dataset, particularly for time windows lower than 12 h. However, the ATT models outperformed the LSTM models when evaluated in the FEWS simulation, achieving similar metrics for all the considered time horizons. This indicates that the attention layer improved the capacity of the model to anticipate flood events and increased its predictive capacity to values closer to 6000 m^3/s. Equivalent results were obtained by Fan et al. (2019) and Wang et al. (2023), who also demonstrated that attention layers improved the performance of their time-series forecasting models.

The improvement in the forecasting capacity of the ATT models is justified by the higher complexity of the validation dataset in which the error metrics were computed. The validation period had a peak discharge of more than 8000 m^3/s, which was overcome only four times in the historical data available for this study. The better performance during the FEWS testing is explained by the peak

discharge being slightly higher than 7000 m^3/s. DL models learn according to the data history; hence, their performance is reduced when forecasting scarce extreme events. Therefore, it can be concluded that the DL model using LSTM layers is reliable for forecasting extreme river discharges. However, its performance varies according to the frequency of the events. The rarer the event, the lower the prediction accuracy.

Regarding the time horizon capacity of the FEWS, it forecasted the January 2016 flood event with the same accuracy, considering 3 h and 36 h in advance using the ATT model architecture. More data with high-intensity events are necessary to reduce error metrics and improve the prediction capacity in extreme situations, thereby providing a more robust river discharge predictor. However, this lack of data is a barrier to the implementation of data models and numerical models.

Moreover, modelling discharges from a dam was one of the challenges overcome in this study. Modelling flows that only depend on natural parameters such as bathymetry, river discharge, and precipitation is easier than forecasting discharges associated with dam operations. The dam's discharge relies not only on the commented variables, but also on energy production and demand, population and industrial growth, and reservoir capacity. Hence, the considered river discharge historical data hide all these parameters and conditions that also evolved over time. Despite this fact, the AI models achieved acceptable error metrics, demonstrating the versatility of the data models in extremely complex situations. Undertaking this task would be more challenging if process-based models were utilised.

The numerical model used in this study had satisfactory metrics during calibration, although a systematic error was observed during the FEWS simulation. This error occurred because the model was calibrated using two weeks of data representing average conditions. The precision of parameter values, such as the Manning coefficient, is more important under flood scenarios than under average conditions. Therefore, calibrating the model exclusively for flood scenarios can reduce the prediction error. Conversely, it could also increase the errors during the average conditions. Thus, two sets of parameters values would ensure low errors in both situations.

Furthermore, using automatic calibration tools, such as *OpenDA*, helped reduce the time necessary to achieve a good result and guarantee a good agreement between the model results and observed data. This procedure allows multiple parameters to be calibrated simultaneously, reducing the human resources necessary to perform this task manually.

Regarding the platform, the ATT model prediction was reliable until a fluvial discharge of 6000 m^3/s. The uncertainties increase for values higher than this, reducing the accuracy of the forecasts. Nevertheless, the model can still identify extreme flood discharges, but its intensity is undervalued. The

numerical model component must be calibrated using extreme event data to reduce the differences between the observed and simulated water levels.

The implementation of real-time FEWS based on ERA5 data relies on data forecasts and availability, which are not instantaneous. ERA5 data is updated twice a day, with forecasts for the next 18 h, which would limit the application of the proposed methodology. Despite this limitation, ERA5 was selected because this widely used database demonstrated to produce accurate precipitation results over the Iberian Peninsula, including extreme events (Gomis-Cebolla et al., 2023; Hénin et al., 2018). Other data that could also be selected for the FEWS construction are the precipitation from the Global Forecast System (GFS) and the Weather Research and Forecast (WRF) models, which forecast precipitation data for the next 16 days and 3 days, respectively, or even results from specific models from the national meteorological agencies, if available. Future studies should assess the quality and uncertainties in the estuarine water levels when considering different precipitation forecast datasets.

6.6. Conclusions

This study combined a DL and process-based models to implement an FEWS for the Douro River estuary. The DL model combined precipitation data from the ECMWF with river discharge data from the SNIRH to forecast discharges at the Crestuma-Lever dam, considering different time windows. The hydrodynamic numerical model used the forecasted river discharges to simulate future water levels in the Douro River estuary.

We successfully implemented an early warning system for the Douro River estuary that can anticipate water levels for 36 h in advance. The combination of AI and numerical models makes it possible to create complex and reliable systems. Douro FEWS can increase the region's resilience against extreme events, reducing losses and saving lives, contributing to the accomplishment of the United Nations sustainable development goal (SDG) 3.d, which aims to strengthen the capacity for early warnings in all countries, SDG 11 which highlights the importance of disaster risk reduction, and SDG 13, which focuses on adaptation and mitigation actions against climate change.

Future research may explore the uncertainties associated with the platform, given that errors in the data are propagated through the DL model's discharge forecasts. Additionally, mapping the most vulnerable regions of the estuary and crossing them with flood maps from the numerical model could increase the usefulness of the platform, allowing for a fast response from emergency teams during flood events.

6.7. References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., ...

Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. http://arxiv.org/abs/1603.04467

Aggarwal, C. C. (2018). Neural Networks and Deep Learning. In Artificial Intelligence. Springer International Publishing. https://doi.org/10.1007/978-3-319-94463-0

Agoshkov, V. I., Ambrosi, D., Pennati, V., Quarteroni, A., & Saleri, F. (1993). Mathematical and numerical modelling of shallow water flow. Computational Mechanics, 11(5–6), 280–299. https://doi.org/10.1007/BF00350088

Alcoforado, M. J., Silva, L. P., Amorim, I., Fragoso, M., & Garcia, J. C. (2021). Historical floods of the Douro River in Porto, Portugal (1727–1799). Climatic Change, 165(1–2). https://doi.org/10.1007/s10584-021-03039-7

APA, A. P. do A. (2014). Relatório de Caracterização: Região Hidrográfica do Douro (RH3). https://apambiente.pt/sites/default/files/_Agua/DRH/ParticipacaoPublica/PGRH/2016-2021/2_Fase/PGRH_2_RH3_Art5.pdf

Bastos, L., Bio, A., Pinho, J. L. da S., Granja, H., & Jorge da Silva, A. (2012). Dynamics of the Douro estuary sand spit before and after breakwater construction. Estuarine, Coastal and Shelf Science, 109, 53–69. https://doi.org/10.1016/j.ecss.2012.05.017

C3S, C. C. C. S. (2024). ERA5 Data Documentation. https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation

Deltares. (2018). Delft3D-Flow: Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments - User Manual.

Des, M., DeCastro, M., Sousa, M. C., Dias, J. M., & Gómez-Gesteira, M. (2019). Hydrodynamics of river plume intrusion into an adjacent estuary: The Minho River and Ria de Vigo. Journal of Marine Systems, 189(May 2018), 87–97. https://doi.org/10.1016/j.jmarsys.2018.10.003

Dozat, T. (2016). Incorporating Nesterov Momentum into Adam. International Conference on Learning Representations. https://cs229.stanford.edu/proj2015/054_report.pdf

Elmilady, H., van der Wegen, M., Roelvink, D., & van der Spek, A. (2022). Modeling the Morphodynamic Response of Estuarine Intertidal Shoals to Sea-Level Rise. Journal of Geophysical Research: Earth Surface, 127(1), 1–26. https://doi.org/10.1029/2021JF006152

Fan, C., Zhang, Y., Pan, Y., Li, X., Zhang, C., Yuan, R., Wu, D., Wang, W., Pei, J., & Huang, H. (2019). Multi-horizon time series forecasting with temporal attention learning. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2527–2535. https://doi.org/10.1145/3292500.3330662

Garcia, M., Ramirez, I., Verlaan, M., & Castillo, J. (2015). Application of a three-dimensional hydrodynamic model for San Quintin Bay, B.C., Mexico. Validation and calibration using OpenDA. Journal of Computational and Applied Mathematics, 273, 428–437. https://doi.org/10.1016/J.CAM.2014.05.003

Gomis-Cebolla, J., Rattayova, V., Salazar-Galán, S., & Francés, F. (2023). Evaluation of ERA5 and ERA5-Land reanalysis precipitation datasets over Spain (1951–2020). Atmospheric Research, 284, 106606. https://doi.org/10.1016/j.atmosres.2023.106606

Granata, F., Di Nunno, F., & de Marinis, G. (2022). Stacked machine learning algorithms and bidirectional long short-term memory networks for multi-step ahead streamflow forecasting: A comparative study. Journal of Hydrology, 613. https://doi.org/10.1016/j.jhydrol.2022.128431

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hallouin, T. (2021). HydroEval: Streamflow Simulations Evaluator (Version 0.1.0). Zenodo.

Hara, K., Saito, D., & Shouno, H. (2015). Analysis of function of rectified linear unit used in deep learning. 2015 International Joint Conference on Neural Networks (IJCNN), 2015-Septe, 1–8. https://doi.org/10.1109/IJCNN.2015.7280578

Hénin, R., Liberato, M., Ramos, A., & Gouveia, C. (2018). Assessing the Use of Satellite-Based Estimates and High-Resolution Precipitation Datasets for the Study of Extreme Precipitation Events over the Iberian Peninsula. Water, 10(11), 1688. https://doi.org/10.3390/w10111688

Iglesias, I., Venâncio, S., Pinho, J. L., Avilez-Valente, P., & Vieira, J. M. P. (2019). Two Models Solutions for the Douro Estuary: Flood Risk Assessment and Breakwater Effects. Estuaries and Coasts, 42(2), 348–364. https://doi.org/10.1007/s12237-018-0477-5

Jevrejeva, S., Jackson, L. P., Grinsted, A., Lincke, D., & Marzeion, B. (2018). Flood damage costs under the sea level rise with warming of 1.5◦C and 2◦C. Environmental Research Letters, 13(7). https://doi.org/10.1088/1748-9326/aacc76

Kim, H. Il, & Kim, B. H. (2020). Flood Hazard Rating Prediction for Urban Areas Using Random Forest and LSTM. KSCE Journal of Civil Engineering, 24(12), 3884–3896. https://doi.org/10.1007/s12205-020-0951-z

Liu, Y., Yang, Y., Chin, R. J., Wang, C., & Wang, C. (2023). Long Short-Term Memory (LSTM) Based Model for Flood Forecasting in Xiangjiang River. KSCE Journal of Civil Engineering, 27(11), 5030–5040. https://doi.org/10.1007/s12205-023-2469-7

Melo, W., Pinho, J., Iglesias, I., Bio, A., Avilez-Valente, P., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2020). Hydro- and Morphodynamic Impacts of Sea Level Rise: The Minho Estuary Case Study. Journal of Marine Science and Engineering, 8(6), 441. https://doi.org/10.3390/jmse8060441

Monteiro, A., Madureira, H., Fonseca, L., & Gonçalves, P. (2017). Plano Metropolitano de Adaptação às Alterações Climáticas. https://cdn.bndlyr.com/stljpltqzpxcsfeb/_assets/documents/2020/01/08/AMP_PMAAC_I6C3uzl.pdf

Muñoz Sabater, J. (2019). ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac

Paprotny, D., Vousdoukas, M. I., Morales-Nápoles, O., Jonkman, S. N., & Feyen, L. (2020). Pan-European hydrodynamic models and their ability to identify compound floods. Natural Hazards, 101(3), 933–957. https://doi.org/10.1007/s11069-020-03902-3

Ramos, A. M., Cortesi, N., & Trigo, R. M. (2014). Circulation weather types and spatial variability of daily precipitation in the Iberian Peninsula. Frontiers in Earth Science, 2. https://doi.org/10.3389/feart.2014.00025

Rentschler, J., Salhab, M., & Jafino, B. A. (2022). Flood exposure and poverty in 188 countries. Nature Communications, 13(1). https://doi.org/10.1038/s41467-022-30727-4

Seshan, S., Vries, D., Immink, J., van der Helm, A., & Poinapen, J. (2024). LSTM-based autoencoder models for real-time quality control of wastewater treatment sensor data. Journal of Hydroinformatics. https://doi.org/10.2166/hydro.2024.167

UNDRR, U. N. O. for D. R. R. (2023, January 19). Sendai Framework for Disaster Risk Reduction: Goal, Targets and Metrics. https://unfccc.int/documents/628865

Venâncio, S. D. S., Pinho, J. L., Vieira, J. M., Avilez-Valente, P., & Iglesias, I. (2019). Analysis of estuarine flood levels based on numerical modelling. The Douro river estuary case study. Revista Eletrônica Em Gestão, Educação e Tecnologia Ambiental, 23, 14. https://doi.org/10.5902/2236117038538

Wang, Y., Wang, W., Chau, K., Xu, D., Zang, H., Liu, C., & Ma, Q. (2023). A new stable and interpretable flood forecasting model combining multi-head attention mechanism and multiple linear regression. Journal of Hydroinformatics, 25(6), 2561–2588. https://doi.org/10.2166/hydro.2023.160

Weber de Melo, W., Pinho, J., Iglesias, I., Bio, A., Avilez-Valente, P., Vieira, J., Bastos, L., & Veloso-Gomes, F. (2022). Flood Risk Assessment at the Douro River Estuary (pp. 37–49). https://doi.org/10.1007/978-981-16-5501-2_4

WMO, W. M. O. (2022). Early Warnings for all - The UN Global Early Warning Initiative for the Implementation of Climate Adaptation - Executive Action Plan 2023-2027. https://library.wmo.int/idurl/4/58209

CHAPTER 7.

