BERGISCHE
UNIVERSITÄT
WUPPERTAL

Dissertation

mit dem Titel

Neural networks for spatial and temporal ocean wave height prediction considering coastal morphodynamics in the East Frisian North Sea

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

im Fach Geographie
durch die Fakultät für Human- und Sozialwissenschaften
der Bergischen Universität Wuppertal

vorgelegt von
Christoph Jörges

Wuppertal, Dezember 2022

Sometimes, you just have to go with the waves.

Acknowledgements

My heartfelt thanks goes to all companions who have supported me during the time of this dissertation.
First of all, I would like to thank Britta Stumpe for the always good and productive collaboration and supervision of this work. As a member of your working group, I was always able to develop personally and professionally. I also want to thank Andreas Keil for the pleasant atmosphere at the institute, and the other members of the review board, Miriam Kuckuck and Bernd Marschner, for their easy cooperation in this work.
Hanno Gottschalk, I would like to thank you for your always good support and supervision not only during the joint paper. It was a pleasure to work with you.
I would like to thank all colleagues of the NLWKN for the friendly hospitality during the research stays at Norderney and the very good and fun cooperation. Especially, I would like to thank Cordula Berkenbrink for her very good advice and expertise in coastal engineering!
I am also very thankful for all of my colleagues at the Institute. In particular, I would like to thank Kathi for the good time we spent together in the office and on conferences, and our professional discussions. Jannis, thank you especially for your always open ear. I have learned a lot from you and enjoyed the time we spend together on unforgettable excursions. Jonas, you have become a good friend and an important support in the institute during my time at BUW, for which I would like to thank you.
Further, I would like to thank my friends and family. You have always been by my side and supported me in all situations. Thanks to my parents and my sister for the continuous mental help, without which this work would not have been possible.
Lastly, I would like to thank Steffi for her outstanding support and understanding in all situations!

Christoph Jörges
December 2022

Contents

List of Figures

List of Tables

Nomenclature

List of Abbreviations

AI	Artificial Intelligence
ANFIS	Adaptive Network-based Fuzzy Inference System
ANN	Artificial Neural Network
BSS	Between Sum of Squares
cGAN	Conditional Generative Adversarial Network
CNN	Convolutional Neural Network
DEM	Digital Elevation Model
DFFNN	Deep Feed-Forward Neural Network
DIA	Discrete Interaction Approximation
DL	Deep Learning
DTM	Digital Terrain Model
DWD	German Meteorological Service (Deutscher Wetterdienst)
ED-CNN	Encoder-Decoder Convolutional Neural Network
ELM	Extreme Learning Machine
EO	Earth Observation
EPUNN	Evolutionary Product Unit Neural Network
ETD	Ebb-Tidal Delta
FFNN	Feed-Forward Neural Network
FFT	Fast Fourier Transformation
GIS	Geographical Information System
GMSL	Global Mean Sea Level
GRF	Gaussian Random Field

IPCC	Intergovernmental Panel on Climate Change
LiDAR	Light Detection And Ranging
LLW	Lowest Low Water
LSTM	Long Short-Term Memory
MAAPE	Mean Arctangent Absolute Percentage Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
metocean	meteorology and oceanography
MHW	Mean High Water
ML	Machine Learning
MLR	Multiple Linear Regression
MLW	Mean Low Water
MRAN	Minimal Resource Allocation Network
MSE	Mean Squared Error
MW	Mean Water
NG	Nordergründe
NHN	German Mean Sea Level 'Normalhöhennull'
NLWKN	Lower Saxony Water Management, Coastal Defence, and Nature Conservation Agency
NOAA	National Oceanic and Atmospheric Administration
NWG	Nordwestgründe
PINN	Physics-Informed Neural Network
P-LSTM	Parallel Long Short-Term Memory
PM	Pierson-Moskowitz
QGIS	Quantum GIS

RBF	Radial Basis Function
ReLU	Rectified Linear Unit
RF	Random Forest
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
RQ	Research Question
SAR	Synthetic Aperture Radar
SDB	Standard Deviation of Bias
SGD	Stochastic Gradient Descent
SL-FFNN	Single-Layered Fully Connected FFNN
SLR	Sea Level Rise
SMS	Surface-Water Modeling System
SSE	Sum of Squared Errors
SVM	Support Vector Machine
SVR	Support Vector Regression
SWAN	Simulating WAves Nearshore
SWH	Significant Wave Height
THW	Tidal High Water
TIN	Triangulated Irregular Network
TLW	Tidal Low Water
UAV	Unmanned Aerial Vehicle
WAFO	Wave Analysis for Fatigue and Oceanography
WL	Water Level
WSA	Waterways and Shipping Administration
WSS	Within Sum of Squares

Notation and symbols

α	Learning Rate
α_{EB}	Eldeberky Proportionality Coefficient
b	Mini Batch Size
β	β-coefficients of Multiple Linear Regression
c_0	Nugget Effect
c_1	Partial Sill Variance
C_t	Memory Cell State
Dir_P	Peak Direction
f	Frequency, Activation Function
F	Matérn Isotropic Covariance
$f_{max_{EB}}$	Eldeberky Maximum Frequency
f_S	Separation Frequency
f_t	Forget Gate
g	Gravity Constant = 9.81 ms^{-2}
γ	Semivariance
Γ	Gamma Function
$H_{m_0},\ H_S$	Significant Wave Height
$H^I_{m_0},\ H^{II}_{m_0}$	Significant Wave Height at Coast I, Coast II
h_t	Hidden State
i_t	Input Gate
k	Lookback of LSTM Input, Feature Map
K_ν	Modified Bessel function of the second kind of order ν
l	Prediction Time Step of LSTM, CNN Layer

L	Loss Function
λ	Regularization Parameter
m_n	n-th Moment of Variance Density Spectrum
n	Total Number of Data
N	Action Density Spectrum
NaN	Not a Number
o_t	Output Gate
r	Pearson Correlation Coefficient
R^2	Coefficient of Determination
S	Variance Density Spectrum
t	Time
$tanh$	Hyperbolic Tangent
Θ	Wave Direction
$T_{m_{0,1}},\ T_m$	Mean Wave Period
$T_{m_{-1,0}},\ T_e$	Energy Period
T_P	Peak Period
U_θ	Wind Direction
$U,\ U_S$	Wind Speed, Wind Speed at Sea
W	Weight Matrice
X	Input Value
X_N	Normalized Input Value
Y	Output Value

Abstract

Climate change impacts such as sea level rise and change in storms particularly threaten coastal regions. Therefore, proper predictions of significant wave height (SWH) become crucial for coastal protection and marine disaster prevention. Especially for the coastal areas of the East Frisian North Sea, the complex morphodynamic variability of ebb-tidal delta (ETD) sandbanks determines the local wave climate along with non-linear wave propagation and tidal currents.
The major aim of this dissertation was to enhance the understanding of coastal morphodynamic impacts on the nearshore sea state by analysis of the ETD damping effect as well as utilizing this knowledge to improve spatial and temporal machine learning predictions of SWH in the research area of Norderney, Germany.
Quantification of the sea state damping revealed a crucial variability of up to $14 \pm 2\%$ due to variations of the coastal morphodynamics, while an average damping effect of $41 \pm 11\%$ was found for the ETD sandbanks. Since the consideration of coastal morphodynamics is hence essential for accurate prediction of SWH, an LSTM neural network using bathymetric data for buoy-based short- and long-term SWH predictions was developed for the first time. With an RMSE of 0.069 m, a parallel LSTM structure was able to improve the SWH reconstruction using bathymetric data by 16.7%. As the accuracy was improved for the single-point predictions including coastal morphodynamics, a further two-dimensional CNN wave height forecasting model was developed for spatial predictions considering the continuously morphodynamically changing ETD sandbanks. Limited by the few costly bathymetric data available, it was required to develop a method to generate sufficient training data. For the first time, geostatistical variogram analysis and random fields were applied for ETD sandbank simulations. Trained with SWAN-modeled and metocean ground truth, the mixed-data CNN surrogate model achieved good accuracy with an spatial averaged RMSE of 0.097 m. Further validation on buoy measurements revealed similar performance of the CNN (RMSE = 0.23 m) and SWAN (RMSE = 0.218 m) models. In addition, the efficient machine learning approach offers the possibility of real-time forecasting and exploration of various morphodynamic and metocean scenarios for coastal protection, as the computation time was reduced by a factor $>300\,000$ compared to the numerical SWAN model.
These results make a decisive contribution to the research of sea state predictions in coastal morphodynamically determined areas. New machine learning and geostatistical methods for future studies are hereby provided.

Keywords: spatio-temporal forecasting, significant wave height prediction, LSTM, CNN, machine learning, coastal morphodynamics, ebb-tidal deltas, SWAN surrogate

Kurzfassung

Auswirkungen des Klimawandels, wie der Anstieg des Meeresspiegels und sich verändernde Stürme, bedrohen insbesondere die Küstenregionen. Korrekte Vorhersagen der signifikanten Wellenhöhe (SWH) sind daher für den Küsten- und Katastrophenschutz von entscheidender Bedeutung. In den Küstengebieten der ostfriesischen Nordsee bestimmen die morphodynamische Variabilität der Sandbänke des Gezeitendeltas (ETD) und eine komplexe Wellenausbreitung das lokale Seegangsklima.
Ziel dieser Arbeit ist es, die Auswirkungen der Küstenmorphodynamik auf den Seegang durch die Analyse des ETD-Dämpfungseffektes zu verstehen und dieses Wissen zur Verbesserung von räumlichen und zeitlichen Vorhersagen der SWH mit maschinellen Lernverfahren im Forschungsgebiet Norderney zu nutzen.
Aufgrund von Variationen der Küstenmorphodynamik zeigte die Quantifizierung der Seegangsdämpfung eine Variabilität von 14 ± 2%, während die ETD-Sandbänke einen Dämpfungseffekt von durchschnittlich 41 ± 11% verursachten. Eine Berücksichtigung der Küstenmorphodynamik ist daher für genaue Vorhersagen der SWH unerlässlich, weshalb erstmals ein neuronales LSTM-Netzwerk mit bathymetrischen Inputdaten für bojenbasierte kurz- und langfristige Vorhersagen entwickelt wurde. Mit einem RMSE von 0,069 m konnte eine parallele LSTM-Struktur die Rekonstruktion der SWH um 16,7% verbessern. Da die Einbeziehung der Küstenmorphodynamik die Genauigkeit punktueller Vorhersagen verbesserte, wurde ein zweidimensionales CNN zur räumlichen Vorhersage der Wellenhöhen, unter Berücksichtigung der sich kontinuierlich morphodynamisch verändernden ETD-Sandbänke, entwickelt. Aufgrund der teuren, begrenzt verfügbaren bathymetrischen Daten wurden erstmalig geostatistische Variogrammanalysen und Zufallsfelder für die Simulation synthetischer ETD-Sandbänke eingesetzt, um Trainingsdaten zu generieren. Trainiert auf gemischten Daten aus numerischen SWAN-Modellen sowie metozeanischen Messungen, erreichte das CNN einen räumlich-gemittelten RMSE von 0,097 m. Eine weitere Validierung mit Bojenmessungen ergab eine vergleichbare Leistung von CNN (RMSE = 0,23 m) und SWAN (RMSE = 0,218 m). Da der Ansatz des maschinellen Lernens die Berechnungszeit im Vergleich zum numerischen Modell um einen Faktor >300 000 reduzieren konnte, bietet er die Möglichkeit der Echtzeitvorhersage sowie der Untersuchung morphodynamischer und metozeanischer Szenarien.
Diese Ergebnisse leisten einen entscheidenden Beitrag zur Erforschung von Seegangsvorhersagen in küstennahen und morphodynamisch bestimmten Gebieten. Sie liefern neue Methoden des maschinellen Lernens und der Geostatistik für künftige Studien.

Schlagworte: räumlich-zeitliche Vorhersage, signifikante Wellenhöhe, LSTM, CNN, maschinelles Lernen, Küstenmorphodynamik, Gezeitendelta, SWAN Substitutionsmodell

1 Introduction

As a foundation of this dissertation, the following chapter introduces the research topic and motivates the objectives. An overview of climate-related hazards to the East Frisian North Sea as the study area of this work is provided, followed by a review of basic related and state of the art research literature. Finally, the research questions and objectives are derived from the identified research gaps.

1.1 Introduction and motivation

More than 10% of today's world population lives in coastal regions that are less than 10 m above sea level. In the future, this number will increase even further (IPCC, 2019). These coastal regions are particularly affected by the present and future impacts of global climate change (Toimil et al., 2020). Due to increasing rates of ice loss from the ice sheets and glaciers, and the unabated thermal expansion of the oceans driven by anthropogenic forced global warming, the global mean sea level (GMSL) is rising at an accelerating rate in recent decades (Hay et al., 2015; IPCC, 2019). In addition, a potential increase in frequency and intensity of storms and a possible trend of increasing ocean wave heights, which contribute to coastal erosion and flooding, threaten the people living in those regions (Grabemann and Weisse, 2008; IPCC, 2019; Mori et al., 2013; Vanem and Walker, 2013; Weisse et al., 2012). Especially for the coastal regions of the tide-dominated North Sea, storm surges pose a perpetual hazard (Bitner-Gregersen et al., 2018; Weisse et al., 2012). About 14% of the land area of Lower Saxony, Germany is deemed to be vulnerable to storm surges (Berkenbrink and Wurpts, 2019). These findings raise special issues for coastal protection on the North Sea coast in light of climate-related hazards (IPCC, 2019; Niemeyer et al., 2018). For proper protection of the coastal regions, their populations, ecosystems, landscapes, and economies, common hard and soft coastal protection measures such as dikes, seawalls, surge barriers, and beach nourishments are often used (Berkenbrink and Wurpts, 2019; IPCC, 2019; Niemeyer et al., 2018). Precise knowledge of the regional wave climate and accurate predictions of related parameters are an essential requirement for building these coastal protection and offshore structures, as well as for planing marine operations, disaster prevention, and feeding climate and wave models (Deo and Naidu, 1999; Ti et al., 2018). In this context, the wave height parameter is one of the most important wave characteristics (besides the wave period), mostly represented by the significant wave height (SWH), which is given by the mean wave height of the highest third of all waves in a defined time interval (Teich et al., 2018).

In order to monitor ocean wave heights, networks of ocean wave buoys can be used (Berkenbrink et al., 2019). Multiple factors may cause gaps in the in-situ measured time series. These include erroneous measurements by buoy floating during most interesting extreme events, technical issues raised by the marine environment (mechanical impacts, marine fouling, and corrosion), maintenance operations, collisions with ships, and vandalism (Cavaleri, 2009; Londhe, 2008). Therefore, sophisticated methods for accurate reconstruction and prediction of ocean wave heights are required for seamless time series.