Interpreting machine learning-based beach profile evolution models with explainable artificial intelligence tools.

Article submitted in July of 2024 to the Journal of Marine Science and Engineering.

Weber de Melo, W., Iglesias, I. and Pinho, J. (2024). Interpreting machine learning-based beach profile evolution models with explainable artificial intelligence tools. JMSE

"If you're teaching today what you were teaching five years ago, either the field is dead or you are."

Noam Chomsky(1928 -) – North American professor and philosopher.

Chapter 7. Interpreting machine learning-based beach profile evolution models with explainable artificial intelligence tools

Abstract

Beach profile evolution is a complex process influenced by the interaction of sediments with waves and winds. Understanding this process is essential for protecting coastal communities and optimising resource management in these regions. Artificial intelligence (AI) models offer alternatives to traditional numerical models. However, the limitation of interpretation and deep under-standing of such models are barriers that reduce their wider application. This study applied three explainable artificial intelligence (XAI) methods to demonstrate how to interpret AI models: partial dependence (PD) plots, permutation importance (PIMP), and the local interpretable mod-el-agnostic explanation (LIME) algorithm. Two Random Forest (RF) models were implemented to simulate the morphological evolution of beach profiles along the northern coast of Portugal. The models were trained using profile measurements, wave action, and sediment size data from nine profiles on five beaches. These XAI methods provide valuable insights into the influence of input variables on the predictions and sensitivity of the outputs to each input variable, contributing to deconstructing the myth that artificial intelligent models are black boxes. It was demonstrated that AI models can be as reliable as process-based models, however adequate techniques are funda-mental for increasing confidence in this type of solution.

7.1. Introduction

Climate change is one of the more significant threats to the delicate balance of coastal environments due to rising sea levels and the increasing frequency and intensity of ex-treme events [1, 2]. These climate change consequences rise the exposure of coastal com-munities to a range of hazards, including coastal erosion, flooding, saltwater intrusion, and storm surges [3]. This jeopardises the livelihoods and well-being of coastal popula-tions and disrupts crucial ecosystem services such as fisheries, tourism, and industrial and leisure activities [1, 4].

Proactive adaptation strategies involving integrated coastal zone management, in-vestment in resilient infrastructure, and community-based approaches are urgently needed to mitigate these risks and ensure the resilience of coastal communities. Using models that can simulate long- and short-term coastal morpho-hydrodynamics is crucial. Process-based models have been widely used to study coastal morphological evolution, having achieved excellent results [5–8]. However, these models are limited by the availa-bility of computational resources, namely processing power and storage capacity, which hinders the wider application of such solutions, particularly when high-resolution mod-els are required [9].

To overcome those limitations, artificial intelligence (AI) models are becoming in-creasingly popular. Instead of solving differential equations that describe the hydro- and morphodynamic variables and their interactions, AI models relate input and output data using relatively simpler mathematical operations with lower computational power re-quirements [10]. The application of AI tools to model coastal environments presents a transformative potential for the representation and forecast of coastal processes [11]. In addition, AI tools can be used as emulators that are trained based on the results obtained by numerical models or data models that are trained entirely on observed data [12, 13]. However, AI models are not unanimously accepted owing to their uncertainties and in-terpretation challenges. They are commonly referred to as black-box models, in which the result is satisfactory, but how the model processes the data to reach that result is un-known.

Explainable AI (XAI) methods aim to increase interpretability and understanding of AI models and can be applied to extract insights from these models. One of the key bene-fits of using these methods is that they allow users to understand and trust the decisions made by AI models, which is particularly important in high-stakes applications such as risk analysis and natural phenomena modelling. Predictive model outputs should be validated, justified, and speculated to enhance the interpretability of the results [14]. Vali-dation means that the evidence supporting the result should be verifiable, justification means that it should be possible to quantify the influence of each input variable in the prediction, and speculation means that the output sensitivity to each input variable should be quantifiable. However, most of the studies that have applied AI techniques to model coastal processes are limited to the performance metrics related with how well they predict the observed data.

XAI techniques, such as the SHapley Additive exPlanations (SHAP) method, have been applied to help interpret large-scale coastal subsidence models [15] and storm surge susceptibility prediction models [16], demonstrating the usefulness of this technique in coastal hydro-morphodynamic topics. This technique explains the individual predictions of a model [17]. However, other techniques, such as partial dependence (PD) plots, per-mutation importance (PIMP), and local interpretable model-agnostic explanation (LIME), are still rarely applied and can extract different insights from AI models. In particular, the PD plots and PIMP methods explain the global performance of a model, which is funda-mental for understanding its mean behaviour.

In this research, we aim to address the lack of examples regarding the application of different XAI techniques in coastal modelling. We implemented two random forest (RF) models to simulate the morphological evolution of coastal beach profiles using measured profile, wave action, and sediment diameter data of nine profiles on the northern Portugal coast. Subsequently, we applied three XAI

methods, PD plots, PIMP, and the LIME algo-rithm, to extract insights from both models, compare their performance, and assess how these RF work. Therefore, the primary objective of this study is to demonstrate how XAI methods can complement AI models, thereby enhancing the confidence in results ob-tained for complex problems such as coastal morphodynamics.

7.2. Materials and Methods

7.2.1. Study area

The northern coast of Portugal, stretching approximately 150 km between the Minho River estuary and Espinho beach, presents diverse coastal habitats, including sandy beaches, rocky shores, and estuaries, contributing to the region's ecological and economic value. The Douro River and Minho River estuaries are the largest sediment sources in the region [18–20], but the construction of dams has reduced the sediment transport capacity of these rivers, particularly the Minho River [21], increasing the erosion risk in the region [22, 23].

The meteo-oceanic characteristics of Northern Portugal are influenced by a variety of factors, including the cyclonic and frontal systems that cross the Atlantic Ocean. The re-gion experiences significant interannual variability in meteorological events, such as heavy precipitation leading to floods, particularly during the winter and autumn seasons [24], and energetic storms during winter [25].

The study area comprises five beaches in northern continental Portugal: Âncora, Amorosa, Neiva, Ofir, and Árvore (Figure 7-1).

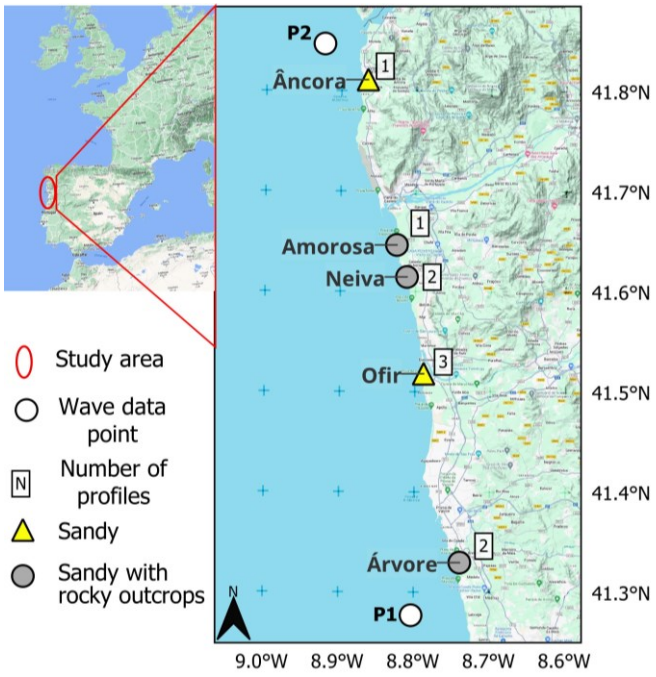


Figure 7-1. Study area, location of the selected beach profiles, wave data extraction poitnas and beach ty-pology. The white rectangles identify the number of profiles at each location in the map.

7.2.2. Random Forest

The RF algorithm is an ensemble learning method for classification and regression that operates by combining several decision trees at training time. The output of this method is the class that is the majority vote of the individual trees or the mean prediction of the trees for regression tasks [26, 27]. It is known for its high-performance classification capabilities, achieved through techniques such as bagging feature selection. At the same time it presents high-speed training and classification [28]. This algorithm is less prone to overfitting than individual decision trees because it combines the individual results of each tree by determining the importance score (weight) for each tree [29].

However, the algorithm's performance can be influenced by the individual trees' quality and the data's balance. For instance, an unbalanced data distribution in a high-dimensional space can lead to the inclusion of suboptimal trees, potentially affecting the outcome [29]. Moreover, the number of trees contributes to performance, with more trees generally leading to better results but at the cost of increased memory usage, which can be a limitation for implementation on small-scale hardware [28].

In summary, the RF algorithm is a robust and versatile machine learning (ML) method that combines the predictions of multiple decision trees to improve the accuracy and handle various data types. In particular, for coastal modelling, RF models have been applied to map flood hazards in urban areas and forecast shoreline changes [30, 31], thereby demonstrating the capacity of this method for coastal environments.

The input data were divided into three categories: profiles, wave action, and sedi-ment. Two models were implemented: M1, which used three data categories, and M2, which used only the profile and wave action data. The difference in the inputs was used to assess the contribution of the sediment data to the model performance. The models used data from the current profile condition and forecasted the next campaign profile, in which the time interval varied between two and four months. The Data Sources section provides more details on how the data were processed.

Both the models used the same hyperparameters. The number of estimators was 200, the depth was 80, the minimum sample split was five, and the minimum sample leaf was four. We used the Friedman MSE function [32] to measure the quality of the split nodes, and the square root criterion was used to limit the maximum number of features. The model was implemented using *Scikit-Learn* Library version 1.2.1 [32]. Figure 7-2 shows a sample decision tree for the M2 model.

The model performance was determined by computing the root mean squared error (RMSE) according to Equation (**7-1**).

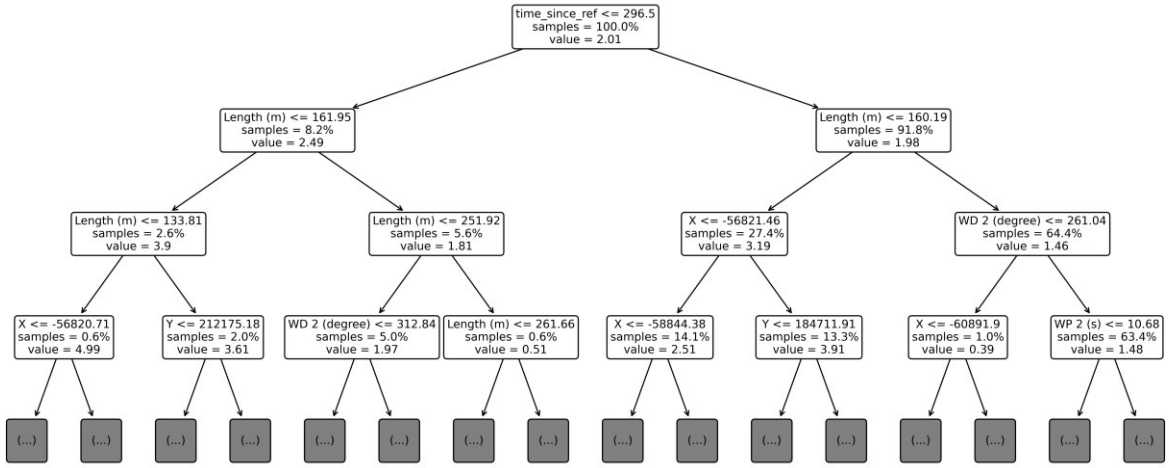


Figure 7-2. Example of a decision tree structure.

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7-1}$$

7.2.3. Data Sources

The COSMO program provided the beaches profiles data that monitors the coastal stretch of continental Portugal [33]. This program used GPS/RTK to survey the emerged profile section and a single-beam probe to survey the submerged section. Figure 7-3 depicts the elevation of the profiles used in this study. It was possible to identify more dynamic profiles, such as Âncora, Ofir, and Árvore, which presented considerable variations dur-ing the sampling period. The profiles of Amorosa and Neiva – P2 presented a more stable behaviour.

The model used four profile variables as inputs: x-coordinate (m), y-coordinate (m), profile length (m), and time (days). The number of days used as input resulted from the date difference between the current and next profiles. Both coordinates are in the ETRS89 reference system. The errors in the profile measurement ranged between 1.1 cm and 2.5 cm horizontally and between 2 cm and 5 cm vertically.

Additionally, a fictional starting point was computed to ensure that the first point was at a fixed location. Hence, the first profile point is 100 m before the first real profile point. A filter was applied to limit the data to the profile section that was more exposed to wave action to reduce the influence of aeolian sediment transport in the simulated beach profiles. The limit was the furthest point from the start of the profile, with an elevation of 6 m.

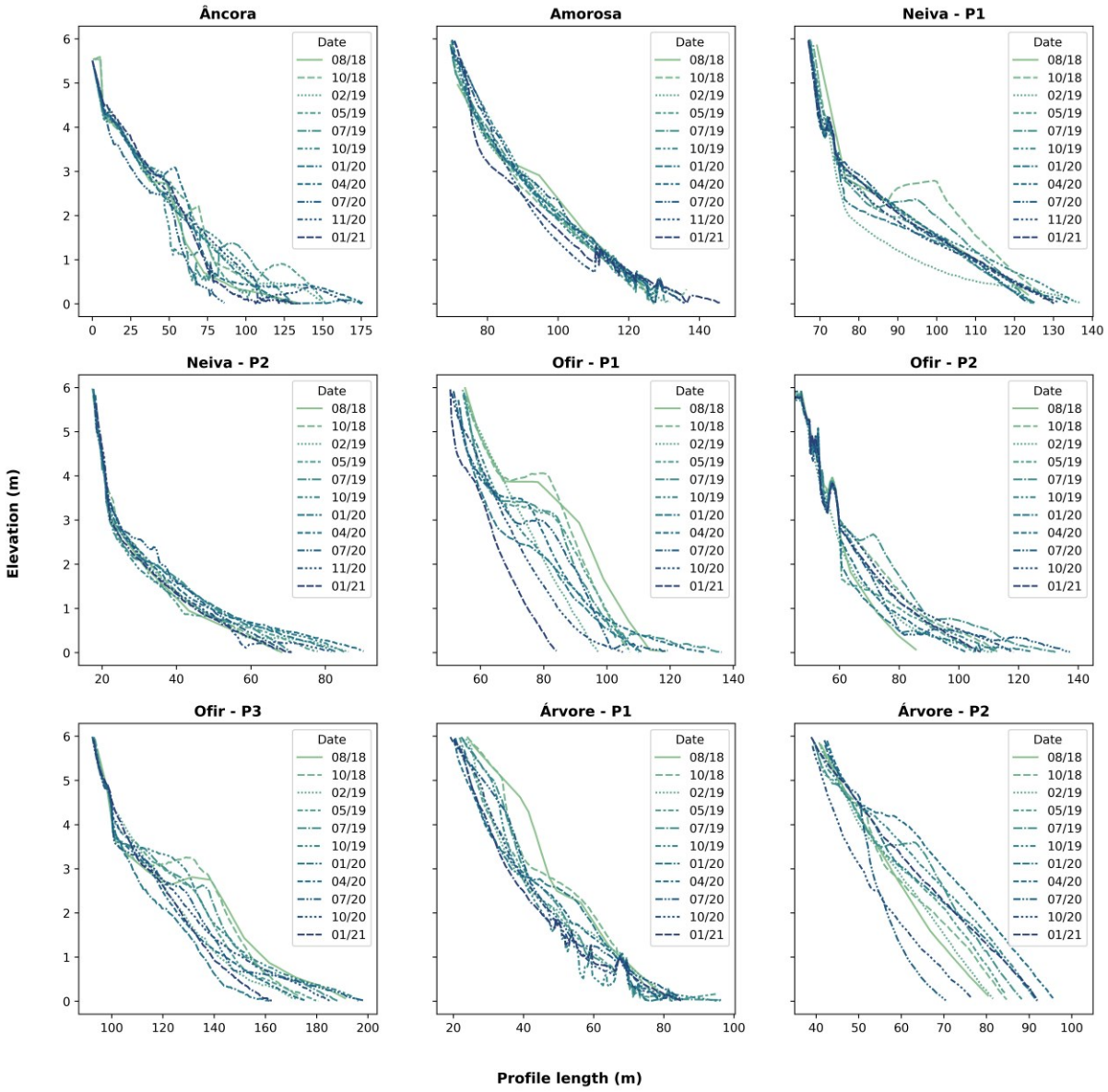


Figure 7-3. COSMO program beach profile data used to train the model.

Wave data were extracted from the Atlantic Iberian Biscay Irish (IBI) Ocean Wave Reanalysis product managed by the Copernicus Marine Service [34]. This product has a spatial resolution of 0.027° × 0.027° and an hourly temporal resolution between January 1993 and December 2021. The white circles in the Figure 1 indicate the locations of the wave action data extraction points (P1 and P2).

Figure 4. IBI wave reanalysis variables from Copernicus Marine Service.Figure 7-4 shows the three wave variables used in this study, namely significant wave hight, wave period, and mean wave direction. A clear predominance from the northwest waves can be observed, accounting for 50% of the observations, with more energetic conditions during the winter and milder conditions during the summer. The significant wave height peak was 8 m, a typical winter value [25, 35]. Approximately 30% of the NW waves had a sig-nificant wave height lower than 2 m, and approximately 5% had a wave height higher

than 3 m. The highest peak period was 24 s, coinciding with significant wave height peaks. 40% of the wave periods were lower than 13 s, and 2.5% overcame 16 s. Moreover, the wave conditions at both points are practically the same, indicating no significant dif-ferences between the selected wave data points.

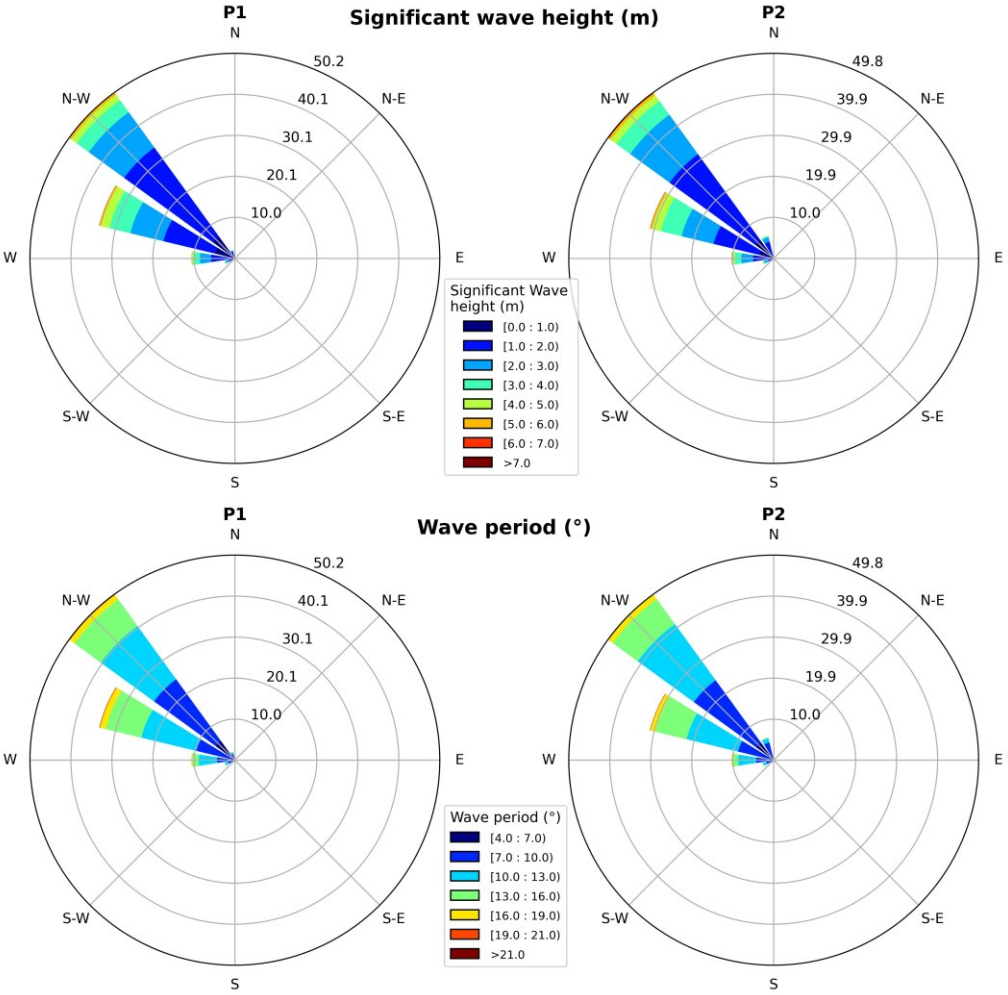


Figure 7-4. IBI wave reanalysis variables from Copernicus Marine Service.

Before feeding the model with the wave variables, the 90th percentile of the significant wave height, wave period, and mean wave direction between the current and next profiles were computed. Due to the fact that the input variables must have the same format, it is not possible to feed the model with hourly wave variable observations and monthly pro-files. Hence, this processing was necessary to adequately use wave information in the models. This limits the use of the model to predict short-term profile changes. For exam-ple, the occurrence of an extreme event greatly affects the profile elevations. However, the wave conditions at that moment would be smoothed by the values during the average conditions, masking their effects. This problem occurs mainly because of the interval be-tween the profile observations, which varies between two and four months.

The means and standard deviations of the wave variables are listed in Table 7-1. These metrics are useful for comparing variables at each wave data point. The mean significant wave height difference was 9 cm, whereas the means of the other variables were practi-cally the same. The standard deviation shows that the variability in the observations fol-lowed the same pattern, with minor differences between the data extraction points.

Table 7-1. Statistical description of wave action variables. Numbers 1 and 2 refer to the wave data points from which the data were extracted.