The buoy-based in-situ measurements are also essential for the setup and validation of spatial wave models. Examples for common state of the art numerical wave models based on the spectral action balance equation are the third-generation wave model Simulating WAves Nearshore (SWAN) (Booij et al., 1999), as well as WAM (The Wamdi Group, 1988), and WAVEWATCH III (Tolman et al., 2002). SWAN was especially designed for coastal and nearshore areas to model two-dimensional irregular and random directional wave fields based on hydrodynamic boundary conditions and an underlying bathymetry (Booij et al., 1999; Hoque et al., 2019; Sebastian et al., 2014; Vieira et al., 2020).
Although SWAN generates precise wave models due to its physical equations of wave propagation and interaction, it is a time- and cost-intensive approach and requires high computational power. Thus, SWAN is not suitable for real-time predictions of storm surge impacts (Hu et al., 2021).

As an alternative to numerical modeling, various studies and methods using a statistical machine learning (ML) approach exist for the reconstruction and prediction of single-point and two-dimensional sea state data (Abed-Elmdoust and Kerachian, 2012; Fernández et al., 2015; Mahjoobi et al., 2008; Malekmohamadi et al., 2011; Mandal and Prabaharan, 2010; Nitsure et al., 2012). Artificial neural networks (ANN) constitute a major group of the soft-computing ML methods for the related tasks (Ali and Prasad, 2019; Deo and Naidu, 1999; Deo et al., 2001; Dixit and Londhe, 2016; Durán-Rosal et al., 2016; Kumar et al., 2017; Law et al., 2020; Londhe, 2008; Londhe and Panchang, 2007; Peres et al., 2015; Zamani et al., 2008). While numerous different ANN architectures can be found (Goodfellow et al., 2016), two of them are of particular focus in the remainder of this thesis:
On the one hand, recurrent neural networks (RNN) such as the long short-term memory (LSTM) neural network are designed to process sequential data or time series (Hochreiter and Schmidhuber, 1997; LeCun et al., 2015). This class of network

is capable of forgetting certain information and overcoming the vanishing gradient problem (Gers et al., 2000). In addition to its typical application of speech and handwriting recognition (Graves and Jaitly, 2014; Graves et al., 2009), machine translation (Wu et al., 2016), image captioning (Aggarwal, 2018), and stock price prediction (Kim and Won, 2018), the LSTM has already been applied to predict single-point SWH with various lead times (Fan et al., 2020; Kagemoto, 2020; Ni and Ma, 2020; Pirhooshyaran et al., 2020).
On the other hand, convolutional neural networks (CNN) are designed to process spatial data like images by learning local patterns (LeCun et al., 1989). In addition to its typical applications of image classification (Sultana et al., 2018; Thai et al., 2012), face and speech recognition (Redmon et al., 2016), and semantic segmentation (Yuan et al., 2020), the CNN has already been applied to predict two-dimensional SWH (Bai et al., 2022; Wei and Davison, 2022). Furthermore, approaches combining LSTM and CNN exist (Zhou et al., 2021).

However, none of these approaches have yet considered coastal morphodynamics for sea state prediction. The ocean builds a complex system at the intersection of hydrosphere, lithosphere, and atmosphere, including several non-linear processes under the influence of the different geospheres. Especially nearshore areas are subject to complex coastal morphodynamics, which should be considered for SWH prediction on a local scale for a particular study area (Kaiser and Niemeyer, 2007; Wang et al., 2012). This dissertation focuses on the study area of the nearshore area of the East Frisian Island of Norderney, Germany (c.f. Section 1.3.3).
As typical for many coastal areas, an ebb-tidal delta (ETD) has developed as a morphological feature in the study area of Norderney (Anthony, 2013; Bergillos et al., 2016; Hansen et al., 2013). ETDs are generated by sediments from the ebb-tidal currents that accumulate seaward of a tidal inlet and form sandbanks (Dallas and Barnard, 2011; Hayes, 1980). The ETD sandbanks off the coast of Norderney exert a great impact on the local wave climate and provide a natural coastal protection against shoreward waves. The resulting damping effect of the ETD sandbanks depends on the incoming SWH and the coastal morphodynamics and morphology of the sandbanks (Niemeyer, 1979; Niemeyer and Kaiser, 1999).
Due to the ebb-tidal currents and even single severe storm surges, the ETD sandbanks are subject to constant and rapid morphological changes by migration, erosion, and accumulation (Castelle et al., 2007; Dallas and Barnard, 2011; Wang et al., 2012). Therefore, the ETD sandbank morphology might be crucial for modeling, reconstructing, and predicting single-point and spatial wave heights (Cooper and

Navas, 2004; Dallas and Barnard, 2011; Eshleman et al., 2007). However, high temporal and spatial resolution data of ETD sandbanks – mostly given by bathymetries – are very limited and cost-intensive due to data acquisition by hydrographic surveys (Spicer et al., 2019).

Based on the above mentioned scientific knowledge and ongoing challenges, this work is concerned with the deeper understanding of the influence of the ETD sandbanks in the nearshore area of Norderney on single-point and two-dimensional predictions of SWH using ML methods, as well as the quantification of the sea state damping effect. The application of different neural network architectures and the synthetic generation of ETD sandbanks will also be analyzed in more detail as a part of this work.

1.2 Emerging climate-related hazards to the North Sea coast

The following sections describe general climate-related hazards to the North Sea coast and highlight the topic of sea level rise. The present measures and development of coastal protection and management are finally presented.

1.2.1 General hazards

The North Sea is one of the most densely populated and major economic coastal region in the world (Quante and Colijn, 2016; Wahl et al., 2013). In the future, the region will be challenged by various climate change impacts like temperature change, changes in solar radiation, precipitation, and evaporation, droughts, sea level rise (SLR), storms, changes in wave conditions, and storm surges (Oost et al., 2017). However, it is quite difficult to make regional or small-scale accurate predictions for specific meteorological and oceanographic (metocean) parameters due to the great natural variability of the North Sea region (Oost et al., 2017; Quante and Colijn, 2016). Ecosystems are often more influenced by anthropogenic factors like land use change, agricultural practices, river management, and pollutant emissions than by climate change itself, even though these factors in turn drive anthropogenic climate change (Quante and Colijn, 2016).

Despite those uncertainties, specific changes exist. The *Norddeutsche Klimabüro* published an information website for the public, containing several climate change scenarios in the Wadden Sea area for temperature, precipitation, wind, and vegetation (`www.coastalatlas.org`).

The yearly mean air temperature has increased in the last decades over the entire North Sea region and is projected to increase between 1.7–3.2°C until the end of the 21st century, with respect to the period 1971–2000 for different IPCC scenarios (RCP4.5 and RCP8.5). Also, the mean surface temperature of the south-east North Sea has warmed by about 1°C since the end of the 19th century. Additional studies point out, that the sea surface temperature of the North Sea will increase further until the end of the century, while the projected magnitude depends heavily on the emission scenario and calculation method (Huthnance et al., 2016; Oost et al., 2017; Schrum et al., 2016).
Another effect of climate change is the decrease in the pH-value of the North Sea in recent decades. In combination with the rising surface temperatures, this leads to a lower uptake of atmospheric carbon dioxide (CO_2) by the North Sea as a sink and thereby amplifies climate change (Huthnance et al., 2016; Schrum et al., 2016).
Extreme wave heights have increased by about 8 cm per decade over the period 1985–2018 in the North Atlantic Ocean due to SLR. They lead to extreme sea level events, which therefore also have increased (IPCC, 2019). Analysis of data from Cuxhaven showed an increase in extreme sea level events for more than 150 years, but the uncertainties for future projected changes in extreme sea levels and extreme waves are large, even if it is likely that they increase in the German Bight towards the end of the 21st century (Weisse et al., 2012). Consequences are coastal erosion and flooding.
The IPCC (2019) claimed a possible increase in global precipitation, winds and storms. Changes in storm surge heights are closely linked to SLR and variations in the wind patterns, although no reliable climate-related changes can be determined for the North Sea region. Only modeling indicates an increase in storm surge heights, with half of the increase driven by natural variability of local weather and the other half by climate change (Oost et al., 2017; Weisse et al., 2012).

The marine and coastal ecosystems of the North Sea are also affected by climate change. Biological processes and organizations are influenced at all scales, leading to changes in the distribution, dynamics, and evolution of populations and the biodiversity of species. For example, species from warmer waters are more abundant today, which will impact sustainable levels of harvesting and various ecosystem services (Brander et al., 2016; Quante and Colijn, 2016). Habitats in dunes and salt marshes may be threatened by SLR, changing wave climate, storm surges, and natural sediment transport, as well as human impacts on geomorphology and sediment transport. Invasive plant species may grow faster and suppress other plants due to competition (Bakker et al., 2016).

These ecosystem changes have an impact on socio-economic factors such as fisheries and agriculture. While fish stocks and communities will change, fisheries and management need to adapt to these changes (Pinnegar et al., 2016). For northern areas of the North Sea region, an increase in agricultural productivity is expected, while the southern areas may suffer from summer droughts. In general, a doubling of atmospheric CO_2 could lead to a yield increase of up to 40% for some crops in the North Sea region, provided that sufficient water and nutrients are available (Olesen, 2016).
Regarding the field of offshore energy systems, several studies have identified an increase in hydropower potential, while the development of the potential of other renewable energy sources is uncertain. Energy and offshore activities are further very vulnerable to extreme weather and wave events, storms, and storm surges (Halsnæs et al., 2016).
Another important socio-economic and climate-sensitive sector is tourism, which plays a major role in the coastal zone of the North Sea region. Summer tourism is expected to benefit from rising air and water temperatures and a longer season, although it must be noted that the tourism sector in turn contributes to climate change (Kreilkamp et al., 2016).

Accurate analysis of climate-related hazards is difficult owing to the high natural variability in the North Sea region. For example, the present coastal erosion is immense but irregular with some coastlines accreting. For scientists, it is therefore a current challenge to isolate and project the regional impacts of climate change from the natural variability in the North Sea region (Quante and Colijn, 2016).
Besides the above described climate-related impacts, the main hazard to the coastal population is the SLR of the North Sea (c.f. Section 1.2.2), which leads to increasing wave heights and wave periods of the depth-limited sea state in the coastal areas (Lewis et al., 2019; Niemeyer et al., 2016).

1.2.2 Sea level rise

The GMSL is rising with an accelerated rate in the recent decades. While the SLR was about 1.4 cm per decade from 1901–1990, the rate increased to about 3.6 cm per decade for the period 2006–2015 (IPCC, 2019). The main cause of the GMSL rise since 1970 is anthropogenically forced, but nevertheless it is the result of different geophysical and climatological processes with various contributions to the total GMSL rise (Frederikse et al., 2020; IPCC, 2019). The main contributions come from ice loss of the Greenland and Antarctic ice sheets and glacier mass loss with about

1.8 cm per decade over the period 2006–2015. This exceeds the contribution of thermal expansion of the oceans with about 1.4 cm per decade for the same period (IPCC, 2019).

While rates on a global scale give an appropriate overview of future trends, rates of SLR on a local scale can still vary widely (Wahl et al., 2013). Meteorological and oceanographic effects have a major influence on local SLR, as well as variations in vertical land movements caused by natural or anthropogenic geological processes (Church and White, 2011; Nicholls et al., 2007; Wahl et al., 2013). For the entire North Sea region, a long-term SLR for the period 1900–2011 of about 1.5 cm per decade was estimated (Wahl et al., 2013). Recent analyses by the Helmholtz-Zentrum *Hereon* based on tide gauge data showed a SLR of about 8 cm over the last 50 years (1.6 cm per decade) for the island of Norderney in the German North Sea and a predicted further rise (Helmholtz-Zentrum hereon).

Taking a closer look at the gauge Norderney (Fig. 1.7), it showed that due to different long-term increases in tidal Mean High Water (MHW) and tidal Mean Low Water (MLW) from 1922–2021, the tidal range has increased by about 1.4 cm per decade in the last century (Fig. 1.1). Albeit natural variability of the SLR, a continuous positive linear trend of about 2.5 cm per decade for the MHW and about 1.1 cm per decade for the MLW can be observed from the data. The mean water (MW) given by the mean of MHW and MLW shows a linear trend of 1.8 cm increase per decade (Fig. 1.1). Significant trend changes are not apparent from Fig. 1.1; the fluctuations around the trend are rather direct consequences of meteorological conditions (Berkenbrink and Wurpts, 2019). More details on the study area are given in Section 1.3.3.

Thus, the SLR of the North Sea region is within the range with the GMSL rise. Extreme sea level events increased mainly due to SLR (Quante and Colijn, 2016). For further projections, the latest increase can be extrapolated linearly, although uncertainties persist regarding the dependency and e.g. according to the IPCC (2019) emission scenarios (RCP2.6 to RCP8.5), a wide range of future rates of increase is conceivable, leading coastal erosion and flooding (Harvey and Nicholls, 2008; IPCC, 2019). SLR also leads to an increase in wave heights and wave periods, which in turn results in a non-linear increase in the loading of coastal protection structures (Holthuijsen, 2007; Lewis et al., 2019).