	HM0 1 (m)	HM0 2 (m)	WP 1 (s)	WP 2 (s)	WD 1 (°)	WD 2 (°)
Mean	1.99	2.08	10.82	10.68	300.82	302.21
Std	1.05	1.06	2.75	2.69	17.96	19.30

The sediment size data were surveyed during the consecution of the MarRisk project [35], which collected and characterised sand samples from several beaches in northern Portugal. This study considered beaches present in both the COSMO and MarRisk da-tasets. Sediment sizes were determined by seizing and digital images and were collected at three positions in the profiles: berm intermediate position (1), berm edge (2), and beach face (3), as summarised in Table 7-2. D50/D90 indicates that 50%/90% of the sand grains are smaller than the D50/D90 value. Given that the COSMO and MarRisk projects were inde-pendent, sand sampling was not necessarily conducted in exactly the same place in the profiles. Hence, it was assumed that all the profiles on the same beach had the same sed-iment characteristics.

Table 7-2. Characteristic sediment sizes (D50 and D90) at the studied beaches.

Beach	D50_1 (mm)	D50_2 (mm)	D50_3 (mm)	D90_1 (mm)	D90_2 (mm)	D90_3 (mm)	Coast type
Âncora	0.58	0.5	0.5	0.92	0.91	0.92	Sandy
Amorosa	0.37	1.48	0.42	0.61	4.67	2.58	Sandy with rocky outcrops
Neiva	0.36	0.29	0.22	0.76	0.49	0.44	Sandy with rocky outcrops
Ofir	0.35	0.45	0.27	0.49	1.38	0.75	Sandy
Árvore	0.64	0.72	0.34	0.98	1.92	0.58	Sandy with rocky outcrops

The sediment characteristics were considered to be constant over time. The idea be-hind using these data is to provide the M1 model with local information that may help to differentiate each beach.

Before training the model, the training and test datasets were determined. Consider-ing that there were few profile measurements, 92% of the data were used to train the mod-el and 8% for testing, resulting in 11 profiles for training and one for testing, which was the most recent profile. It is advisable to employ a training rate within the 65-80% range and a testing rate within the 5-20% [12, 36, 37]. However, the scarcity of data hinders the practical application of this suggestion.

7.2.4. XAI methods

AI models are popular because of their impressive capacity to associate inputs with outputs in multidimensional spaces, resulting in models with diverse uses, ranging from identifying patterns in images to flood prediction [38–40]. However, many of these models are considered black boxes and are not completely trusted by the research community, which is concerned with the models' limitations, interpretability, and uncertainties de-rived from the lack of knowledge of the internal processes that takes place when AI mod-els are applied.

Explanation methods should provide an understanding of the input variables and model response, which are essential parameters for the interpretability of the model [41]. In addition, these methods should be locally faithful, which refers to the faithfulness of an explanation that reflects the model's decision-making process for a specific input instance. The explanation captures why the model makes a particular prediction [42]. Moreover, an explainer should be model-agnostic, meaning it is reliable regardless of the model [41].

In this context, XAI tools can help overcome the interpretability barrier and increase acceptance of AI models. XAI methods promote the transparency of AI-based systems and consist of techniques, processes, and strategies that explain complex ML systems' output [43, 44]. These techniques can be divided into global and local explanations. Global ex-planations describe the model behaviour in all training/testing datasets, which helps un-derstand the average prediction performance of the model. Local explanations focus on a single prediction, which helps to understand the contribution of each input to the output [14].

We applied three XAI methods to extract insights from the data models applied in this study (RF). The PD profile method, also known as the PD plot method, shows the de-pendence of the model output on the input variables [14, 45], and is expressed through a partial dependence function (**7-2**). For example, if the resulting PD plot for a random varia-ble is a straight line, the variable has a linear influence on the output.

$$\hat{f}_{s(x_S)} = E_x\left[\hat{f}(x_S, x_C)\right] = \int \hat{f}(x_S, x_C) dP(x_C) \quad (7\text{-}2)$$

where x_s are the input features for the partial dependence function that will be defined, x_c are the other features of the ML model, $\hat{f}$ are the random variables, E_{xc} is the covariance matrix of the inputs, and $P(x_c)$ is the joint density of the inputs. The partial function $\hat{f}_s$ is estimated using the Monte Carlo method.

PIMP estimates the relevance of a model feature to the resulting output error after permuting the feature values [46, 47]. This process breaks the bond between the input and output data. This algorithm first computes the error of the original model. Second, it permutes the values of each feature and

estimates a new error in the model. Finally, it computes the feature importance based on the difference between the model errors before and after permuting.

The LIME algorithm approximates the black-box model to a more interpretable glass-box model. This method can explain the predictions of any model in an interpretable manner [41]. The method determines a local explanation considering fidelity and complexity terms (7-3)

$$s(x) = \underset{g \epsilon G}{\operatorname{argmin}} L(f, g, \pi_a) + \Omega(g) \tag{7-3}$$

where *s(x)* represents a local explanation for the prediction made by model f at a specific input x, and *argmin* denotes an argument that minimises a function. In this case, the minimized function is the sum of the two terms: $L(f, g, \pi_a)$ and $\Omega(g)$. $L(f, g, \pi_a)$ is a fidelity term that measures how well the local model *g* approximates the original model *f* around the input *x*, π_a represents a weighting function that can be used to focus the explanation on particular parts of the input, and $\Omega(g)$ is a complexity term that penalises the complexity of the local model g, helping to prevent the local model from overfitting to the data around *x*, and *G* is the set of all possible local models that can be used to explain the prediction of *f* at *x*.

7.3. Results

7.3.1. Model performance

The results of the model for the test dataset are shown in Figure 7-5. The maximum RMSE obtained was 52 cm at Ofir – P1. The findings revealed that M1 surpassed M2 in seven out of the nine profiles, suggesting that the inclusion of sediment data enhanced the model's performance. The M1 lowest error was achieved at Amorosa, Neiva – P1, and Neiva – P2, with an RMSE of 3 cm, whilst M2 lowest RMSE was 6 cm at Amorosa and Neiva – P1.

In addition, the beach did not affect the performance of the model. It is worth noting that the model performance on beaches with multiple profiles varied, and the model did not compute a general profile approximation that could present an average performance for both locations. For example, at Neiva - P1, the simulated profiles presented a harsh slope change at 175 m length with a flat section at 160 m, which both models identified. At Neiva – P2, the profile started on a cliff with only a few meters between 6 m and 2.5 m elevations, again identified by both models. This demonstrated that the models could ac-curately predict the profiles and correctly identify each profile at the same beach.

Comparing the obtained results with those of other studies, both models achieved satisfactory metrics. Birrien et al [48] implemented a coupled hydrodynamic equilibrium beach profile model that obtained a RMSE of 0.75 m, Want et al [49] used a process-based numerical model that solves equations for wave energy conservation, roller energy con-servation, wave setup calculation, flow velocity, sediment

transport, and bottom profile evolution using the Exner equation obtaining a RMSE of 0.11 m. Otto et al [50] forecasted the beach profile evolution using a spatiotemporal function, obtaining RMSE between 0.25 m and 0.9 m.

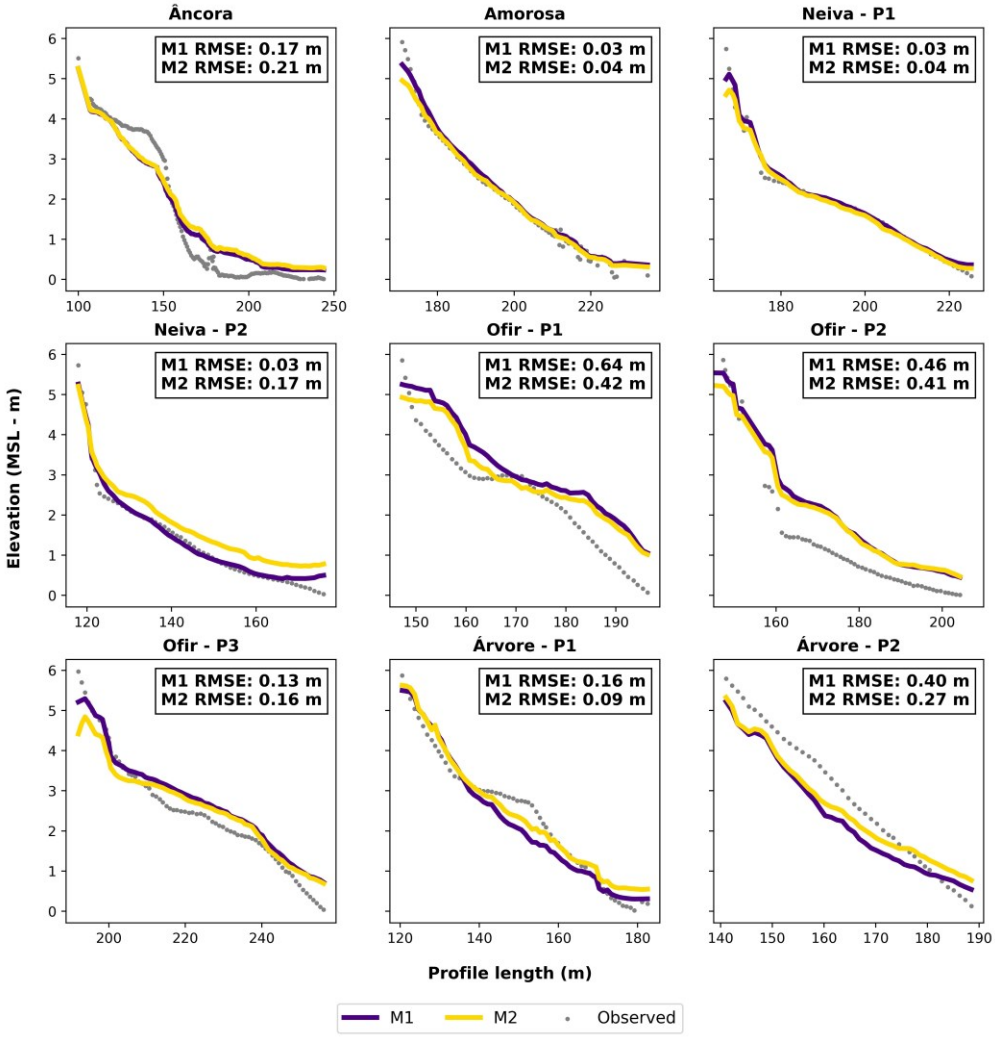


Figure 7-5. Local model performance for the test dataset.

In summary, both models demonstrated good performance, achieving error metrics similar to those obtained by process-based models. It should be highlighted that the lack of data did not significantly affect model performance, and both models adequately sim-ulated the profiles.

7.3.2. Partial dependence profiles

The PD profiles of the M1 and M2 models are depicted in Figure 7-6 and Figure 7-7, respectively. Both images demonstrate the usefulness of the method for interpreting the global behaviour of the models. For instance, the most important input variables were the x-y-coordinates and length, owing to the highest dependence of the model on these features. The x-y-coordinates presented flat areas and explicit breaks, which occurred due to the nonlinearity of these variables. The horizontal intervals represent areas with no data because the dataset used data from specific beaches. For example, the *M*

shape in the y-coordinate profile coincides with Ofir coordinates. However, it can be observed that the x-coordinates are more relevant than the y-coordinates. This could be physically associated with the proximity of the profile to the shoreline. Lower x values indicate points closer to the ocean, whereas higher values indicate points closer to the dune crest. Hence, the model was more sensitive to points farther from the ocean.