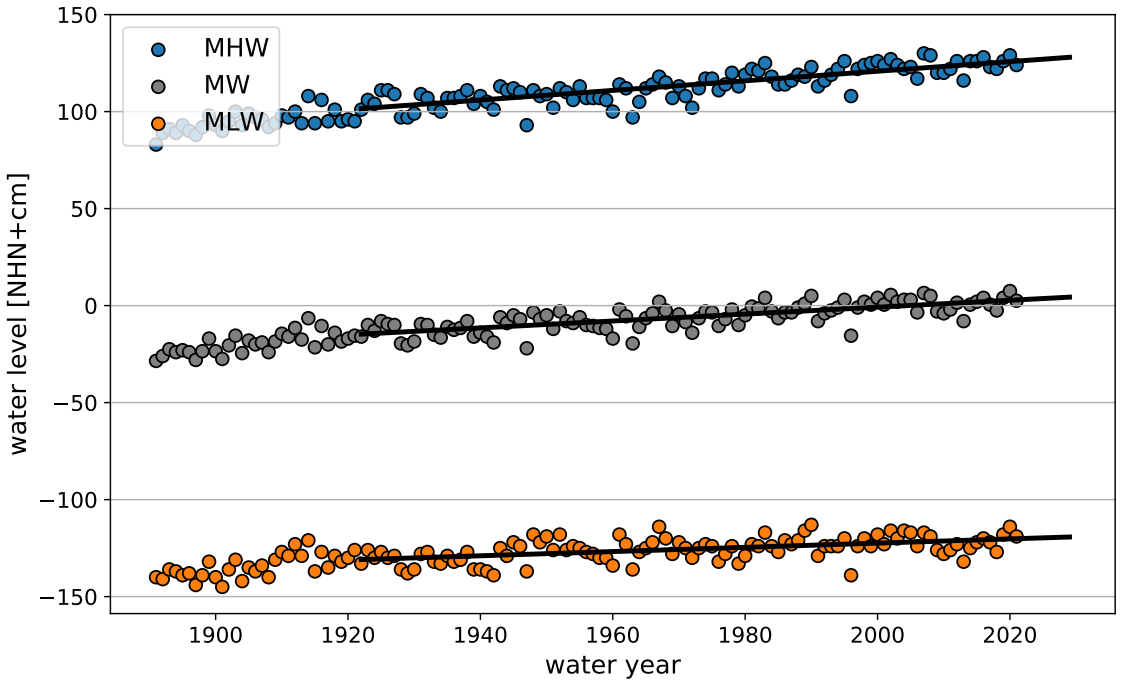

Figure 1.1: *Time series of the tide gauge level Norderney Riffgat, showing the Mean High Water (MHW), Mean Low Water (MLW), and Mean Water (MW) which is equal to the mean of MHW and MLW (gauge level operator and data source WSA-Emden).*

1.2.3 Coastal protection

Coastal areas are under tremendous threat from extreme events such as storms, storm surges, and SLR, which contribute to flooding, erosion, coastal retreat, and ecosystem loss (c.f. Sections 1.2.1; 1.2.2). These events can impose significant costs on coastal communities, while there is considerable regional variation (Nicholls et al., 2007). Large parts of the North Sea coastline are therefore highly defended by coastal protection structures as part of an integrated coastal protection system (Wahl et al., 2013). Every North Sea country has its own coastal protection strategy with specific geographical boundary conditions, planning schedules, and governmental regulations. But all countries face the same challenge of climate change, making them consider a general adaption strategy for future coastal protection: Coastal protection structures must be economic to construct in the short-term, on the one hand, and easily adaptable in the future, on the other hand, to remain flexible with respect to climate change impacts that are unknown today (e.g. masked due to high natural variability). The use of soft measures, which can be temporarily effective, should also be implemented, as they might prevent counteracting natural trends (Niemeyer et al., 2016).

Coastal protection on the North Sea coast has a long tradition of more than 1 000

years. Most parts of the southern North Sea countries of Denmark, Germany, the Netherlands, Belgium, and the United Kingdom are protected by a line of dikes, solid structures, dunes, or cliffs. The challenges of climate change lead to adaptations in coastal protection strategies, such as the do-nothing approach (i.e. release specific areas to the ocean), retreat of the protection line by managed realignment, reinforcement and maintenance of the existing protection line, wetland restoration, and a flexible coastal protection response to erosion. Erosion is traditionally controlled by sand nourishment. To hold the current coastline under future SLR, the nourishment approach is required to be applied more frequently and on a larger scale. Considering suitable sediment reservoirs with sufficient volume for sand nourishment raises the question of sustainable management (Dronkers and Stojanovic, 2016; Niemeyer et al., 2016).
All North Sea countries have developed adaption strategies to cope with the yet known impacts of climate change and are prepared for the challenge of a SLR of 1 m until 2100. Although there is already a lot of expertise in these countries, many knowledge gaps still exist, and the countries would therefore benefit from joint research management (Dronkers and Stojanovic, 2016).
Especially for morphodynamic complex areas like the East Frisian Islands, it will be a major challenge to model and predict morphodynamics of future climate change scenarios (not forecasts), besides the consideration of SLR, to meet the knowledge required for the design of adaption measures for coastal protection (Niemeyer et al., 2016; Ritzmann and Kaiser, 2010).

In Germany, coastal protection is controlled by the federal states. The Coastal Research Station of the Lower Saxony Water Management, Coastal Defence, and Nature Conservation Agency (NLWKN) is responsible for planning and construction of the federal island and coastal protection system in Lower Saxony. As a reaction to the severe storm surge of 1953, a change from a reactive coastal management strategy to a precautionary strategy was performed in Germany, and to the severe storm surge of February 1962, the State of Lower Saxony passed the Lower Saxony Dike Law (dt.: Niedersächsisches Deichgesetz (NDG)). This makes it the only German federal state with its own dike law (Lippe). Since than, dikes are build under consideration of further changes of the wave climate – mostly forced by climate change. Although there is no clear evidence of an increase in the frequency of storm surges, SLR currently is taken into account by increasing the precautionary measure of the design water level from 50 cm to 100 cm (Berkenbrink and Wurpts, 2019). Since 1997, the numerical model SWAN (c.f. Section 1.3.2) is used for the calculation of the design sea state for coastal protection structures, considering a

design water level given by the single-value method on basis of a severe storm surge. The design sea state and the resulting design dike height also take into account the expected highest tidal high water and the local wave run-up (Niemeyer, 2001).
Every year, the federal state of Lower Saxony invests more than 60 million euros for coastal protection (Stolz, 2021). In 2020, about 1.1 million people and a damage potential of about 129 billion euros were protected by the main dikes and barrier structures (NLWKN Pressestelle, 2020). Since 1955, a total amount of about 3 billion euros was put into Lower Saxony's coastal protection (Lippe).
The main dike line of the entire federal state of Lower Saxony is about 610 km long (Stolz, 2021). The dikes of the East Frisian Islands of Lower Saxony have a total length of 35 km, while the dunes on the islands have a total length of about 97 km (NLWKN, 2022). The highest dikes are up to 9 m high (Lippe). In case of any further scientific evidence, the dikes can be raised by one additional meter. For a sustainable and safe coastal protection, it is essential to check coastal protection structures regularly and include new scientific knowledge about the climate-related hazards continuously (Berkenbrink and Wurpts, 2019).

1.3 State of the art and basic related work

The following chapter provides an overview of the statistical, technical, and physical basics that are employed in the remainder of this work. For further details, the reader is referred to read the respective research studies (c.f. Chapters 2, 3, 4). The study area of the research studies, mainly characterized by the ETD sandbanks, is finally presented in detail.

1.3.1 What waves are and how they are measured and modeled

The spatio-temporal propagation of ocean waves is important for coastal management, protection, and operations (Ti et al., 2018). Ocean waves can be defined as all (directional) vertical motions of the ocean surface over space and time (Holthuijsen, 2007). To study waves, proper observations of wave height, wave period, and wave direction are essential (Teich et al., 2018). Besides visual observations of the sea surface by experts, advanced observation techniques with instruments are commonly used. One distinguishes between *in-situ* measurements taken directly in the ocean by floating wave buoys, pressure transducers, current meters, echo-sounder at the sea bottom, or wave pole measurement platforms and *remote sensing* techniques taken from above the sea surface by drones, airplanes, satellites, or land-based imaging with radar or laser (Holthuijsen, 2007).

All measurement techniques, including wave buoys, record time series of the (directional) vertical motion of the sea surface. Wave buoys can measure the three-dimensional acceleration of the surface water by moving up-and-down with the waves including the horizontal surge-and-sway motion for directional information. (Holthuijsen, 2007; Malcherek, 2010)

For the processing of buoy measurements and the determination of wave parameters, a continuous variance density spectrum $S(f)$ of frequency f can be calculated by a Fourier transformation of the buoy-measured sea surface elevation data (Fig. 1.2). Commonly, a period of 30 minutes is used for the spectral computation. The unit of $S(f)$ is $[m^2/Hz]$.

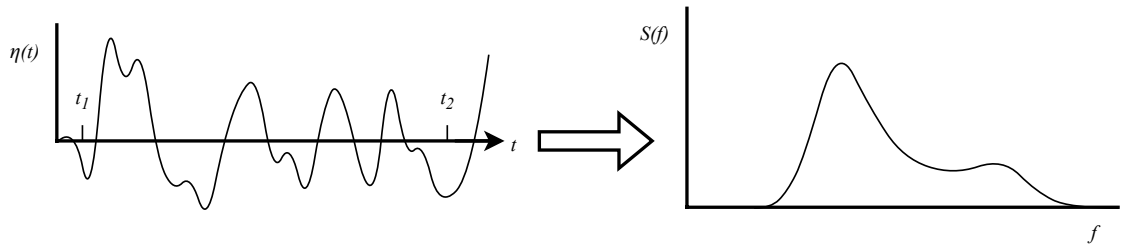

Figure 1.2: *Transformation of the surface elevation $\eta(t)$ of period $[t_1, t_2]$ to the variance density spectrum $S(f)$ with frequency f.*

Based on $S(f)$, the wave parameters significant wave height H_{m_0}, mean wave period $T_{m_{0,1}}$, energy period $T_{m_{-1,0}}$, peak period T_P, and wave direction Dir_P can be determined as follows:

$$H_{m_0} = H_S = 4\sqrt{m_0} \tag{1.1}$$

$$T_{m_{0,1}} = T_m = \frac{m_0}{m_1} \tag{1.2}$$

$$T_{m_{-1,0}} = T_e = \frac{m_{-1}}{m_0} \tag{1.3}$$

$$T_P = argmax(S(f))^{-1} \tag{1.4}$$

where m_n is the n-th moment of $S(f)$ given by

$$m_n = \int f^n S(f) df \tag{1.5}$$

and Dir_P is defined as the wave direction of the $S(f)$ peak (Bacon and Carter, 1991; Holthuijsen, 2007; Sheng and Li, 2017; SWAN, 2019b).

While wave buoys only provide single-point data, spatial wave fields can be provided by numerical wave modeling methods. In turn, single-point buoy data are used to calibrate and validate model parameters in numerical modeling. Commonly, the numerical wave model SWAN is used for complex coastal areas. Wave propagation in the third-generation wave model SWAN (Booij et al., 1999) is computed by the spectral action balance equation

$$\frac{\partial}{\partial t}N + \frac{\partial}{\partial x}c_x N + \frac{\partial}{\partial y}c_y N + \frac{\partial}{\partial f}c_f N + \frac{\partial}{\partial \theta}c_\theta N = \frac{S}{f} \tag{1.6}$$

where $N = N(f, \theta)$ is the action density spectrum with frequency f and wave direction θ, S the variance density spectrum, t the time, and x and y the spatial propagation directions with velocities c_x and c_y, respectively. The fourth term describes the changes in depth and currents, while the fifth term represents directional changes caused by refraction (Booij et al., 1999; SWAN, 2019a,b). As coastal processes like shoaling, refraction, diffraction, reflection, and wave breaking are considered by SWAN, it is a suitable model for the nearshore area of the East Frisian North Sea (Holthuijsen, 2007).
By using unstructured computational grids (compared to regular/rectangular or curvilinear grids), a specific focus can be put on certain morphodynamic features in the bathymetry and an efficient computation can be performed concurrently (SWAN, 2019b). A detailed description of the SWAN model parameter and calibration process is given in Sections 2.4.3 and 4.2.4. While most studies and applications use only static bathymetries (Akpinar et al., 2016; Hoque et al., 2019; Kutupoğlu et al., 2018; Sierra et al., 2017), dynamically changing bathymetries can be used to analyze morphological impacts on the wave climate (Dallas and Barnard, 2011; Eshleman et al., 2007).

1.3.2 Neural networks and their application for wave height prediction

Machine learning

Artificial neural networks (ANN) are a class of learning algorithms of deep learning (DL), which belong to the field of artificial intelligence (AI). The data-driven machine learning (ML) algorithms are inspired from the functionality of the human brain. Albeit ANNs have been known for many decades (first attempts date back to 1943 (McCulloch and Pitts, 1943)), their vast potential has only been unleashed by efficient learning algorithms, increased computational power, and optimized architectures. Moreover, the availability of training data in many domains has increased enormously in recent years (Goodfellow et al., 2016). Thus, ANNs have developed

tremendously and revealed outstanding performance in many fields of ML like computer vision for image classification (Sultana et al., 2018; Thai et al., 2012), face recognition and object detection (Redmon et al., 2016), and semantic segmentation (Yuan et al., 2020) as well as non-linear regression problems like time-series prediction (Hastie et al., 2009; Kuhn and Johnson, 2013; Song et al., 2020).
ML algorithms are categorized in supervised, unsupervised, semi-supervised, reinforcement, and transfer learning, while prominent examples for the supervised and unsupervised learning are the classification/regression task and clustering, respectively (Hastie et al., 2009). The algorithms used in this dissertation belong to the supervised learning approach.

Basics of neural networks

A simple feed-forward neural network (FFNN) or multilayer perceptron consists of an input, output, and hidden layer as well as bias units without inputs and forms the basis of DL models. In case a FFNN has multiple hidden layers, it is called a deep FFNN (DFFNN). An exemplary DFFNN architecture is shown in Fig. 1.3. The strength of the connection between two neurons is expressed by weights, which also represent the learnable parameters (degrees of freedom). If a connection between all neurons of consecutive layers exists, the network is called fully-connected, as depicted in Fig. 1.3. The hidden and output layers additionally contain a so-called activation function f, which is chosen depending on the type of function to be approximated. Typical non-linear activation functions are the sigmoid function, hyperbolic tangent (tanh), and rectified linear unit (ReLU) function (Aggarwal, 2018; Goodfellow et al., 2016; Hastie et al., 2009; Khan et al., 2020).
By means of a loss function L, the prediction error of the ANN can be estimated based on the ground truth data and then minimized by backpropagation with an optimization function (Goodfellow et al., 2016). The mean squared error (MSE) with $L2$ regularization for reducing overfitting is widely used as loss function L for regression problems:

$$L(X, W) = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 + \frac{1}{2} \lambda \|W\|_2^2, \tag{1.7}$$

where n is the size of data, $\hat{y}$ the predicted value, y the ground truth value and λ the regularization parameter of the L2-norm $\|W\|_2$ of the weight matrices W (Goodfellow et al., 2016; Ng, 2004; Nie et al., 2010). Common optimization algorithms for the backpropagation of the gradients of the loss function are the adaptive moment estimation (Adam) (Kingma and Ba, 2015), stochastic gradient descent (SGD) (Bottou, 1998), AdaGrad, RMSProp, and AdaDelta.