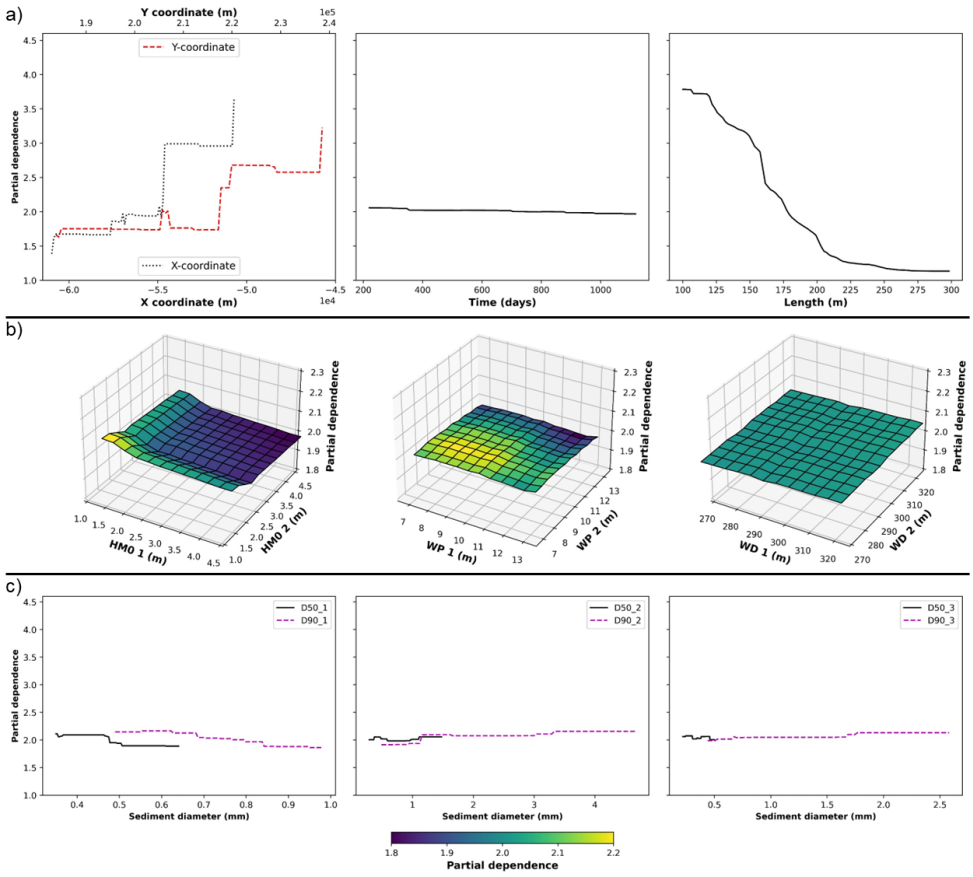


Figure 7-6. The PD profile results for M1. a) Effect of profile variables (coordinates, time, and length). b) Influence of wave action variables (HM0, wave period, and wave direction). c) Influence of sediment size variables at berm edge, intermediate position, and beach face.

The same pattern was observed when the profile length was assessed. The dependence of the models on this variable can be approximated by a second-degree function, indicating that the profile elevation is influenced more by actions that occur in the first 200 m of the profile. The influence of time on the output exhibited a linear tendency in both models.

Regarding the wave action variables, both wave data points similarly influenced the model. The 3D-plots crossed the wave variables at each data extraction point with their respective PD profiles. Significant wave heights between 1 m and 2 m influenced the profile elevation slightly more than those between 3.5 m to 4 m. This likely occurred because of the frequency of the observed wave heights and the absence of an extreme event during the period in which the data were surveyed. The wave period had

a linear influence on the output of both models, and its influence was represented by an almost horizontal plane, indicating a linear relationship between the wave period and the model result. The wave direction PD profile was a tilted plane with the highest vertex at 320°. This indicates that waves coming from the northwest are more relevant to the profile elevation than those from the west.

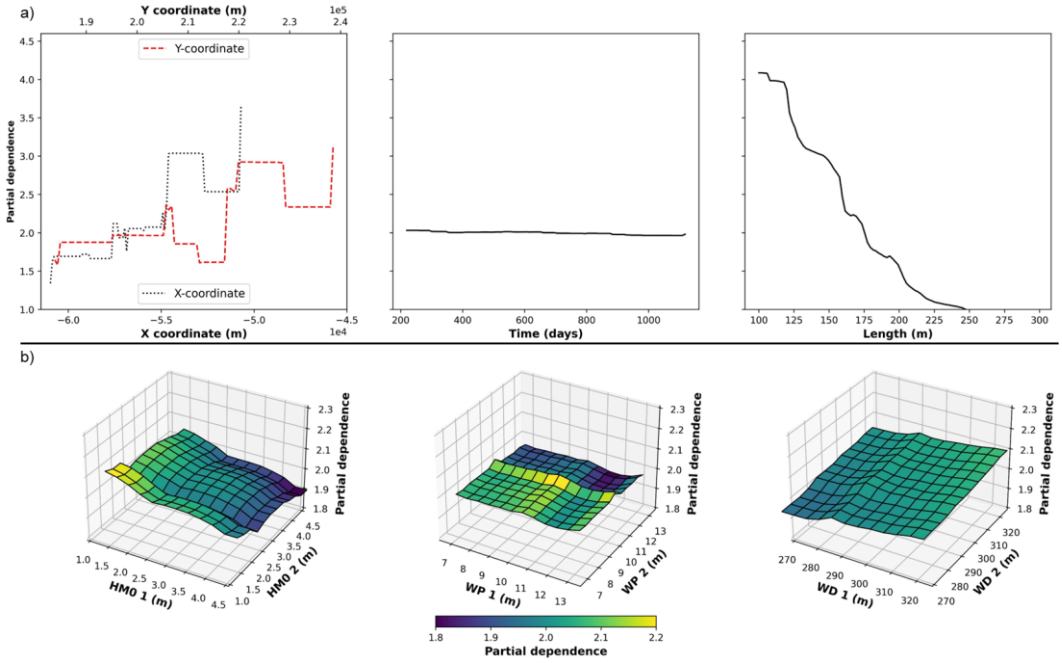


Figure 7-7. The partial dependence profile results for M2. a) Effect of profile variables (coordinates, time, and length). b) Influence of wave action variables (HM0, wave period, and wave direction).

In addition, the influence of the sediment size data on the M1 results indicated a linear dependence between sediment size and profile elevation. The horizontal line indicates that the relevance of this variable is constant, regardless of the sediment size and location. However, it must be stressed that the sediment size were considered constant over time, and the results of these PD profiles reflect the data availability. Were sediment samples collected over time as profile data, the sediment PD profiles would probably differ from those in Figure 7-6c.

7.3.3. Permutation importance

The first analysis of PIMP divided the inputs according to the variable types. Therefore, it was possible to assess the importance of wave action, profile, and sediment size variables separately (Figure 7-8). It can be observed that the profile data were the most important for both models. The baseline bar permutes all data simultaneously, whereas the full model bar is the result without permutation. For example, when permuting the profile data, the RMSE of the model increased by 2.1 m. Hence, the higher the loss, the higher the significance of that variable in the model.

A comparison between the models allowed us to conclude that the sediment variables were more important than the wave action variables regardless of adopting the sediment size constant over time. The sediment loss was 1.19 m, which is more than three times the wave loss. M2 compensated for the lack of sediment data by increasing the importance of the profile and wave variables, which was expected because the model had similar metrics.

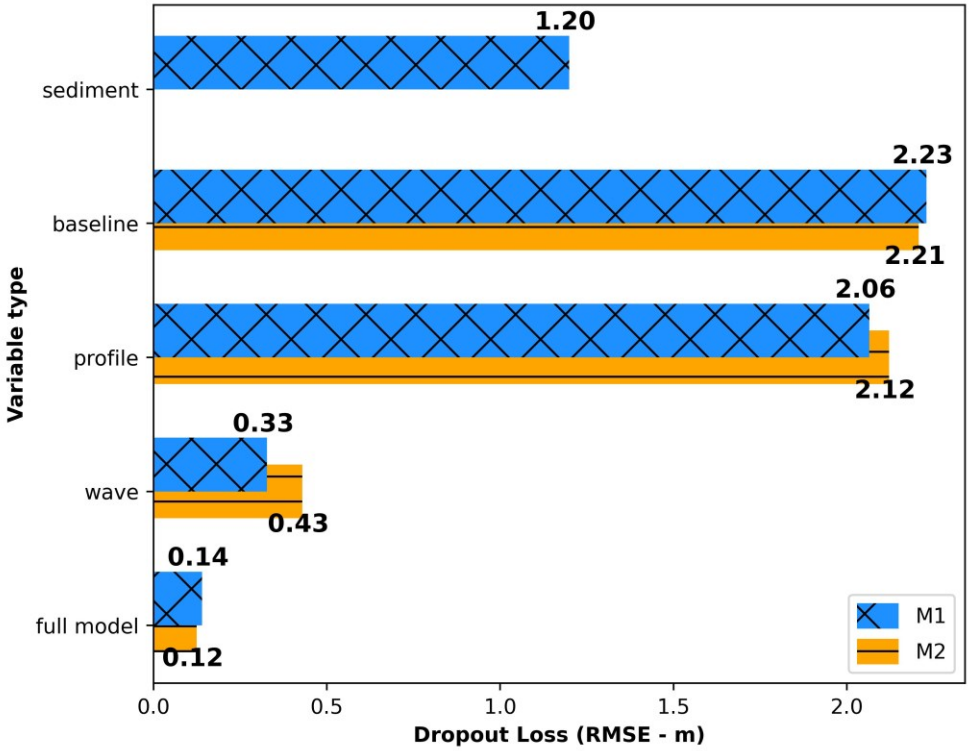


Figure 7-8. Permutation importance of variable groups.

Figure 7-9 shows the PIMP for each input variable. This figure helps determine the importance of each variable, helping the decision-making process regarding the model data. For instance, it is possible to eliminate irrelevant variables to simplify future versions of the models or even focus on efforts to gather data on the most important variables. This result followed the PD profile behaviour, with the profile length, x-coordinate, and y-coordinate variables being the most relevant to the output.

The comparison of the PIMP results reinforces the importance of the sediment data for M1 predictions. The loss associated with the sediment diameter varied between 0.24 m and 0.40 m, while the wave action variables loss varied between 0.13 m and 0.24 m. The sediments sampled at an intermediate position were more important than those at the other two positions. In addition, the sum of the D90 sediments scored 1.03 against 0.85 of the D50, indicating that D90 was more relevant to the output. Moreover, an assessment of the wave variables showed that the wave data points did not influence the model's results. Both points had equivalent results for all variables.

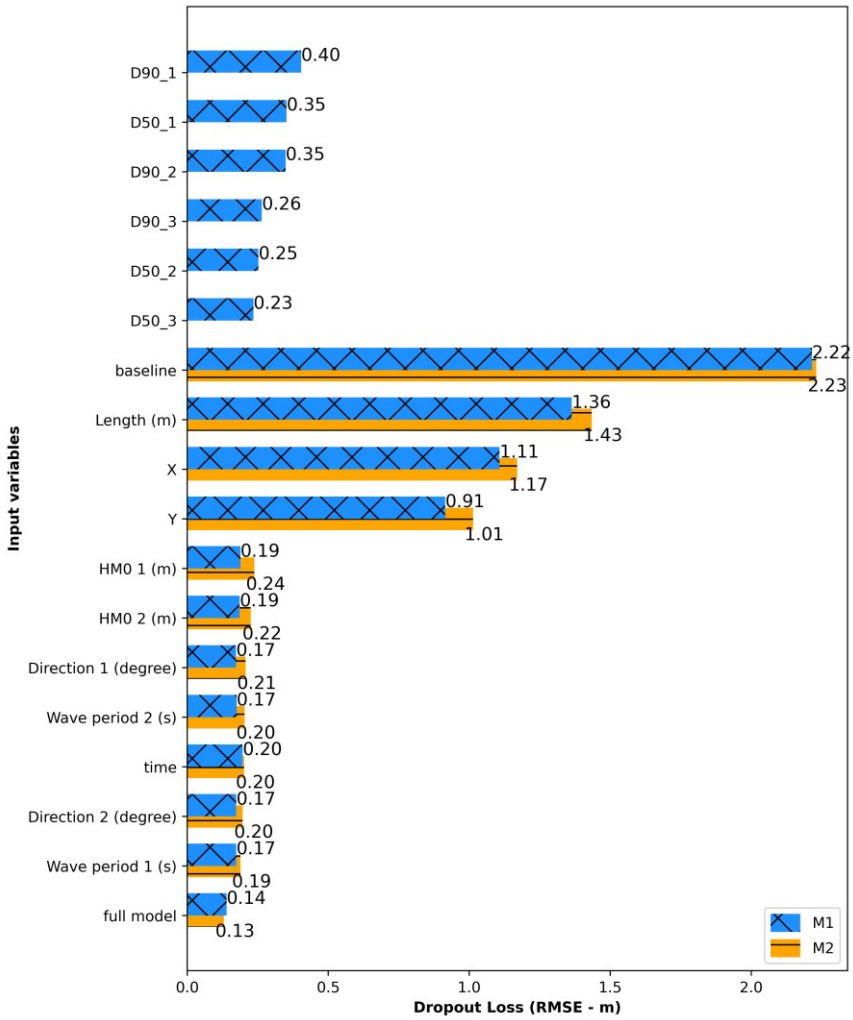


Figure 7-9. Permutation importance of input variables.