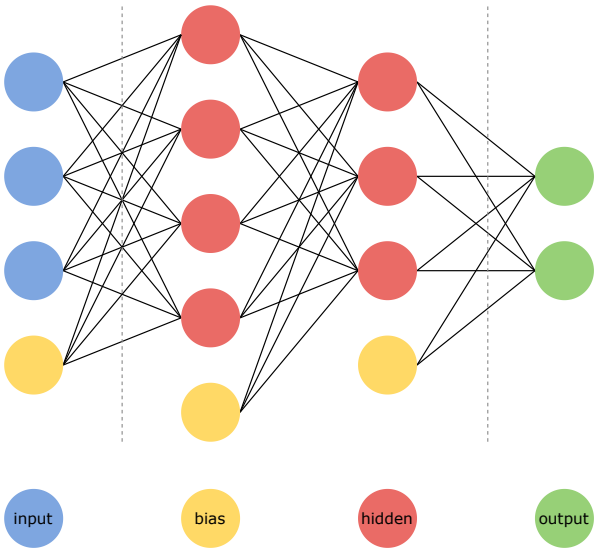

Figure 1.3: *Deep feed-forward neural network (DFFNN) with input neurons (blue), two hidden layers with different numbers of neurons (red), output neurons (green), and bias units (yellow), which is fully-connected.*

To validate the accuracy of ANN predictions, error metrics like the root mean squared error (RMSE) (Eq. 1.8), the mean absolute error (MAE), the mean error (bias), the mean arctangent absolute percentage error (MAAPE) (Eq. 1.9) (Kim and Kim, 2016), the Pearson correlation coefficient r (Eq. 1.10), and the coefficient of determination R^2 (Eq. 1.11) are often used.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{1.8}$$

$$MAAPE = \frac{1}{n}\sum_{i=1}^{n} arctan\left(\left|\frac{\hat{y}_i - y_i}{y_i}\right|\right) \tag{1.9}$$

$$r = \frac{cov(\hat{y}, y)}{\sqrt{(var(\hat{y})var(y)}} \tag{1.10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1.11}$$

where $cov(x, y)$ is the covariance of x and y, $var(x)$ the variance of x, and $\bar{y}$ the mean of the ground truth.

Long-Short term memory neural networks

Long-Short term memory (LSTM) neural networks are special recurrent neural networks (RNN) designed by Hochreiter and Schmidhuber (1997). While all RNNs

include feedback connections to prior inputs in order to transfer information over time (LeCun et al., 2015), the LSTMs are able to learn long term dependencies and coping with the so-called vanishing gradient problem due to their cell structure (Gers et al., 2000; Hochreiter and Schmidhuber, 1997; Pascanu et al., 2013). Each LSTM cell (Fig. 1.4) consists of four units, the forget gate f_t, input gate i_t, memory cell state C_t, and output gate o_t at time t, as defined in Eq. 1.12 to Eq. 1.17.

$$f_t = \sigma(W_f h_{t-1} + W_f x_t) \tag{1.12}$$

$$i_t = \sigma(W_i h_{t-1} + W_i x_t) \tag{1.13}$$

$$\tilde{C}_t = tanh(W_C h_{t-1} + W_C x_t) \tag{1.14}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{1.15}$$

$$o_t = \sigma(W_o h_{t-1} + W_o x_t) \tag{1.16}$$

$$h_t = o_t \cdot tanh(C_t) \tag{1.17}$$

where x_t, h_t, $\tilde{C}_t$, W, and σ are the input data, hidden state, temporary cell state, weight matrices, and sigmoid activation function, respectively. The Hadamard matrix product operation is denoted as $\cdot$. Biases are not added explicitly, since they can be integrated to the weight matrices. The final output $\hat{y}_t = W_y h_t$ can be used in the loss function from Eq. 1.7 (Gers et al., 2000; Hochreiter and Schmidhuber, 1997; Qing and Niu, 2018).

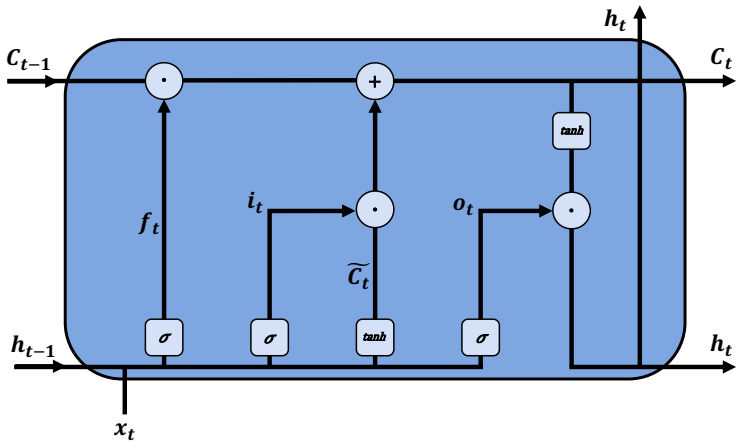

Figure 1.4: *LSTM cell with forget gate f_t, input gate i_t, memory cell state C_t, and output gate o_t at time t, and the sigmoid and tangent hyperbolic activation functions σ and tanh, respectively.*

Convolutional neural networks

Convolutional neural networks (CNN) are especially designed for data on grid topologies (e.g. images) by learning local patters and shift invariants (LeCun et al., 1989). CNNs are therefore often applied for image classification, face and speech recognition, and natural language processing (Aggarwal, 2018; Goodfellow et al., 2016). The principle of a convolution operation by a so-called kernel filter is exemplarily shown in Fig. 1.5. By moving over the input, the convolution operation learns latent features and extracts them to the resulting output feature maps (Goodfellow et al., 2016).

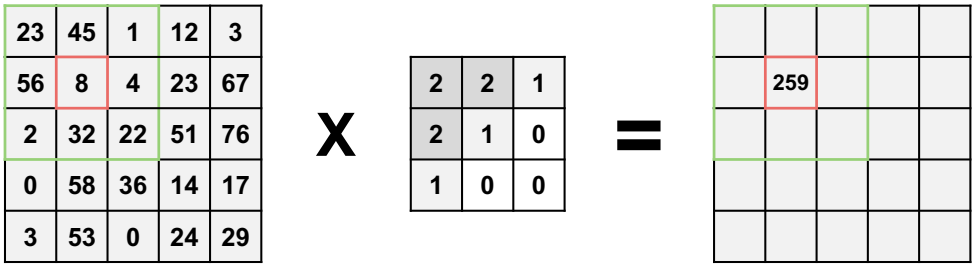

Figure 1.5: *Example of a convolution with a kernel of size 3 × 3. The output of the convolution is called the feature map.*

A convolution layer is defined by the number and size of the convolution kernels, the step size of the kernel (stride), and the number of rows and columns added to the boundaries of the input feature matrices for adjusting the size of the output feature map (padding). The learnable parameters of a CNN are the weight matrices. The convolution layer operation $*$ for an layer l and output feature map k with input feature matrix $X_k = (x_{ijk})_{i,j=1}^{n,m} \in \mathbb{R}^{n \times m}, n, m \in \mathbb{N}$ and weight matrix of the convolution kernel $W_k = (w_{i'j'kk'})_{i',j'=-r}^{r} \in \mathbb{R}^{(s \times k) \times (s \times k')}, s = 2r + 1, r \in \mathbb{N}$ is performed by

$$z_{ijk}^{(l)} = (W * X)_{i,j,k}^{l} = \sum_{k'=1}^{d^{(l-1)}} \sum_{i',j'} x_{i+i',j+j',k'}^{(l-1)} \cdot w_{i',j',k',k}^{(l)} \tag{1.18}$$

$$x_{ijk}^{(l)} = f\left(z_{ijk}^{(l)}\right) \tag{1.19}$$

where $Z_k = (z_{ijk})_{i,j=1}^{n,m} \in \mathbb{R}^{n \times m}$ is the output feature map of the convolution and $d^{(l)}$ the number of feature maps of layer l.

The architecture of a CNN can be composed of several additional layers like pooling layer, batch normalization, dropout layer, and fully-connected layer. Pooling layer are commonly used to reduce the size of the feature maps and thereby saving computational cost and generalize by extracting the main features. The most popular

variants of pooling are the max pooling and average pooling (Fig. 1.6). More information on CNNs can be found in Sections 4.2.6 – 4.2.8.

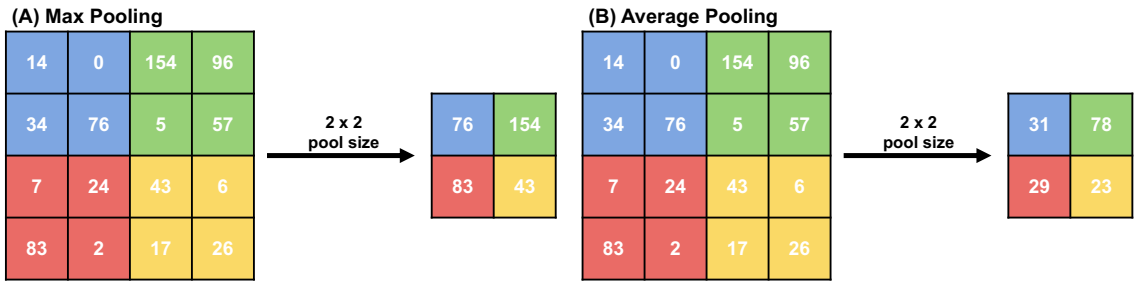

Figure 1.6: *Example of the (A) max pooling and (B) average pooling approach with a kernel of size 2 × 2 and a stride of 2, respectively.*

Neural networks for wave height prediction

Neural networks have long been used to predict environmental and climate data. Consequently, there are several studies that applied ANNs for wave height prediction (Ali and Prasad, 2019; Deka and Prahlada, 2012; Dixit and Londhe, 2016; Durán-Rosal et al., 2016; Kumar et al., 2017; Londhe, 2008; Stefanakos, 2016). Even though many different architectures, such as FFNN (Deo and Naidu, 1999; Deo et al., 2001; Herman et al., 2007, 2009; Law et al., 2020; Londhe and Panchang, 2007; Peres et al., 2015; Zamani et al., 2008), genetic algorithms (Altunkaynak, 2013), evolutionary product unit neural networks (Durán-Rosal et al., 2016), or extreme learning machine (Kumar et al., 2018) have been developed for sea state prediction, the focus in the following literature review will be on LSTM and CNN architectures as they are used in this dissertation.

Since LSTMs are particularly suitable for time series data, they have already been implemented for reconstructions and predictions of sequential environmental data such as wind speed (Li et al., 2018), sea surface temperature (Xiao et al., 2019), solar irradiance (Qing and Niu, 2018), and weather forecasting (Salman et al., 2018). First applications of a RNN for sea state prediction date back to Balas et al. (2004). Pirhooshyaran et al. (2020) first designed a more sophisticated encoder and decoder LSTM neural network to reconstruct and predict deep ocean wave heights in the North Pacific and North Atlantic Ocean. They achieved an R^2 of up to 0.81 for the SWH reconstruction at a deep ocean buoy near Los Angeles. Kagemoto (2020) focused on the prediction of water-surface wave trains and the motion of a containing

floating body, while Ni and Ma (2020) concentrated on the prediction of polar westerlies wave heights based on few measurements available, both using LSTM neural networks. Fan et al. (2020) also applied LSTMs for up to 6 h ahead predictions of SWH at the deep ocean with an R^2 of 0.97. Additionally, they combined the LSTM with data of the numerical wave model SWAN and reached an improvement by more than 65% for a single-point prediction. Another study of James et al. (2018) also utilized FFNNs trained on SWAN-modeled data, here for the Monterey Bay area. With a reduction of the computational time by a factor of more than 100 compared to the physics-based SWAN model, they achieved a satisfactory result for the spatial SWH surrogate prediction model. Feng et al. (2020) applied SWAN-modeled input data from 2005–2014 to a FFNN for the prediction of SWH and wave peak period of Lake Michigan. They were able to reduce computational time by a factor of up to 20 000 while keeping comparable accuracy to SWAN. Chen et al. (2021b) on the other hand developed a random forest surrogate model for the spatial prediction of nearshore wave fields using SWAN-modeled ground truth data. The performance for single-point predictions on buoy data was better for the surrogate model ($R^2 = 0.91$) than for the corresponding SWAN model ($R^2 = 0.85$), while computation time was additionally reduced by a factor of 100. Zhou et al. (2021) presented a two-dimensional convolutional LSTM neural network for SWH prediction in the South and East China Seas with lead times up to 24 h. Trained on WAVEWATCH III data, they achieved a MAPE of 61% for the 24 h forecast. Zhou et al. (2021) have combined the approaches of LSTM for temporal data and CNN for spatial data in their network architecture.

CNNs have been rarely used to predict two-dimensional wave fields since the training data usually need to be generated by numerical models. Bai et al. (2022) designed a CNN model for spatial SWH prediction in the South China Sea using ERA5 reanalysis data. Forecasts with lead times from 12–72 h were performed with MAPEs between 8.55% and 19.48%, respectively. Another study of Wei and Davison (2022) used a CNN model to predict waves and hydrodynamics in a nearshore area by applying SWASH-modeled ground truth data. Their model accurately predicted wave propagation and wave breaking over sandbanks. Choi et al. (2020) applied the CNN on raw ocean images captured on wave buoys for real-time predictions of SWH due to their efficiency on image data.

1.3.3 Study area and ETD sandbanks

The study area of this dissertation covers the nearshore area of the East Frisian Island of Norderney in the German Bight of the North Sea (Fig. 1.7). The area extends about 7.05–7.3° E and 53.65–53.85° N. Norderney is separated by the tidal inlets Norderneyer Seegat and Wichter Ee from the islands Juist in the west and Baltrum in the east, respectively. With a length of about 15 km, a maximum width of 2.5 km, and a resulting total area of 26.3 km^2, Norderney is the second largest of the seven East Frisian Islands. While about two thirds of the island are dune and sand landscapes, the urban area is located on the western end, reflecting the earlier eastward migration by erosion.

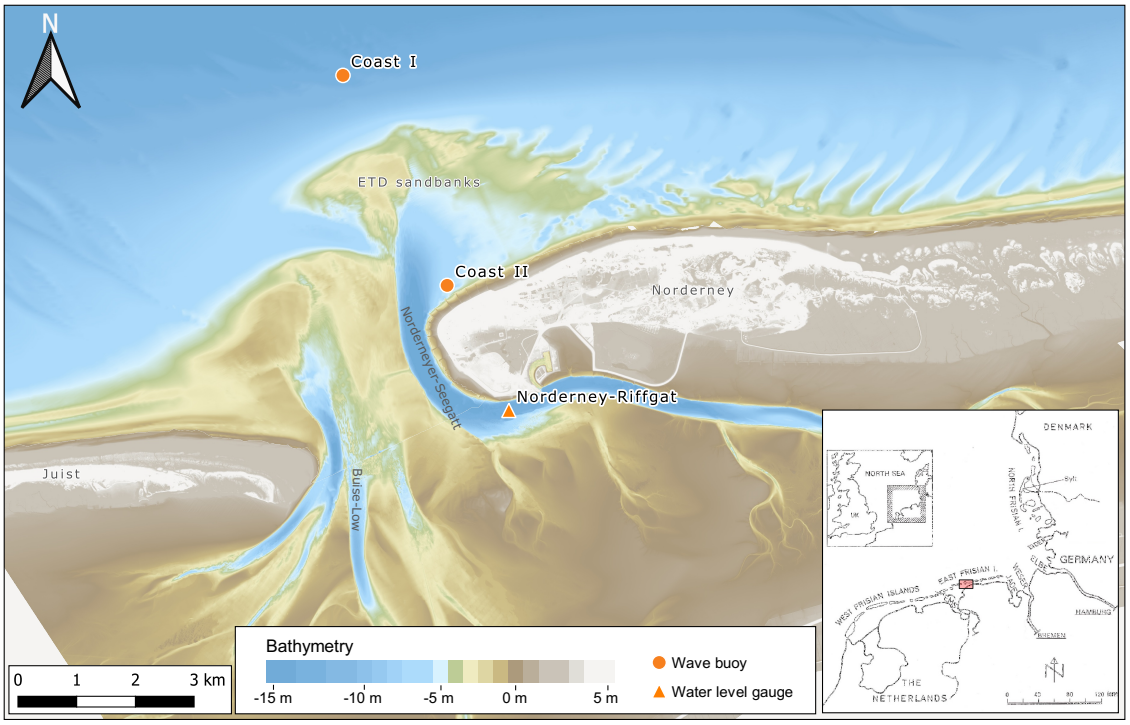

Figure 1.7: *Bathymetry map of the study area, showing the ETD sandbanks, wave buoy positions, and the tide gauge level Norderney Riffgat.*

Located in the temperate climate zone, the predominant wind direction in the study area is southwest (Fig. 1.8). The long-term MHW at the water level gauge Norderney Riffgat is 1.23 m above the German mean sea level (NHN) and has a mean tidal range of 2.47 m (Fig. 1.7). Due to the oceanic influence, the current mean water temperature is about 14°C. More details on SLR in the study area are given in the previous Section 1.2.2.

The genesis of the East Frisian Islands started about 8 500 years ago with the warming after the Last Glacial Period of the Weichselian glaciation. By eustatic SLR, the Dogger Bank and English Channel were flooded. Driven by the Coriolis force, an eastward tidal flood at the East Frisian North Sea has then developed. When the rate of SLR decreased significantly about 4 000 years ago, the North German upland moor and peat were formed. After alternating phases of transgression and regression, the interaction of tidal currents and waves resulted in the final accumulation of sand sediments that today form the East Frisian Islands (Meyer, 2014; Petersen, 2005; Quedens, 2000).

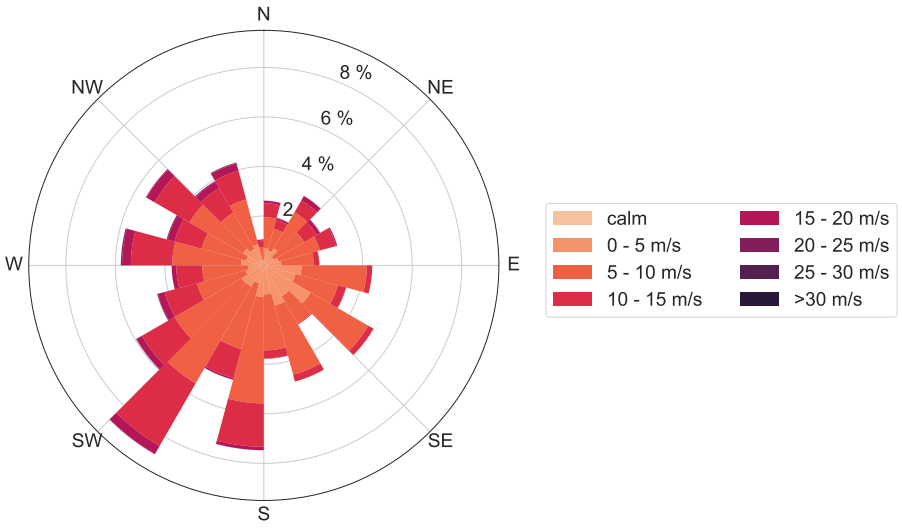

Figure 1.8: *Windroseplot of averaged data from 2004–2017 of the DWD station Norderney. Wind direction, speed, and frequencies are shown.*

Norderney is a typical East Frisian barrier island with strong tidal currents leading to a massive drive of the sediment transport of the ETD sandbanks in the tidal inlets. The eastward sediment transport parallel to the coast is disturbed by the tidal currents in the main channel Norderneyer Seegat. On the west coast, the sand subsequently accumulates to sandbanks. By wave breaking and energy dissipation, the sandbanks protect the western part of the island. On its curved way to the eastern part of Norderney, the sand of the ETD additionally supplies the beaches (NLWKN, 2019). The mean velocity of the different plates of the ETD varies between 250–430 my^{-1}, depending on the plates position in the cycle of about 37 years in total (Bremermann and Meyer, 2012). Older studies quantified a mean movement of 400 my^{-1} (Fitzgerald et al., 1984) and 325 my^{-1} in the west and 500–550 my^{-1} in the east (Gaye and Walther, 1935). Ebb-tidal currents and storm surges change the morphodynamics of the ETD continuously and thus dominate the local wave climate (Wang et al., 2012).

The western plates north of the Buise-Low (Robbenplate and Osterriff) are periodically incised or separated by the Buise-Low (Fig. 1.7). Outflow of the Norderneyer Seegat therefore varies from east of the northern plates and between the two plates of the northern ETD sandbanks, if they are separated (Fig. 1.7). Thus, both plates of the northern ETD underlay huge ongoing morphodynamic changes. Accretion of the ETD sandbanks is constant in the central western part of Norderney (Meyer, 2014). Besides the morphodynamic processes, the interaction of tidal range and waves is important for the evolution of the barrier island Norderney (Quedens, 2000).
Continuous morphodynamic change and natural variability are highly dynamic and must be considered in further analyses of the study area. A detailed analysis of the morphodynamic changes in shape and size of the ETD sandbanks since 1995 is given in Section 2.3.2. For further information on general morphological patterns in tidal inlets, the reader is referred to Hayes (1980).

1.4 Research questions, aims, and objectives

The previous review of state of the art knowledge and recent publications highlighted the great potential of neural networks for precise predictions of sea state data. Moreover, the importance of morphological changes of the ETD sandbanks off the East Frisian Islands for the nearshore sea state climate was underlined.
Based and motivated by the prior work, the following research questions (RQ) were inferred, aiming to contribute to better understand spatial and temporal predictions of nearshore wave height with neural networks considering coastal morphodynamics in light of coastal protection and climate risks:

RQ-a. How can the influence of coastal morphodynamics and its variability on the nearshore sea state be quantified?

RQ-b. How can data of coastal morphodynamics improve the single-point reconstruction and prediction of significant wave height by neural networks?

RQ-c. How can simulated bathymetries be used to predict two-dimensional significant wave height with convolutional neural networks?

In addition to application-related implications for the research area, methodological aspects were in the focus of the respective research. Different neural networks were designed to cope with coastal morphodynamics in wave height prediction. The methods were compared to other common machine learning methods for spatial and temporal predictions and the speed-up compared to numerical modeling using SWAN was considered.

In the following, the above mentioned open research questions (a–c) are addressed in three studies within this dissertation, each covering full methodological design, model experiments, discussion of results, and challenges:

I. *(RQ-a) How can the influence of coastal morphodynamics and its variability on the nearshore sea state be quantified? What is the magnitude of the sea-state damping effect in the Norderney research area, what does it depend on, and how is it changing over time?*

   Understanding the coastal morphodynamics in the East-Frisian North Sea area is crucial for precise predictions of the nearshore sea state. As stated in Section 1.3.3, the ETD sandbanks exert a huge impact on the wave climate (Niemeyer, 1979; Niemeyer and Kaiser, 1999). However, the damping effect has not yet been precisely quantified. Also differences between swell-waves and wind-waves were not yet investigated. In Chapter 2, these objectives were addressed by validated SWAN simulations with dynamic bathymetries (c.f. Section 1.3.1) for five representative storm events at tidal high water. This allowed isolating the effect of different bathymetries on the nearshore sea state to distinguish it from natural variability for the first time. In addition, a GIS-based approach was used to analyze the ETD sandbank dynamics by movement velocities and changes in size and shape to account for the problem of rapid changes by migration, erosion, and accretion (Castelle et al., 2007; Wang et al., 2012). Therefore, the bathymetries of the years 1995, 2005, 2006, 2008, and 2015 were used to cover a long interval of the complete live-cycle of a single plate moving from west to east (c.f. Section 1.3.3).

   Buoy-measured in-situ data are important for model validation (c.f. Section 1.3.1), thus data off the Norderney coast from 1992–2017 were compared to the isolated model results.

II. *(RQ-b) Are LSTM neural networks and an ensemble of parallel LSTMs suitable for single-point reconstruction and short- and long-term prediction of significant wave height? How can data of coastal morphodynamics (bathymetries) improve the model's performance?*

   Many studies already investigated the single-point reconstruction (hindcast) and prediction (forecast) of significant wave height with machine learning methods – mostly neural networks (c.f. Section 1.3.2) – but non of them considered complex coastal morphodynamics yet. Pirhooshyaran and Snyder (2020) introduced the sequence-to-sequence LSTM neural networks for

deep ocean wave height reconstruction and prediction. Fan et al. (2020) used LSTMs to improve numerical prediction of SWAN wave model (c.f. Section 1.3.1). In Chapter 3, the consideration of coastal morphodynamics for state of the art single-point significant wave height reconstructions and predictions was addressed in a coastal area for the first time using LSTM and ensemble of parallel LSTM neural networks. The LSTM was fed by time series of metocean input data coming from two wave buoys off the coast of Norderney and a measurement station of the German Meteorological Service (DWD), as well as linear bathymetry data of the ETD sandbanks from 2004–2017. The LSTM was compared to several machine learning algorithms for benchmarking and the proposed method was validated on well-studied offshore buoys off the U.S. west coast. By feature selection, the improvement of the LSTM with included bathymetric data was of special interest regarding the aim of this dissertation, albeit this study was limited to single-point reconstructions and predictions.

III. (RQ-c) How can bathymetries of coastal morphodynamics be simulated and used to predict two-dimensional significant wave height with convolutional neural networks? What speed-up compared to common numerical modeling can be reached using machine learning?

Since only limited data of ETD sandbank morphologies are available from hydrographic surveys (Spicer et al., 2019), most models in the literature – both numerical and machine learning – consider static bathymetries only (c.f. Section 1.3.1). In Chapter 4, this problem was addressed by simulating ETD sandbanks using a geostatistical approach based on variogram analysis and random fields simulation. Simulated ETD sandbanks were used for the first time to train a two-dimensional mixed-data deep CNN (c.f. Section 1.3.2), since the aforementioned methodological approach can extend the amount of coastal morphodynamic data. This aimed to generally consider arbitrary coastal morphodynamics in SWH predictions. Ground truth wave fields were modeled by SWAN (c.f. Section 1.3.1) and respective metocean data from 2004–2017, thus the resulting CNN model formed a SWAN surrogate model. Other studies like Zhou et al. (2021) reached immense speed-ups by machine learning (especially CNN) surrogates, compared to common SWAN numerical model. Reduction of computational time is of special interest for rapid coastal protection applications (Hu et al., 2021). The proposed CNN spatial model was used for the investigation of SLR scenarios, as well as the impact of possible bad ETD morphologies on coastal protection at the Norderney coastline,

since spatial morphodynamic dependencies are essential for accurate predictions. Validation was performed on 59 buoy measurements by comparing the performance of SWAN numerical model and proposed mixed-data CNN surrogate model.

The three introduced studies are the framework of this dissertation and will be elaborated in the following three Chapters 2, 3 & 4. In Chapter 5, an overall discussion of the results and findings is given and conclusions are made. Fig. 1.9 summarizes the framework of this dissertation. All research discussed in this dissertation has already been published in a peer-reviewed scientific journal or is under review (Appendix III: List of publications).

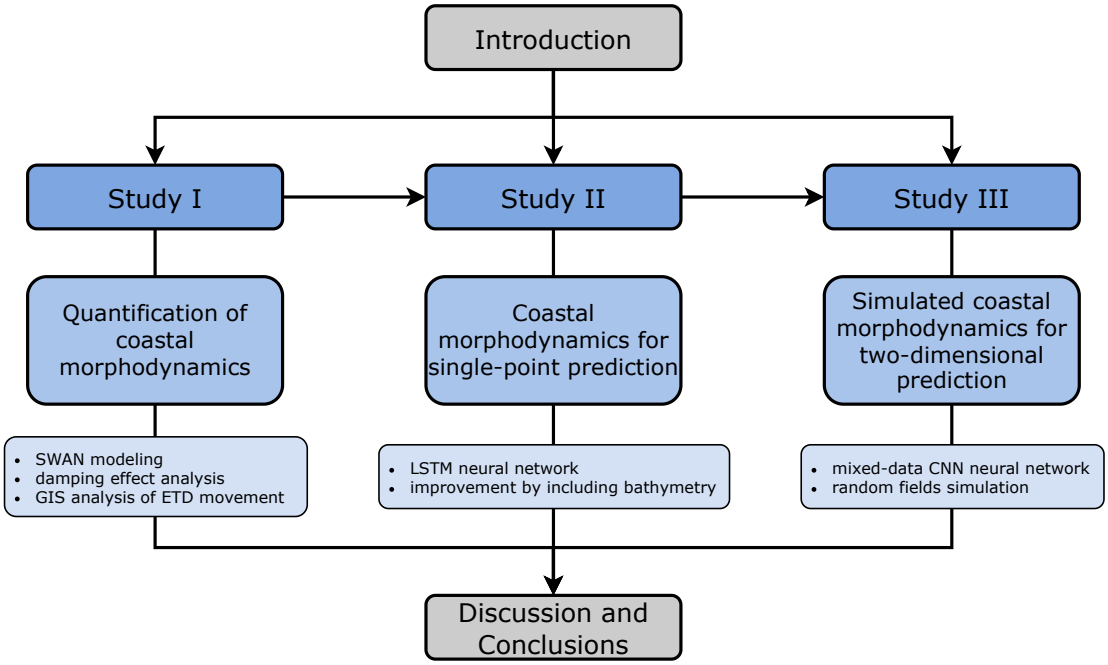

Figure 1.9: *Flowchart diagram representing the structure of the main part of this dissertation including the three consecutive studies with an overview of the corresponding methodologies applied in each one. The arrows indicate the direction of information flow between the research studies.*

2 Research Study I

Quantification of morphodynamic variability and sea state damping of plates at the nearshore area in the East Frisian North Sea

Co-Authors: Cordula Berkenbrink, Jannis Heil & Britta Stumpe

Published in Coastal Engineering (2021), 164 (accepted 30.08.2020)

Abstract

At the nearshore area of Norderney, the ebb-tidal delta (ETD) sandbanks exert a huge impact on the local wave climate, which is of particular interest for coastal protection. Besides an eastward migration of the sandbanks with a mean velocity from 435–491 my^{-1}, depending on the plates position, the ETD sandbanks had a high variability in shape and size. Therefore, the influence of those morphodynamic changes on the sea state damping effect was quantified by dynamic bathymetries using validated SWAN simulations. The impacts of the bathymetries from the years 1995, 2005, 2006, 2008, and 2015 with highest morphological variability were considered for five different boundary conditions, representing storm surge events of the Norderney nearshore wave climate. A mean variation of 14 ± 2% of the relative sea state damping effect was found for the different bathymetries of the ETD sandbanks at the various boundary conditions. Swell waves showed a range of 16 ± 2% for the damping effect, which increased with increasing incoming wave height. Buoy data off the Norderney coast from 1992 to 2017 were investigated and a mean buoy-measured damping effect of 0.37 was determined over the entire period. General results of SWAN modeling were also found in the in-situ measurements.