7.3.4. LIME

The LIME method was applied at two points on each profile: one placed at the shoreline (Figure 7-10) and the other at the berm edge (Figure 7-11**Erro! Fonte de referência não encontrada.**). In addition, the method was applied considering wave action and sediment size variables because of their physical relevance to the morphological profile evolution. The higher the weight of the variable, the higher the importance of that variable to the output.

The LIME results at both points showed that sediment size variables were more important than wave action variables. The weight of the wave action variables varied between -0.1 and 0.1 for both models, whilst the sediment variables ranged between -0.25 and 0.19. Both models considered the wave action similarly, as demonstrated in the results of PIMP and LIME methods, suggesting that the data that really improved the model profile elevation predictions is the sediment size.

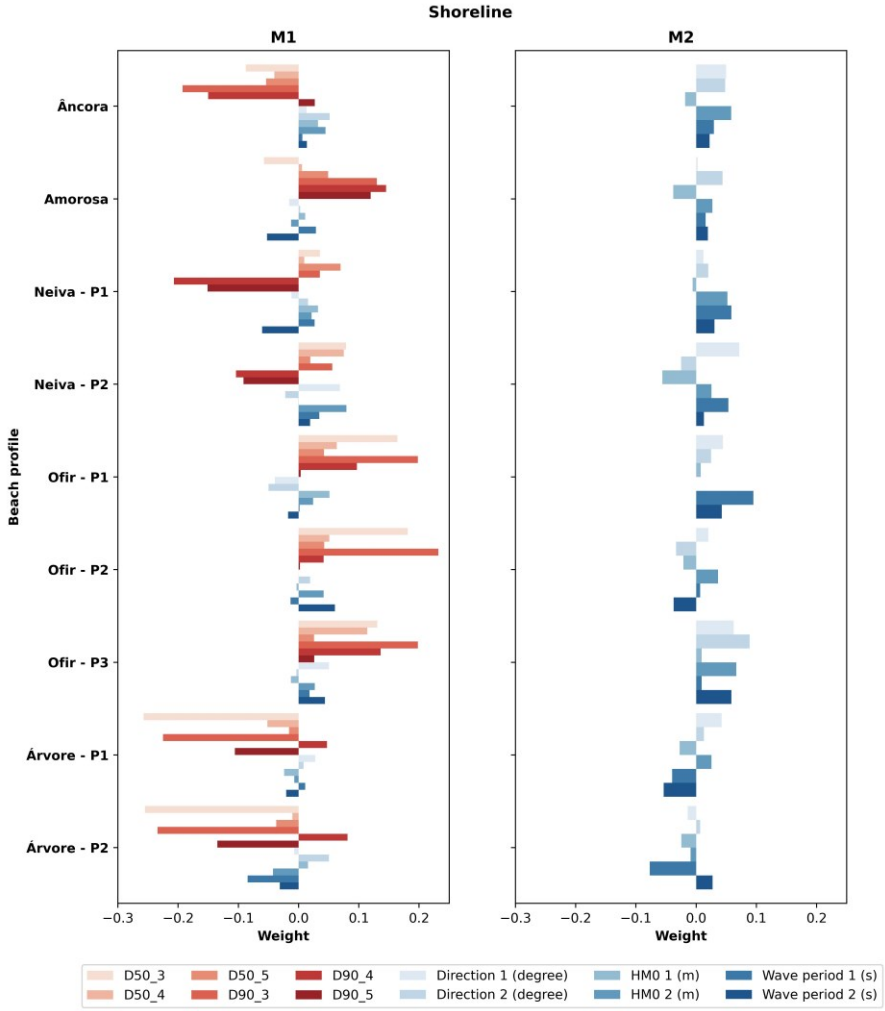


Figure 7-10. LIME method results for shoreline points.

Additionally, it was possible to identify the most relevant variables at each point. For example, at Âncora, the wave direction had higher significance than the wave period and significant wave height. In contrast, the direction and significant wave height had similar relevance at the berm edge. This result also indicates that it is possible to simplify the model by removing one of the wave data points because there is no clear pattern between the wave action variables and their respective weights. Hence, reducing one of these points would help to better associate each input's relevance with the obtained output.

The M1 results show how sediment size was considered at each beach. It can be observed that profiles at the same beach considered the sediment size equally, whereas the wave variable weights did not present any pattern. Physically, it can be interpreted that the sediment component relevance is constant, varying only according to the beach, whereas the influence of the wave variables may differ according to the profile location and orientation with the wave incidence and local bathymetry.

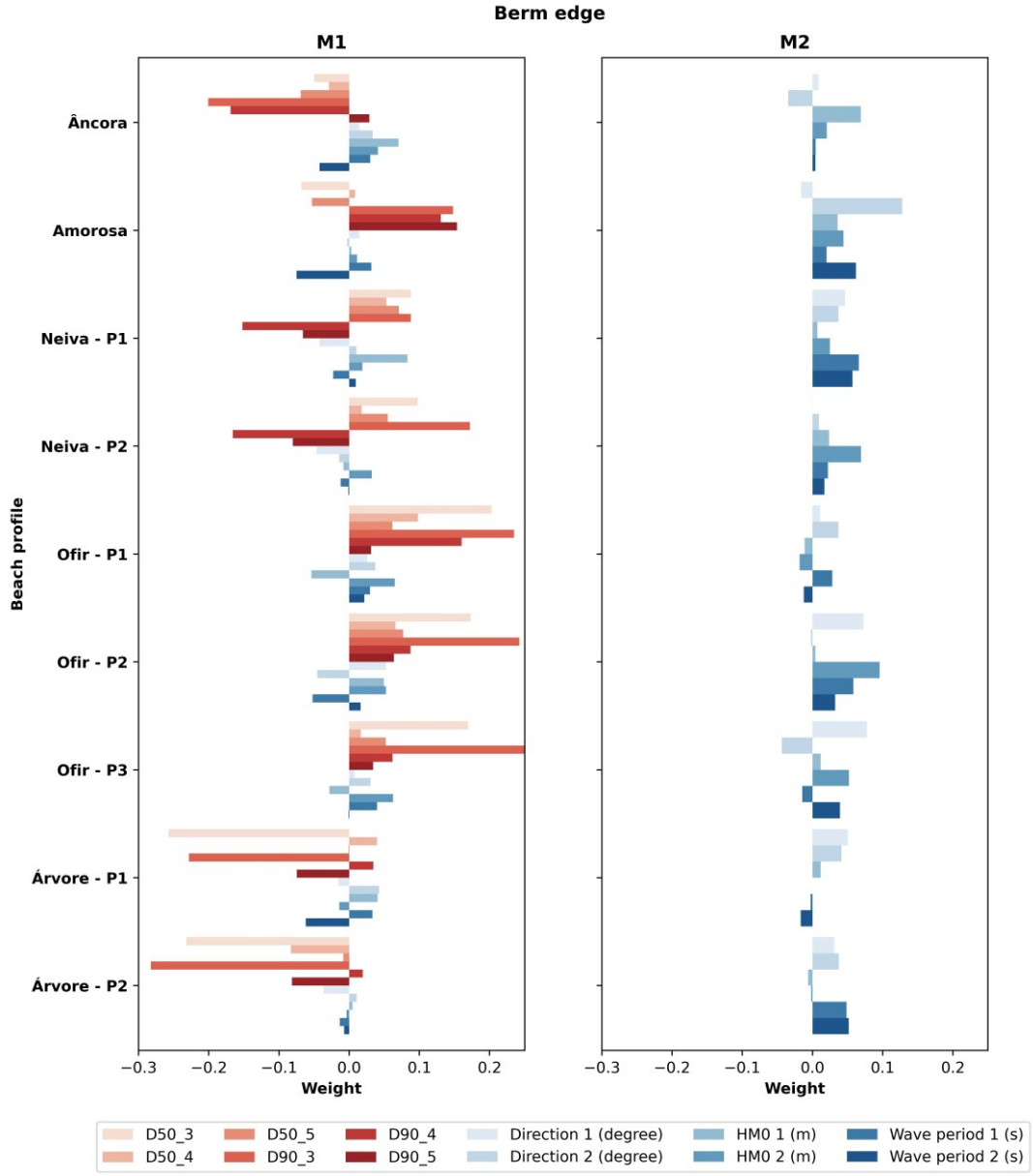


Figure 7-11. LIME method results for berm edge points.

Moreover, the influence of the wave variables on the edge of the berm was reduced compared to their effect on the shoreline for both models. For instance, the sum of the significant wave height weights at both wave data points was reduced for all profiles except for Âncora and Árvore, whose influences remained approximately the same at both profile positions. This behaviour is expected, considering that the influence of wave height is reduced at positions farther from the ocean.

It should be noted that this analysis focused on wave action and sediment size variables. The profile variables had a greater influence on the output, as demonstrated by the PIMP results. However, this analysis aims to demonstrate how the LIME method can extract knowledge from an RF model's punctual results. Therefore, we restricted our analysis to variables with an already established physical relationship to the profile evolution. However, this method can also be used to discover the inputs most

relevant to the output and then compare them with physical models, which can provide insights into how to consider each variable and their relevance to the model outcome.

Finally, considering both points, the mean, maximum, and minimum weights of each variable were determined (Table 7-3). The wave action variables at the shoreline had higher maxima and lower minima than those at the berm edge for M1, which was the expected physical behaviour. For M2, the wave variable weights presented a similar pattern between the points at the shoreline and the berm edge. Moreover, the sediment variables did not exhibit any patterns. The mean weights must be carefully considered because they include positive and negative values in their calculus. Hence, a mean near 0 indicates a symmetric weight distribution and not that the variable is irrelevant.

Table 7-3. LIME method mean, maximum and minimum weights for each input variable.