Keywords: sea state damping, SWAN, North Sea wave climate, ebb-tidal delta sandbank morphology, sea state separation, sea state variability

2.1 Introduction

Coastal areas are of great interest as present and future climate change impacts, such as sea level rise, will put increasing pressure on humans living in those regions. Also the possibility of an increase in extreme storm events, which may damage coastal areas, is conceivable (Bengtsson et al., 2009; Weisse et al., 2012). For the North Atlantic and the North Sea, various studies that have analyzed the wave climate identified an increasing trend in ocean wave heights (Bacon and Carter, 1991; Grabemann and Weisse, 2008; Mori et al., 2013; Vanem and Walker, 2013; Weisse and Gunther, 2007). To sufficiently assess coastal management risks and evaluate human interferences, detailed knowledge about the wave climate is of particular interest, not only for coastal protection, but also for offshore structures and operations (Kaiser and Niemeyer, 2007; Teich et al., 2018; Wang et al., 2012).

For such purposes, the state of the art third generation wave model SWAN (Simulating WAves Nearshore) is often applied for modeling design waves, wave climate, and special events in shelf seas, coastal, and nearshore areas (Booij et al., 1999; Hoque et al., 2019; Sebastian et al., 2014; Vieira et al., 2020). In many cases, static bathymetries are used for those applications (Akpinar et al., 2016; Hoque et al., 2019; Kutupoğlu et al., 2018; Sierra et al., 2017). But SWAN modeling with dynamical bathymetries is also an important and promising approach in coastal risk assessment. Thus, some studies estimated morphological impacts on the wave climate in coastal and nearshore areas (Dallas and Barnard, 2011; Eshleman et al., 2007). Due to local changes in the morphodynamic conditions and the interaction of numerous different influences, global wave climate issues are very heterogeneous on a local scale (Kaiser and Niemeyer, 2007; Wang et al., 2012). Various studies analyzed those morphological conditions and sediment dynamics of ebb-tidal deltas (ETD) in coastal areas (Anthony, 2013; Bergillos et al., 2016; Hansen et al., 2013). ETDs consist of sediments from ebb-tidal currents, deposited seawards of the tidal inlet (Dallas and Barnard, 2011). They are subject to rapid changes by migration, erosion, and accretion (Castelle et al., 2007).

The ETD sandbanks off the Norderney coastline have a major impact on the local wave climate, since they are a natural barrier against incoming waves, especially during extreme events for which only few in-situ buoy data are available (Niemeyer, 1979; Niemeyer and Kaiser, 1999). Wave breaking causes the sea state energy to be dissipated on the sandbanks before reaching the islands beaches, while many influences affect the damping, whose complexity is not yet entirely known even due to missing measurements (Niemeyer and Kaiser, 1999). In order to consider local morphodynamic influences on the sea state, small-scale and high-resolution analy-

ses are necessary. Until now, no quantification of this damping effect, especially for the specific impact on swell waves, has yet been done. In particular, the influence of specific small-scale morphological changes or rather gaps in the ETD sandbanks on the sea state in this area is of great interest for coastal protection, as the ETD sandbanks are constantly exposed to morphological changes.
Thus, the main objective of this study was to quantify the influence of the ETD sandbank dynamics on the sea state and to analyze their damping effect by validated SWAN simulations for representative storm events at tidal high water (THW) and for swell in specific.
In Section 2.2 the research area and dataset are presented. In Section 2.3 the ETD sandbank dynamics (migration velocities and changes in size and shape) are analyzed with GIS based on five different bathymetries from 1995, 2005, 2006, 2008, and 2015. Section 2.4 presents the methods for data analysis and numerical modeling with SWAN. Based on the previous analyzes, Section 2.5 focuses on the modeling results of the nearshore sea state of Norderney with dynamic bathymetries in SWAN to isolate and quantify the ETD influences on the sea state damping. Therefore, five representative storm cases are simulated and the respective damping effects are discussed. The Norderney nearshore wave climate is also characterized in this chapter to distinguish natural variability from morphological influences. Further evaluations transfer some general findings of the modeled sea state damping effect to the buoy measurements, unaffected and affected by the morphodynamic of the ETD sandbanks off the Norderney coast. Section 2.6 then provides appropriate conclusions.

2.2 Research area and data

2.2.1 Research area

The research area covers the nearshore zone of the coastal area of the East Frisian island of Norderney in the northern German North Sea. It extends from about 7.05°–7.3° E and 53.85°–53.65° N. Norderney is separated from the island of Juist in the west by the Norderneyer Seegat and from Baltrum in the east by the Wichter Ee (Fig. 2.1). The main wind direction of the research area is southwest. The long-term Mean High Water (MHW) is 1.24 m above the German reference height (NHN), which is equal to the German mean sea level.

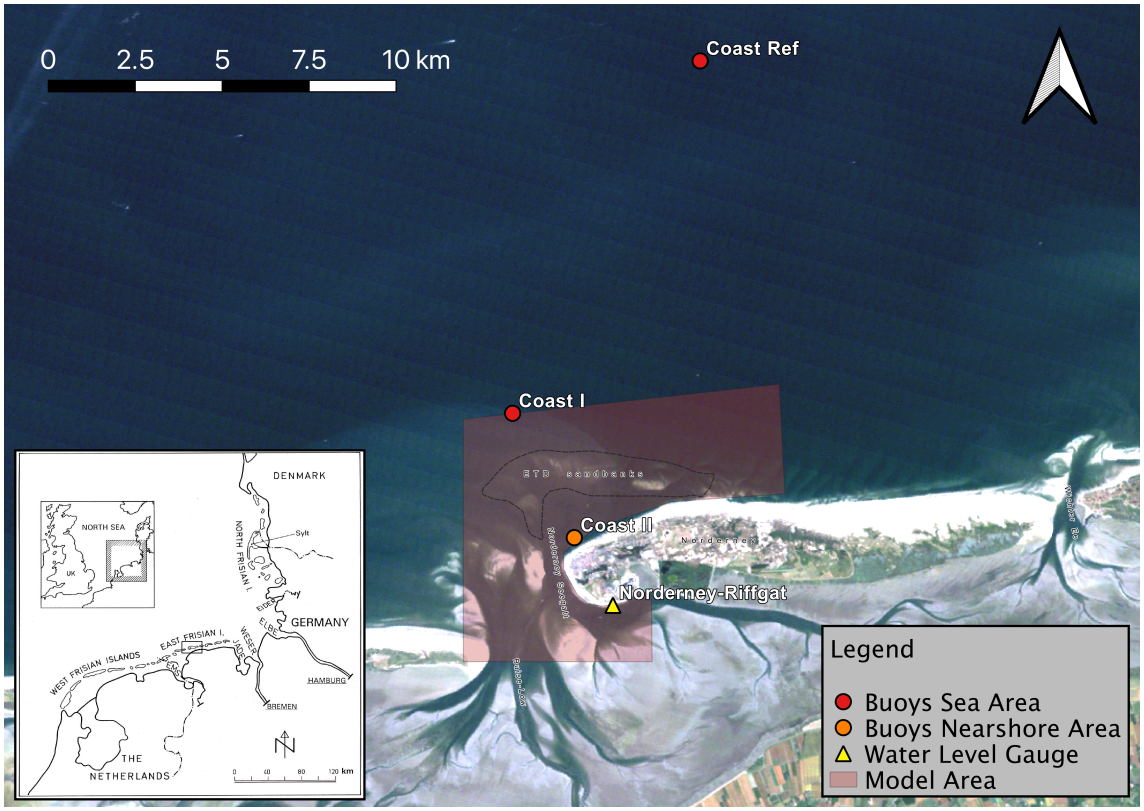

Figure 2.1: *Map of the research area and the positions of the buoys. The buoys are clustered according to their absolute variability. The SWAN model area (red) and the extend of the relevant ETD sandbanks (dashed line) are shown on the map.*

2.2.2 Data

Measurement data of three buoys located in the research area were provided by the Coastal Research Station of the Lower Saxony Water Management, Coastal Defence and Nature Conservation Agency (NLWKN), Norderney. Two buoys (Coast Ref & Coast I) are located north of the ETD sandbanks – the reef arch west and north of Norderney – and one buoy (Coast II) is located south of the ETD sandbanks and north of the Norderney coastline (Fig. 2.1). Due to their position north and south of the ETD sandbanks and the amount of data collected (1992–2017), the two buoys Coast I and Coast II were selected for the analysis of the sea state damping effect (Fig. 2.1). All buoys are of the type Datawell Waverider, both directional and non-directional, which measure heave, north (directional only) and west (directional only) acceleration with a frequency of 1.28 Hz. Wind and water level data were given by the German Meteorological Service (Deutscher Wetterdienst, DWD) for the station Norderney and the Waterways and Shipping Administration (WSA) for the measurement level Norderney-Riffgat, respectively (Fig. 2.1). The mean tidal water level at gauge Norderney-Riffgat was about 0.04 m NHN in the observed years.

For the analysis of the ETD sandbank dynamics digital terrain models (DTMs) and aerial photos of the ETD sandbanks were also provided by the NLWKN Coastal Research Station. The DTMs included the years 1995, 2005, 2006, 2008, and 2015, while the aerial photos were available for 1995 and 2005 to 2008. The DTMs were also used as bathymetries for SWAN modeling.

2.3 Morphodynamic analysis of the ETD sandbanks by GIS

2.3.1 ETD dynamics analysis methods

In a first step an analysis of the ETD sandbank movements and dynamics was performed with QGIS using the DTMs and aerial photos (QGIS Development Team, 2019). A detailed description of the meta data of the DTMs and the aerial photos is given in Tab. 2.1. The five different bathymetries from the DTMs were selected for the highest degree of variability in their characteristics of the sandbanks size and shape. Differences between the aerial photos and the DTM analysis can be explained by the varying dates of survey (Tab. 2.1). Different water levels of the aerial photographs caused a coarse approximation of plates below sea level.
The hydrographic surveying for the DTMs was performed with a sampling rate of 3–10 kHz. For a more accurate measurement, the measurement profiles should be perpendicular to the depth contours. However, this cannot always be guaranteed due to the partly round shape of the ETD sandbanks (Bremermann and Meyer, 2012). The DTMs were calculated in QGIS by triangulation via a Triangulated Irregular Network (TIN) on an even grid with a grid point distance of 5 m. If necessary, corresponding data points were recalculated after error correction (Bremermann and Meyer, 2012). Because the Mean Low Water (MLW) is about -1.21 m NHN and the Lowest Low Water (LLW) is about -2.48 m NHN, the ETD sandbanks were defined as all areas above -3.5 m NHN, to ensure that all areas were taken into account. Bremermann and Meyer (2012) used this value as the lower boundary for the definition of sandbanks in order to include both single plates as well as extended plate structures. To consider differences in sea level, the bathymetries have been clipped to a reference area defined by the extend of the ETD sandbanks over all years (Fig. 2.1) and were MLW-corrected by subtracting the annual MLW of the respective years. The distribution of the water depths at MLW provided indications of how wave breaking was influenced and the plates developed. To validate the results of the DTMs, which include measurements of several months, the shape of the ETD sandbanks was visually detected and calculated by image segmentation with k-means algorithm (see Section 2.4.2) off the aerial photo snapshots (Dhanachandra et al., 2015).

Table 2.1: *Information about the survey of the DTMs and aerial photos.*

		hydrographic surveying	
year	institution	survey time	point distance [m]
1995	BSH	23/05–13/09	200 × 200 (100 × 100)
2005	NLWKN	02/03–22/11	10 × 150
2006	BSH	–	75 × 100
2008	NLWKN	–	2 × 100
2015	BSH	–	5 × 5
		aerial photos	
year	institution	scale	recording time MEZ	water level [m NHN]
1995	WBG	1:28 000	10/07 15:15	-1.52
2005	Hansa	1:28 000	22/09 08:38	-1.66
2006	WBG	1:28 000	01/07 09:27	-1.35
2007	WBG	1:28 000	03/10	–
2008	WBG	1:28 000	31/07 16:57	-1.66

2.3.2 ETD dynamics analysis results and discussion

The results of the long-term variability of the ETD sandbank dynamics were evaluated by analyzing their movement, size and shape based on the five topographies as given below.

To obtain a general overview of the movement and size, Fig. 2.2 shows the different height distributions of the sandbanks against the MLW-corrected water depths of the relevant years. The histogram qualitatively describes the ETD sandbanks displacements, since differences between the years are significant. In 1995 (blue) the distribution was skewed to the left or lower water depth, while being skewed to the right or higher water depth since 2005 (orange). In the interval of 1 to -3 m, fewest values were found in 1995 (Fig. 2.2).

Following, we analyzed the specific movement velocities and patterns of the plates Nordwestergründe (NWG I/NWG II) and Nordergründe (NG) (Tab. 2.2, Fig. 2.3). In general, the ETD sandbanks move from the eastern end of Juist in a curve to the north, east, and south-east and accumulate at the northern beach of Norderney (Fig. 2.3). The mean velocity of the plates moving east is about 435 ± 180 my^{-1} for NWG I/NWG II and 491 ± 388 my^{-1} for NG with a high standard deviation of annual velocities that will not allow for a consistent periodic pattern (Tab. 2.2). Bremermann and Meyer (2012) indicated a displacement velocity varying between 250 and 430 my^{-1} depending on the current position of the plates, whereby a period of approximately 37 years was calculated for a complete cycle of a plate. Other

studies quantified the average displacement velocity at 400 my^{-1} (Fitzgerald et al., 1984) and at 325 my^{-1} in the west and 500–550 my^{-1} in the east (Gaye and Walther, 1935).

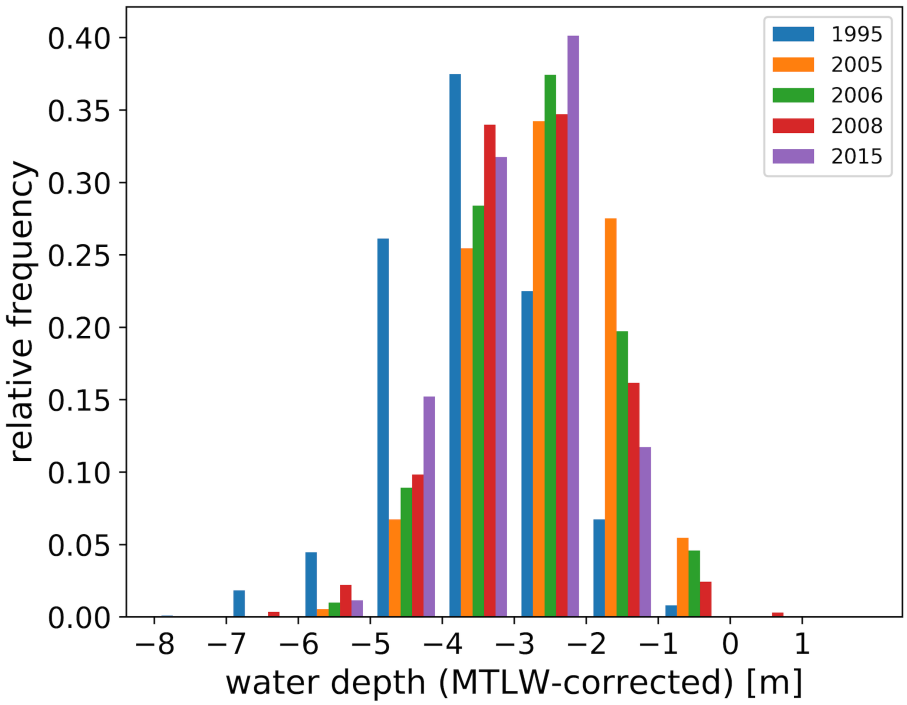

Figure 2.2: *Distribution of the ETD-clipped and MLW-corrected water depths for the years 1995, 2005, 2006, 2008, and 2015.*

The sizes of the ETD sandbanks vary over time with a mean area of 55 ± 24 ha or 126 ± 83 ha for NWG or NG, respectively. This could be caused by non-periodic, extreme storm surges, which can lead to abrupt changes of the ETD sandbanks shape and size and rather periodic influences such as currents and tidal inlets (Dallas and Barnard, 2011; Wang et al., 2012).

The qualitative analysis of the plates shape was carried out for individual years. In 1995, the shape of the ETD sandbanks was narrow between Coast I and Coast II (three gaps were visible), while they were most developed in 2005 (Fig. 2.3). In 2005, NWG II and NG merged. From 1995 to 2005, the small plate in the west (northwest of Coast II) has moved south. In 2005, the entire area between Coast I and Coast II was filled with plates. In 2006 there was a small gap between NWG I/NWG II and NG, between 320° and 332°. NWG I had moved north in 2006, getting closer to NWG II and causing a great gap south of NWG I. The Buise-Low has extended to the north in 2006 (Fig. 2.3). In 2008, the northern gap has moved between 345° and 10°, while NWG I and NWG II merged during this year. Thus, the gap has become

13° larger. South of the new NWG plate a huge gap still existed. From 2005 to 2008 the development of an accumulating plate at the east of the Norderney beach was clearly visible (Fig. 2.3). This small plate connected to the beach of Norderney in 2008. Areas falling dry periodically were identified in 2008 (Fig. 2.3). In 2015, the small gap has moved to the west with a range of 325°–345°. It was 5° smaller than the year before, but a new huge gap was developing from the east (Fig. 2.3).
Furthermore, the ETD sandbanks did not become higher with sea level rise, since a significant correlation of the minimum, maximum, mean, and standard deviation of the uncorrected depths and the MHW and MLW of the respective years was not found (Tab. 2.3). However, potential growth effects may be superimposed by the natural dynamics of the ETD sandbanks.
The dynamics of the ETD sandbanks described above provide a basis for further analyses of its influences on the sea state. Respective methods are explained in the next chapter. In summary, the ETD sandbanks were small in 1995, while being most developed in 2005. In 2006, 2008, and 2015 small gaps between plates were visible (Fig. 2.3).

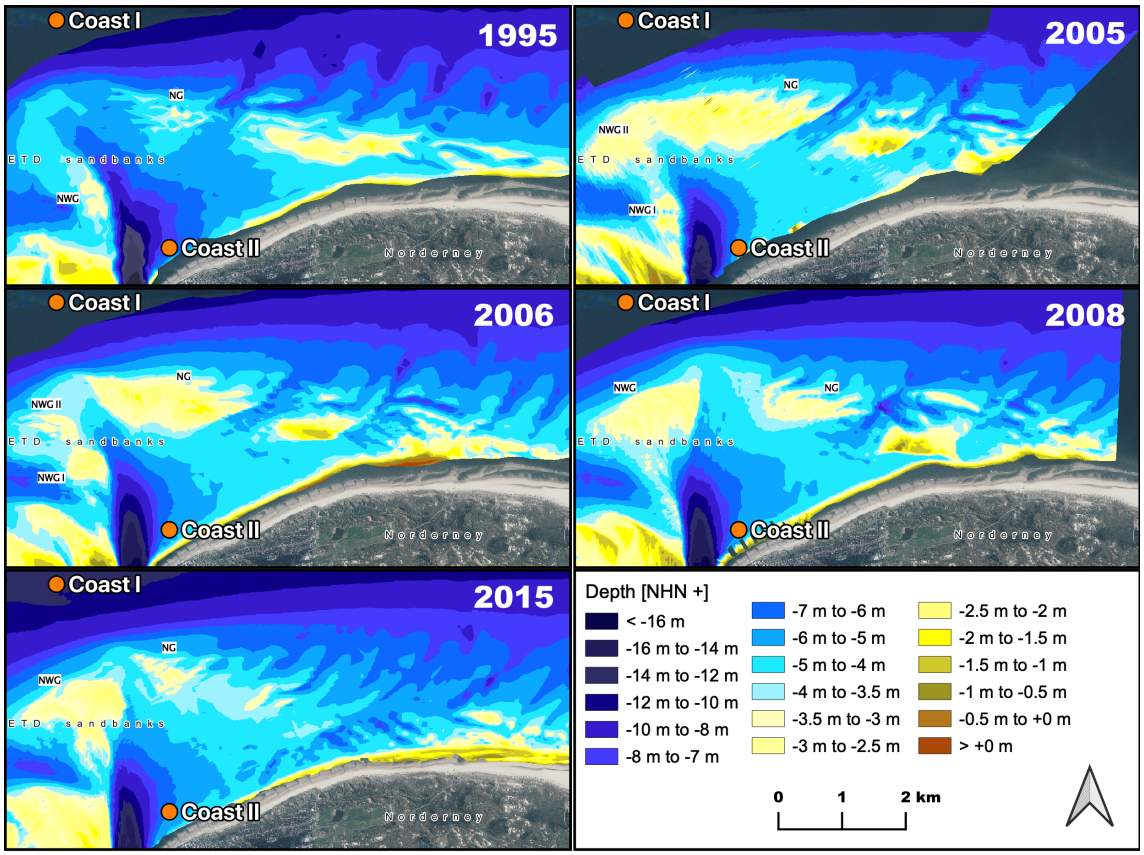

Figure 2.3: *Bathymetries of the research area for the years 1995, 2005, 2006, 2008, and 2015. The bathymetries were used for SWAN simulations.*

Table 2.2: *ETD sandbank displacement velocities and sizes of the (a) DTMs (bathymetry) and (b) aerial photos.*

year	**movement velocity & direction [my^{-1}]**			**area [ha]**		
	NWG I	**NWG II**	**NG**	**NWG I**	**NWG II**	**NG**
2005/06[a]	*663 (N)*	*354 (NE)*	*256 (SW)*	*33/19*		*258/85*
2005/06[b]	*559 (N)*	*191 (N)*	*278 (E)*	*24/29*	*58/23*	*168/164*
2006/07[b]	*626 (NE)*		*1 160 (E)*	*29,23/51*		*164/147*
2007/08[b]	*315 (NE)*		*258 (SW)*	*51/72*		*147/30*
2006/08[a]	*335 (NE)*		*503 (E)*	*19/79*		*85/30*
mean	*435*		*491*	*55*		*126*
std	*180*		*388*	*24*		*83*
min	*191*		*256*	*19*		*30*
max	*663*		*1 160*	*82*		*258*
n	*7*		*5*	*7*		*7*

Table 2.3: *Correlation between mean water level (MHW and MLW) and ETD sandbank heights (uncorrected depth) for data of the years 1995, 2005, 2006, 2008, and 2015.*

Water Level	ETD min	ETD max	ETD mean	ETD sd
MHW	*-0.74*	*0.11*	*-0.44*	*-0.03*
MLW	*-0.25*	*0.64*	*0.21*	*0.29*

2.4 Analysis methods of ETD sea state damping

2.4.1 Buoy data

2.4.1.1 Wave parameter postprocessing

The buoys postprocessing software W@ves21 provided wave parameters from a spectral analysis at 30-min time periods (history data files) (Datawell, 2014). Since all data should be processed consistently and the buoys software had changed over the years, the raw (directional) surface elevation data of the 30-min periods were used in addition to the history data files. Based on the buoys raw data, a data preprocessing was performed with the MATLAB toolbox WAFO to calculate the continuous variance density spectrum for a range of 0–0.65 Hz by a Fast Fourier Transformation (FFT) of the surface elevation values of the 30-min periods (WAFO, 2011). Various outlier detection tests were applied to remove and reconstruct spurious surface elevations (Holthuijsen, 2007; WAFO, 2011).
The statistical wave parameters significant wave height H_{m_0}, energy period $T_{m_{-1,0}}$, peak period T_P, and the wave direction Dir_P were determined half-hourly from the spectral variance density. They are defined as

$$H_{m_0} = 4\sqrt{m_0}, \quad T_{m_{-1,0}} = \frac{m_{-1}}{m_0}, \quad T_P = argmax(S(f))^{-1} \tag{2.1}$$

where m_n is the n-th moment of the variance density spectrum defined as

$$m_n = \int f^n S(f)\, df \tag{2.2}$$

where $S(f)$ is the variance density spectrum of frequency f. The wave direction Dir_P is defined as the wave direction at the variance density spectral peak (Bacon and Carter, 1991; Holthuijsen, 2007; Sheng and Li, 2017; SWAN, 2019b; Young, 1999).
The data preprocessing and the calculated sea state parameters were validated with the history data files (Fig. 2.4). As a result, the parameters calculated in Eqs. 1 and 2 were used in the following. Some measurement periods had to be excluded

from further analysis due to buoy failure and incorrect measurements caused by storms or maintenance. A qualitative verification of the measured data, followed by an inspection of extreme storm events, and wind data was conducted (Axer et al., 2005-2008). Only extreme wave events with a corresponding storm event were kept in the final data set.

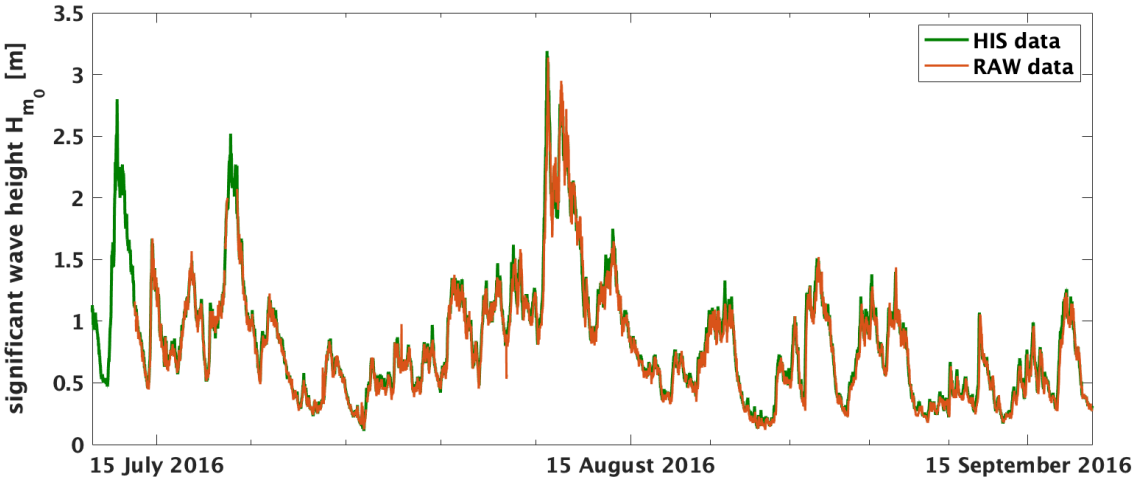

Figure 2.4: *Comparison of buoy history data (green) and the processed raw data (red) for the parameter H_{m_0} for three month of the year 2016. The parameters calculated from the variance density spectrum of the raw data were used in this study.*

2.4.1.2 Spectral separation

To identify swell and wind sea, a partitioning and separation of the variance density spectra was performed. The applied partitioning algorithm is based on Hanson and Phillips (2001) which is an implementation of the concept described by Hasselmann et al. (1996). The idea is to consider the wave spectrum as mountains and to define different regions by matching every spectral grid point to a local peak with highest local spectral energy. This is also known as the watershed algorithm. From the eight surrounding neighbor grid points, the one with the direction of the steepest ascent was iteratively chosen. All grid points assigned to the same local maxima were clustered into a partition of the wave spectrum. Individual partitions were not merged to swell or wind sea systems as described in Hasselmann et al. (1996) and Portilla et al. (2009), because the total swell or rather wind sea energy was of interest. To remove fine scale noise, a H-maxima transformation was applied to suppress all maxima of the wave spectrum whose height was less than $h = 0.028$ (Hanson and Phillips, 2001; Hanson et al., 2009; Portilla et al., 2009).
After partitioning, two different methods for spectral separation, namely wave age criterion and separation frequency, were compared. The wave age criterion uses two-

dimensional spectral information and depends on wind speed U $[ms^{-1}]$ and direction U_θ [°]. Since it is difficult to determine the phase speed c of the wind sea in shallow waters for each wave event depending on the wave period T, the phase speed for deep water was applied. Wind sea was defined after the wave age criterion by

$$f \geq \frac{g}{2\pi} \left(1.5\ U \cdot cos(\delta)\right)^{-1} : \quad 0 \leq \delta \leq \frac{\pi}{2} \tag{2.3}$$

where $f = T^{-1}$ is the frequency of the waves, $g = 9.81$ the gravity constant and δ the angle between U_θ and the measured wave direction Dir_P. If the peak of a partition falls within the described parabolic area of Eq. 2.3, the entire partition was defined as wind sea, otherwise as swell (App. I.1). Factor 1.5 was used in Hanson and Phillips (2001) to ensure that all wind sea peaks are obtained.
The second method of the separation frequency f_S uses the wind speed U only and, therefore, does not depend on the directional information of the buoy to determine a frequency which defines the swell at lower frequencies and the wind sea at higher frequencies (App. I.1). For this method the following applies

$$f_S = 0.104\ \frac{g}{U} \tag{2.4}$$

based on the peak frequency of the Pierson-Moskowitz spectrum (PM-peak method) (Pierson and Moskowitz, 1964; Portilla et al., 2009). To ensure that all wind sea peaks were included, the separation frequency in Eq. 2.4 was applied. It is shifted to lower frequencies compared to other methods, which, e.g., use the relationship of the peak frequency of the steepness function and the wind speed (Portilla et al., 2009; Wang and Hwang, 2001).
Compared to the wave age criterion, the separation frequency provided better results for events with low wind speeds and a high difference between wind and mean wave direction, and was therefore used in this study.