Model	Metric	HM0 1	HM0 2	WP 1	WP 2	WD 1	WD 2	D50 1	D50 2	D50 3	D90 1	D90 2	D90 3
M1 shoreline	mean	0.011	0.020	0.003	-0.005	0.010	0.008	-0.008	0.024	0.016	0.022	0.010	-0.034
	min	-0.024	-0.042	-0.085	-0.061	-0.039	-0.050	-0.257	-0.052	-0.054	-0.234	-0.207	-0.151
	max	0.051	0.079	0.034	0.060	0.068	0.052	0.181	0.114	0.070	0.232	0.145	0.120
M1 berm edge	mean	0.018	0.029	0.007	-0.016	-0.003	0.011	0.014	0.021	0.021	0.047	0.001	0.001
	min	-0.054	-0.014	-0.052	-0.075	-0.047	-0.045	-0.257	-0.083	-0.069	-0.283	-0.169	-0.081
	max	0.083	0.065	0.040	0.022	0.052	0.042	0.203	0.098	0.077	0.250	0.160	0.154
M2 shoreline	mean	-0.019	0.031	0.017	0.014	0.032	0.021						
	min	-0.056	-0.009	-0.077	-0.054	-0.014	-0.033						
	max	0.009	0.067	0.095	0.059	0.072	0.089						
M2 berm edge	mean	0.016	0.032	0.026	0.026	0.031	0.023						
	min	-0.011	-0.018	-0.015	-0.017	-0.016	-0.044						
	max	0.069	0.096	0.066	0.062	0.078	0.128						

7.4. Discussion and conclusions

This study demonstrated the use of XAI methods to interpret and select ML models. Two RF models were implemented: M1 used profile, wave action, and sediment variables as inputs, and M2 used the same variables, except for the sediment variables. Both models achieved metrics similar to those obtained by numerical models, indicating that the accuracy of AI-based models can equalise or even overcome process-based models. However, one of the main barriers to achieving a more generalised use of AI models is the distrust on these methodologies, being called black-box models. Despite the methods' good results and computational performance [10, 12], the difficulty in interpreting and understanding the model is one of the main reasons for their avoidance.

In this study, we applied XAI methods to extract knowledge from two ML models, demonstrating, at the same time, that they can be as accurate as process-based models. The PD profiles show how the

model inputs influence the output. This method can individually show linear and nonlinear relationships between variables, which is fundamental for better understanding the model results and comparing them with other types of models. Additionally, the PD profile method generates valuable insights that can guide the use of process-based models. For instance, the relationship between variables that do not have a mathematical definition can be conceptualised after understanding how an ML model relates to these variables. This allows for an increase in the complexity of the models according to the data availability.

The PIMP method shows the behaviour of the models' error after permuting the input variables. This method is fundamental for optimising AI models because less important variables can be identified and removed from an overly-complex model. Reducing the number of inputs is one of the easiest ways to enhance interpretability. This method can also be used to determine the key features in a specific situation. For example, assuming that there were sufficient data relating the profile morphological evolution to the wind and wave variables. The PIMP method can show which phenomena are more relevant to the output, and it is possible to know how and where to act in hazardous situations. Considering a profile with a severe erosive trend, it would be possible to identify which variables are leading the erosive process and act according to their importance.

The LIME method can generate insights similar to the PIMP method; however, the first is a local method, whereas the second is global. This means that the LIME method is useful for better understanding a specific model prediction, whereas the PIMP method helps interpret the model itself. This study used the LIME method to understand how each variable influences the points at the shoreline and beach berm. Therefore, it was possible to identify the most relevant variables at each location for the same model. Considering the analogy used to explain the importance of the PIMP method, the LIME method would be useful for determining the positions in the profile in which each variable dominates the erosive process. Hence, it would be possible to act locally instead of assuming that all points have the same trend.

In conclusion, this study demonstrated that AI-based models can be as accurate as numerical models and confirmed that the interpretation barrier of AI models can be overcome using adequate methods. Both models, numerical and AI-based, have positive and negative features, but the major advantage of AI models is their capacity to solve complex phenomena as long as they were feed with sufficient and good quality data. This novel techniques completely changed the modelling paradigm, in which explanations (equations) were used to generate the output data. With AI models, the output is already known; hence, the challenge is determining the equations that lead from the inputs to the output and ensuring that the model is capable of generalisation. This is the inverse of the classical approach, in

which equations are determined based on empirical methods and observations. AI models are already guaranteed to be capable of reproducing data, but the real challenge is how to interpret this model to increase its reliability.

Therefore, the objective of showing how to explore and extract knowledge from black-box models was successfully achieved. Using XAI methods is fundamental for increasing our understanding of how the data models work and expanding their applications. It was demonstrated that it is possible to enhance their applicability with adequate methods without restricting the model performance to computing only the error metrics.

7.4.1. Methodology limitations

The performance of AI-based methodologies depends on the quality and quantity of the data. Hence, the model is reliable based on the characteristics of the data used to train the parameters. These models are expected to exhibit a decrease in performance when using data or forecast events outside the range of the training data. For example, in this study, the wave action data did not present extreme events during the model training. Hence, the model is unreliable for assessing the profile evolution after a storm.

Similarly, wave data preprocessing smoothed the effects of less common events when the 90th percentiles of significant wave height, wave period, and the mean wave direction were computed. This means that extreme events would be masked during the data preprocessing phase. To overcome this limitation, the frequency of profile samples must be increased. Therefore, when increasing the number of field surveys is not viable, using numerical models to simulate the profile evolution between field campaigns can solve this limitation. In this case, the input/output data would combine the observed data and numerical model results, thereby increasing the dataset size.

7.5. References

1. Pörtner H-O, Roberts DC, Tignor M, et al (2022) IPCC, 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovern-mental Panel on Climate Change. Cambridge, Uk and New York, NY, USA
2. Oppenheimer M, Glavovic B (2019) Sea Level Rise and Implications for Low-Lying Islands, Coasts and Communities. In: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate
3. Bai J, Li X (2023) The evolution of China's policies on marine and coastal ecosystems in climate change adaptation. Front Mar Sci 10:. https://doi.org/10.3389/fmars.2023.1190132
4. IPCC (2019) The Ocean and Cryosphere in a Changing Climate. A Special Report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change 1–765
5. Biolchi LG, Unguendoli S, Bressan L, et al (2022) Ensemble technique application to an XBeach-based coastal Early Warning System for the Northwest Adriatic Sea (Emilia-Romagna region, Italy). Coastal Engineering 173:. https://doi.org/10.1016/j.coastaleng.2022.104081
6. Rutten J, Torres-Freyermuth A, Puleo JA (2021) Uncertainty in runup predictions on natural beaches using XBeach nonhydrostatic. Coastal Engineering 166:103869. https://doi.org/10.1016/j.coastaleng.2021.103869
7. Matheen N, Harley MD, Turner IL, et al (2021) Bathymetric data requirements for operational coastal erosion forecasting using xbeach. J Mar Sci Eng 9:. https://doi.org/10.3390/jmse9101053

8. Roelvink D, Reniers A, van Dongeren A, et al (2009) Modelling storm impacts on beaches, dunes and barrier islands. Coastal Engineering 56:1133–1152. https://doi.org/10.1016/j.coastaleng.2009.08.006
9. Gharagozlou A, Anderson DL, Gorski JF, Dietrich JC (2022) Emulator For Eroded Beach And Dune Profiles Due To Storms. J Geophys Res Earth Surf 127:. https://doi.org/10.1029/2022JF006620
10. Weber de Melo W, Pinho J, Iglesias I (2022) Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model. Journal of Hydroinformatics. https://doi.org/10.2166/hydro.2022.068
11. Abouhalima M, das Neves L, Taveira-Pinto F, Rosa-Santos P (2024) Machine Learning in Coastal Engineering: Applications, Challenges, and Perspectives. J Mar Sci Eng 12:638. https://doi.org/10.3390/jmse12040638
12. Weber de Melo W, Pinho J, Iglesias I (2023) Coastal morphodynamic emulator for early warning short-term forecasts. Environmental Modelling & Software 165:105729. https://doi.org/10.1016/j.envsoft.2023.105729
13. Goldstein EB, Coco G (2015) Machine learning components in deterministic models: hybrid synergy in the age of data. Front Environ Sci 3:1–4. https://doi.org/10.3389/fenvs.2015.00033
14. Biecek P, Burzykowski T (2021) Explanatory Model Analysis. Explore, Explain, and Examine Predictive Models. With examples in R and Python. Chapman and Hall/ CRC, New York
15. Qiao X, Chu T, Krell E, et al (2024) Interpretation and Attribution of Coastal Land Subsidence: An InSAR and Machine Learning Perspective. IEEE J Sel Top Appl Earth Obs Remote Sens 17:4768–4783. https://doi.org/10.1109/JSTARS.2024.3361391
16. Alshayeb MJ, Hang HT, Shohan AAA, Bindajam AA (2024) Novel optimized deep learning algorithms and explainable artificial intelligence for storm surge susceptibility modeling and management in a flood-prone island. Natural Hazards 120:5099–5128. https://doi.org/10.1007/s11069-024-06414-6
17. Molnar C (2024) Interpretable Machine Learning A Guide for Making Black Box Models Explainable, 2nd ed
18. Mendes R, Sousa MC, DeCastro M, et al (2016) New insights into the Western Iberian Buoyant Plume: Interaction between the Douro and Minho River plumes under winter conditions. Prog Oceanogr 141:30–43. https://doi.org/10.1016/j.pocean.2015.11.006
19. Santos AI, Oliveira A, Pinto JP, Freitas MC (2021) Hydrodynamic and Sediment Transport Patterns in the Minho and Douro Estuaries (NW Portugal) Based on ADCP Monitoring Data: Part 1-Tidal Sediment Exchanges. Coasts 1:31–55. https://doi.org/10.3390/coasts1010003
20. Balsinha MJ, Santos AI, Oliveira ATC, Alves AMC (2009) Textural composition of sediments from Minho and Douro estuaries (Portugal) and its relation with hydrodynamics. J Coast Res 1330–1334
21. Melo W, Pinho J, Iglesias I, et al (2020) Hydro- and Morphodynamic Impacts of Sea Level Rise: The Minho Estuary Case Study. J Mar Sci Eng 8:441. https://doi.org/10.3390/jmse8060441
22. Bio A, Gonçalves JA, Iglesias I, et al (2022) Coastal erosion: a northern-portuguese case study. In: Martí Ezpeleta A, Lorenzo González N, Royé D, Díaz Poso A (eds). Asociación Española de Climatología, Madrid
23. Vousdoukas MI, Ranasinghe R, Mentaschi L, et al (2020) Sandy coastlines under threat of erosion. Nat Clim Chang in press: https://doi.org/10.1038/s41558-020-0697-0
24. Alcoforado MJ, Silva LP, Amorim I, et al (2021) Historical floods of the Douro River in Porto, Portugal (1727–1799). Clim Change 165:17. https://doi.org/10.1007/s10584-021-03039-7
25. Oliveira T, Cagnin E, A. Silva P (2020) Wind-waves in the coast of mainland Portugal induced by post-tropical storms. Ocean Engineering 217:108020. https://doi.org/10.1016/j.oceaneng.2020.108020
26. Dai B, Chen R-C, Zhu S-Z, Zhang W-W (2018) Using Random Forest Algorithm for Breast Cancer Diag-nosis. In: 2018 International Symposium on Computer, Consumer and Control (IS3C). IEEE, pp 449–452
27. Breiman L (2001) Random Forests
28. Mishina Y, Murata R, Yamauchi Y, et al (2015) Boosted Random Forest. IEICE Trans Inf Syst E98.D:1630–1636. https://doi.org/10.1587/transinf.2014OPP0004
29. Cutler A, Cutler DR, Stevens JR (2012) Random Forests. In: Zhang C, Ma Y (eds) Ensemble Machine Learning. Springer New York, New York, NY, pp 157–175
30. Kim H Il, Kim BH (2020) Flood Hazard Rating Prediction for Urban Areas Using Random Forest and LSTM. KSCE Journal of Civil Engineering 24:3884–3896. https://doi.org/10.1007/s12205-020-0951-z
31. Gunawan PH, Iryanto, Ghozali AL, et al (2023) Data-driven Shoreline Change Forecasting on Eretan Beach Using Random Forest. Institute of Electrical and Electronics Engineers (IEEE), pp 1–4

32. Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12:2825–2830
33. APA (2020) Programa de Monitorização da Faixa Costeira de Portugal Continental (COSMO). https://cosmo.apambiente.pt. Accessed 15 Apr 2020
34. Copernicus Marine Service (2022) Atlantic -Iberian Biscay Irish- Ocean Wave Reanalysis. https://data.marine.copernicus.eu/product/IBI_MULTIYEAR_WAV_005_006/description. Accessed 4 Feb 2023
35. Bio A, Gonçalves JA, Pinho J, et al (2020) Indicadores de vulnerabilidade de erosão costeira: Um estudo de caso no Norte de Portugal. Revista de Gestão Costeira Integrada 20:197–209. https://doi.org/10.5894/rgci-n337
36. Aggarwal CC (2018) Neural Networks and Deep Learning. Springer International Publishing, Cham
37. Motamedi M, Sakharnykh N, Kaldewey T (2021) A Data-Centric Approach for Training Deep Neural Networks with Less Data. In: 35th Conference on Neural Information Processing Systems
38. Nearing G, Cohen D, Dube V, et al (2023) AI Increases Global Access to Reliable Flood Forecasts
39. Santhanam V, Morariu VI, Davis LS (2016) Generalized Deep Image to Image Regression. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2017-Janua:5395–5405. https://doi.org/10.1109/CVPR.2017.573
40. Nearing G, Klotz D, Frame JM, et al (2022) Technical note: Data assimilation and autoregression for using near-real-time streamflow observations in long short-term memory networks. Hydrol Earth Syst Sci 26:5493–5513. https://doi.org/10.5194/hess-26-5493-2022
41. Ribeiro MT, Singh S, Guestrin C (2016) "Why Should I Trust You?": Explaining the Predictions of Any Classifier
42. Alvarez-Melis D, Jaakkola TS (2018) On the Robustness of Interpretability Methods
43. Ridley M (2022) Explainable Artificial Intelligence (XAI). Information Technology and Libraries 41:. https://doi.org/10.6017/ITAL.V41I2.14683
44. Adadi A, Berrada M (2018) Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). IEEE Access 6:52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052
45. Friedman JH (2001) Greedy function approximation: A gradient boosting machine. The Annals of Statistics 29:1189–1232. https://doi.org/10.1214/aos/1013203451
46. Altmann A, Toloşi L, Sander O, Lengauer T (2010) Permutation importance: A corrected feature im-portance measure. Bioinformatics 26:1340–1347. https://doi.org/10.1093/bioinformatics/btq134
47. Fisher A, Rudin C, Dominici F (2019) All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously
48. Birrien F, Baldock T (2021) A coupled hydrodynamic-equilibrium type beach profile evolution model. J Mar Sci Eng 9:. https://doi.org/10.3390/jmse9040353
49. Wang G, Li Y, Zhang C, et al (2023) Effects of Wave Height, Period and Sea Level on Barred Beach Profile Evolution: Revisiting the Roller Slope in a Beach Morphodynamic Model. Water (Switzerland) 15:. https://doi.org/10.3390/w15050923
50. Otto P, Piter A, Gijsman R (2021) Statistical analysis of beach profiles – A spatiotemporal functional ap-proach. Coastal Engineering 170:. https://doi.org/10.1016/j.coastaleng.2021.103999

CHAPTER 8.

Conclusions and final remarks.

"The good thing about science is that it's true whether or not you believe in it."

Neil deGrasse Tyson (1958 -) – North American astrophysicist and scientific communicator.

Chapter 8. Conclusions and final remarks.

During the consecution of this PhD thesis, several AI methods were implemented to model HMD processes using hydrodynamic variables to simulate the morphological evolution of estuarine and coastal areas. Two distinct approaches were applied, each one of them yielding promising outcomes. The first approach involved combining diverse data sources with training a data model. Using only hydrodynamic variables and geographical information as inputs, it was possible to simulate the evolution of several beach profiles. This comprehensive approach demonstrates the potential of AI for a full-data solution. The second approach, known as emulation, employed artificial data generated using morpho-hydrodynamic numerical models. This strategy efficiently reduces the computational resources required for the simulations. Both approaches have demonstrated exceptional efficacy, underscoring the potential of AI as a pivotal tool for advancing HMD research.

Applying the RF algorithm demonstrated that simpler models, such as ML models, can yield reliable results if the available data has sufficient quality and quantity. Chapter 3 underscored the significance of reliable data sources and elucidated the advantages that open data sources could offer in accelerating the advancement of scientific comprehension in the field of HMD.

While the RF algorithm performed well, the major drawback of this method is the difficulty in interpreting its results. The model does not consider each input explicitly, and it is necessary to adopt explainable AI methods to overcome this limitation. Despite that, the model achieved error metrics in the same order as other methods. Moreover, it was possible to extract beach features from the simulated profiles, namely the beach slope and advance/retreat in specific elevations, demonstrating the usefulness of the method.

Chapter 4 and Chapter 5 applied a similar DL architecture to emulate the morphodynamic processes within estuarine and coastal areas. These models effectively replicated the results of numerical simulations and demonstrated success across diverse temporal scopes, physical phenomena, and computational resource reduction requirements. Chapter 4 describes how the DL model reproduced the long-term morphological evolution of the intertidal zone of an estuary forced by tidal elevations and fluvial discharges. In Chapter 5, the model correctly replicated the effects of a storm in a coastal area forced by tides, currents, and waves. Regardless of the higher complexity of the second case, the model also achieved satisfactory results, demonstrating the inherent adaptability of the DL models.

Furthermore, the DL model emulated the morphological evolution in 2D, whilst the RF model was 1D. It once more demonstrated the capacity of AI algorithms to compute different variables in different domains. Even increasing the number of dimensions and processes, a satisfactory performance was

achieved by both methods. Hence, it is concluded that AI models are suitable tools for modelling HMD processes, regardless of the domain and complexity involved.

Chapter 6 presented a framework to combine AI and process-based models to implement an FEWS. The chapter demonstrated the capacity of the AI models to forecast fluvial discharges considering time horizons of up to 36 hours. It was shown that attention layers optimized the fluvial discharge forecasting capacity and increased the accuracy of predictors considering a time window of 36 hours.

Finally, Chapter 7 explored possible interpretations of a ML model using XAI concepts. It was demonstrated that the interpretability barrier of AI models can be overcome by adopting appropriate methods that demonstrate how they relate inputs to outputs, thereby enhancing the reliability of this type of solution.

In summary, this thesis demonstrates the proficiency of AI models, their flexibility with available data, and the improvements they can deliver to HMD modelling. Data-driven models can encompass all variables influencing the output, irrespective of our knowledge of each variable. In theory, the impact of every variable is embedded in the output, even if complete comprehension has not yet been achieved. In contrast, numerical models are based on the assumption that the applied differential equations comprehend all phenomena that influence the physical process. Although numerical models can achieve exceptional outcomes, the associated uncertainties in their results must be considered cautiously, demanding, in some cases, the implementation of supplementary techniques, such as ensembles, to mitigate them.

For instance, we can consider an estuary in which all the necessary data to implement numerical and data models are available. To perform a forecast with a numerical model, it will be necessary to include all the details of the bathymetry in the grid, the changes in the sediment characteristics in each section, the depth of the sediment layers, and the existence of plants at the bottom to determine the bottom friction coefficient adequately. For example, areas with coral reefs have completely different hydrodynamic patterns than the ones observed in areas with mangroves. The assumption of uniformity in the bottom friction coefficient across the entire area augments model uncertainty. Conversely, the comprehensive integration of these diverse factors escalates the efforts to implement the model.

In the case of implementing a data model in the same area, all the particularities of the domain are reflected in the output variable. Consequently, the AI model can learn how to consider each detail in the study area by adjusting its parameters. Achieving highly detailed models is relatively easier in AI models than in numerical models. In a hypothetical scenario where the output variable is the water level, the data model is apt to decipher the distinct behavioural nuances of this variable across various sections

of the domain. Importantly, this understanding evolves without requiring manual demarcation of individual regions as it learns how to consider each region during training.

Additionally, the bed shear stress coefficient varies according to the water level. For instance, different values of this parameter can lead to different outcomes during flood events. Therefore, it would be necessary to have a numerical model configuration for the average conditions and another configuration for the extreme conditions. In data modelling techniques, if there are data about both average and extreme events, the model can learn that, during an extreme event, the relationship between the inputs and outputs changes.

Another advantage of AI models is their computational performance. They compute the results faster than numerical models, which is extremely important when the processing time is a limiting factor. For example, the numerical simulation component is critical during the forecasting phase in early warning systems. The higher the computational time the model requires, the later the forecast is generated, which could negatively affect the adoption of early actions. Thus, optimising the modelling component can enhance the application of real-time EWS by reducing the computation time.

However, one of the limitations of AI models is the necessity of high-quality data. AI solutions are as good as the data provided to the model, and deeply understanding the data is fundamental to recognize the constraints of the model. For instance, an AI model trained on average conditions cannot be expected to predict extreme events accurately. The model would still compute a result, but its accuracy would decrease. Therefore, analysing the data is crucial to understand the scope of application of AI models.

Among the similarities between the applied AI methods, they require a certain degree of data pre-processing. At this stage, graphics and descriptive statistics are essential for extracting insights from the data and drawing a picture of what the model learns. Extreme value analysis, data distribution, and variable correlation help the modeller to simplify the model and understand when it performs well, the uncertainties related to the inputs and outputs, and to optimise the model. In addition, identifying data outliers is paramount for ensuring the data quality used to train the model. At this stage, it is possible to identify which variables are most relevant to our problem and exclude those with low importance. The simplification of data models is fundamental for understanding how the model works.

In the instance of the RF method, this technique necessitates minimal pre-processing effort, but data analysis remains a crucial component. Normalization is indispensable to attain satisfactory outcomes with NNs (LSTM and CNN). In the NNs implemented in this thesis, normalisation layers were selected to allow the use of raw data as inputs. However, it is also possible to normalise the data before setting the inputs.

Regarding data science and AI, this thesis demonstrates the importance of quality over quantity when implementing data models. The estuarine emulator used 365 images per variable, whereas the coastal emulator used 109 images. This represents a 70% reduction in the number of images, and the accuracy of the model was not significantly affected. Hence, the hypothesis that AI algorithms demand large data was rejected. This conclusion agrees with the data-centric approaches for AI implementation that have become popular in the last few years. The performance of the model will be higher when trained with fewer good-quality data than when trained with many poor-quality data.

Hence, it is concluded that AI models are as reliable as numerical models, with particular advantages and disadvantages as any other modelling methodology. Both model types are a simplification of reality, and understanding the limitations of each methodology is fundamental for potentializing the use of each model. This was explicitly demonstrated in **Erro! Fonte de referência não encontrada.**, i n which AI and numerical methods were combined to implement the modelling component of a FEWS. The AI component was fundamental for increasing the forecasting time horizon of the model, whereas the numerical component was essential for estimating the water levels in the study area. This thesis also demonstrates that different AI algorithms can be successfully applied to model HMD processes, from forecasting fluvial discharges and water levels to beach profiles and long- and short-term morphological changes.

In addition, this thesis accomplished the main objective of the research project, which was to implement AI methods for operational forecasting modelling. It was demonstrated that AI methods can be applied directly to field observations and emulate numerical models, reaching errors similar to those of the numerical models and reducing the simulation computational time.

8.1. Future research

Future research on HMD modelling can apply AI to generate digital twins to create a virtual representation of real systems of the world. AI is used as the engine of digital twins owing to its capacity to manage big data efficiently. Digital twins are widely used in industry to control and manage different processes and operations. One example is the Digital Twin Ocean produced by the Mercator Ocean, which promises to combine real-time data and predictive models to generate valuable insights to support the decision-making process of human interventions in the ocean. This would allow us to quantify the human impact on the ocean and adapt our interventions by contributing to a more sustainable use of the ocean's natural resources.

The development of the Douro Estuary FEWS can potentially be further enhanced. The focus of this study was limited to the modelling component of an EWS, which could be improved by identifying

additional monitoring variables, particularly for areas with limited data availability. Additionally, a comprehensive assessment of the risks in the area, implementation of a dissemination and warning system, and evaluation of response capabilities are necessary to ensure the effective operation of a fully functioning EWS.