2.4.2 Statistical methods of natural sea state variability

While the analysis of the ETD sandbank dynamics was presented in the previous chapter, the statistical methods to describe the natural variability of the sea state are described below. To get a first overview of the data and their temporal variability, two time-series were plotted and a linear regression was performed on buoys Coast I and Coast II to find a linear trend in the wave height data. Afterwards, the general natural variability of the sea state was described by means and standard deviations of the buoy data. To quantify the intensity of possible non-linear influ-

ences on wave propagation in this area, the linear Pearson correlation coefficient r between the wave heights of the different buoys was determined (Cowpertwait and Metcalfe, 2009; Ruggiero et al., 2010).
In addition to the time series analysis, a k-means clustering of the data was carried out, in order to identify groups of buoy positions with equal temporal variability (Jain, 2010). Therefore, the clustering was performed according to the H_{m_0} standard deviation of each buoy. The k-means clustering algorithm is an unsupervised method to group a dataset into a fix number of $k \in \mathbb{N}$ clusters. The centroid and objects of each cluster are determined iteratively by assigning objects to clusters by their nearest centroid and calculating new cluster centroids, whereby the initial centroids are randomly selected. Iteration stops, when the sum of squared errors (SSE) between the mean of every cluster centroid and its assigned objects is minimized. In this study the Euclidean metric was used to calculate the distances (Herrero et al., 2015; Jain, 2010; MacQueen, 1967). To find the optimal k, the sum of squared distances within a cluster (WSS) and the sum of squared distances between the clusters (BSS) were calculated and minimized and maximized for $k = 1,\ 2,\ 3$, respectively (Hartigan and Wong, 1979). The number of clusters at the elbow point of the respective quality indicators was selected to ensure both good interpretability as well as separation of the clusters (Hastie et al., 2009).
After determining spatial groups of similar wave height variability, we looked for seasonality in the clustered data by periodicity analysis. Therefore, a periodogram was calculated by FFT of the half-hourly H_{m_0} data (Venables and Ripley, 2002). The periodogram shows the (smoothed) spectral density of periodic components with a certain period of a time series. Thus, harmonic periodic patterns and their amplitudes in the data were used to determine the main seasonal components (Bloomfield, 2000; Shumway and Stoffer, 2006; Vanem and Walker, 2013). Clearly identifiable peaks in the periodogram indicate specific periodic signals in the sea state data. The original data can be written as the sum of sinusoid functions determined by the FFT. A period of e.g. 1 h in the periodogram refers to a sinus curve with the respective period and an amplitude given by the spectral density.

2.4.3 Numerical modeling

2.4.3.1 SWAN wave model

Nearshore sea state modeling was performed with the third-generation wave model SWAN with an unstructured grid. SWAN is made for simulations of random, short-

crested waves in coastal regions (Booij et al., 1999). It calculates the spreading of the sea state by solving the spectral action balance equation

$$\frac{\partial}{\partial t}N + \frac{\partial}{\partial x}c_x N + \frac{\partial}{\partial y}c_y N + \frac{\partial}{\partial f}c_f N + \frac{\partial}{\partial \theta}c_\theta N = \frac{S}{f} \tag{2.5}$$

where $N = N(f, \theta)$ is the action density spectrum with frequency f and wave direction θ. The action density spectrum is equal to the variance density spectrum divided by the frequency f. t denotes the time, x and y represent the geographical propagation with velocities c_x and c_y, respectively. The fourth term in Eq. 2.5 represents changes in depth and currents of the computational grid. The fifth term of Eq. 2.5 represents directional changes caused by refraction (Booij et al., 1999; SWAN, 2019a). The unstructured computational grid must have 4–10 cells around every node and the size of the inner angles must not exceed 143° to run properly in SWAN (SWAN, 2019b). Unstructured grids can be fitted to the topography and simultaneously reduce the computational costs.

2.4.3.2 Model settings

The unstructured grid was adjusted to the average ETD sandbanks shape and was therefore applicable for all bathymetries, while the number of nodes was kept low at 32 055. The maximal spatial resolution of the unstructured grid was about 16 m around the ETD sandbanks, while the resolution had its minimum of about 154 m at the model boundaries (Fig. 2.5). The unstructured grid was created with the Surface-water modeling system (SMS) and the paving method (Surface-Water Modeling System, 2012).

All computations with SWAN were performed in stationary mode of SWAN model version 41.20AB. The wind input and whitecapping generation was chosen according to van der Westhuysen (2007). Triad wave-wave interactions of Eldeberky (1996) were used with proportionality coefficient $\alpha_{EB} = 0.1$ and a maximum frequency $f_{max_{EB}} = 2.5\ f_{m01}$ with f_{m01} mean frequency, as proposed in van der Westhuysen (2012). Quadruplet wave-wave interactions were modeled using the Discrete Interaction Approximation (DIA) according to Hasselmann et al. (1985). Bottom friction was represented by Hasselmann et al. (1973) with a friction coefficient of 0.038 $m^2 s^{-3}$ as proposed by van der Westhuysen (2012). For depth-induced wave breaking the $\beta - kd$ parameterization of Salmon et al. (2015) was applied, which shows comparable results to the ϕ parametrization of van der Westhuysen (2009, 2010) and is recommended for complex coastal environments with horizontal bathymetries and local wave generation like the nearshore area of Norderney (Salmon and Holthui-

jsen, 2015). The proportionality coefficient α for the rate of dissipation of the $\beta - kd$ parametrization was tested during calibration of the respective events between 0.05 and 1.00 (see 2.4.3.4).

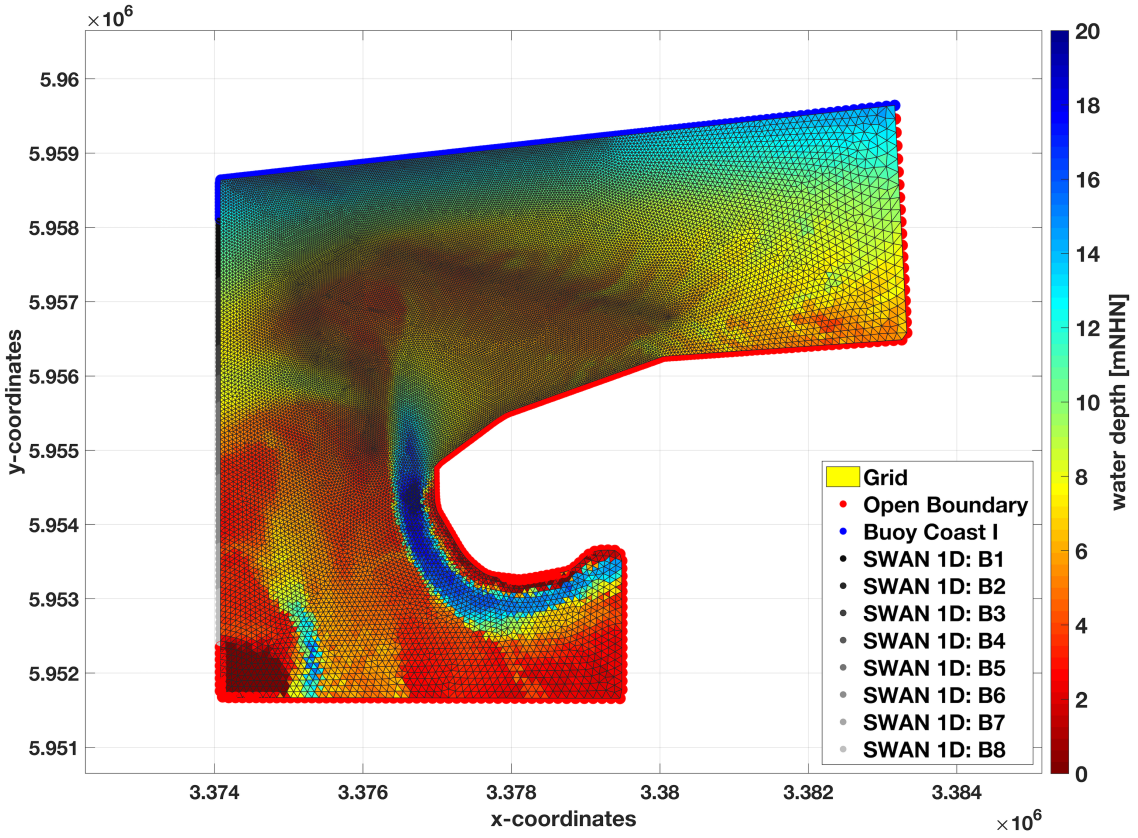

Figure 2.5: *SWAN model boundary conditions and grid properties. The northern boundary is equal to the sea state signal of Coast I. The western boundaries (B1 to B8) are modeled by a one-dimensional SWAN simulation with the respective bottom profile and incoming wave signal of Coast I. The southern and eastern boundaries are considered as open boundaries. In the background the topography of 2015 is shown.*

A frequency resolution of 73 frequencies between 0.025 Hz and 1.0 Hz and a directional resolution of 2.5° over the whole circle were applied. To ensure full convergence of the SWAN computations, the curvature-based criterion of Zijlema and van der Westhuysen (2005) was used. Accordingly, computations were stopped if the curvature of the normalized H_{m_0} iteration curve was less than 0.005 and either the absolute or relative change in H_{m_0} changed by less than 0.005 m or 0.01 respectively, for over 98.5% of the computational wet grid points (Salmon and Holthuijsen, 2015). A maximum of 80 iterations was set.

2.4.3.3 Boundary conditions

Tab. 2.4 presents the selection of measured cases for the boundary conditions of SWAN modeling. The buoy-measured signals of Tab. 2.4 at position Coast I were used as the northern boundary for the respective cases (Fig. 2.5). The western boundary was divided into eight different sections (Fig. 2.5), which were each modeled by a one-dimensional simulation with SWANOne with the same hydrodynamical and bathymetric conditions as the main simulation (Verhagen et al., 2008). The southern and eastern boundaries were defined as open boundaries.
For each simulation, the wind input and water level were assumed to be constant according to the measurements presented in Tab. 2.4 (Witting et al., 2014). The calculated 10 m wind speed at sea was used. The bathymetries analyzed in Section 2.3 were used for the respective modeling cases. Currents were not considered in this study, since only simulations and observations at THW were taken into account. In the coastal area, the direction of the dominant tidal currents change near peak water levels, thus the current velocities at these times, i.e., to THW and TLW, can be neglected (Kaiser and Niemeyer, 1999; Koch and Niemeyer, 1978; Mai et al., 2004). All selected cases are representative storm events from different years (different bathymetries) with different wind speeds and wind directions at the respective THW time, because coastal protection is focused on such events. The quantification of the ETD damping effect is of great interest especially during such events.
With 12 cm, the variation of the observed MHW at Norderney-Riffgat in the computed years 1995, 2005, 2006, 2008, and 2015 was significantly higher than the theoretical mean sea level rise of about 28 cm per century so that sea level rise was not considered in the simulations and an area-constant water level was used for each modeled case.

2.4.3.4 Model calibration

The nine selected cases in Tab. 2.4 were divided into four calibration and five validation cases. Calibration was performed to $f_{max_{EB}}$ of the triad wave-wave interactions, bottom friction coefficient, and the proportionality coefficient α of the $\beta - kd$ parametrization. For $f_{max_{EB}} \in [2.00, 5.00]$ f_{m01} best results in terms of absolute percentage error of the wave parameters and qualitative analysis of the spectra were achieved for $f_{max_{EB}} = 2.5$ f_{m01} as initially set according to van der Westhuysen (2012) (see 2.4.3.2). The bottom-friction coefficient was tested for the interval $[0.01, 0.1]$ with the best results for the initial parameter 0.038 $m^2 s^{-3}$ by (van der Westhuysen, 2012) (see 2.4.3.2). Fig. 2.6 shows the results of the triad wave-wave interactions and bottom friction calibrated model wave spectra at position Coast II.

Table 2.4: *Selection of Buoy Measurement Cases of Coast I and Coast II for Calibration and Validation (Test Cases) for the boundary conditions of SWAN modeling.*

Case Name	Date and Time (UTC+1)	U_{See} $[ms^{-1}]$	U_θ [°]	Water Level [m NHN]	Tidal Phase	H_{m_0} Coast I [m]	$T_{m_{-1,0}}$ Coast I [s]	*Dir* Coast I [°]
Calibration #1	15-Nov-2005 10:00	*10.2*	*320*	*1.99*	*THW*	*2.00*	*6.02*	*318*
Calibration #2	16-Dec-2005 10:00	*18.5*	*343*	*2.18*	*THW*	*4.82*	*8.28*	*333*
Calibration #3	07-Nov-2015 21:30	*16.0*	*247*	*1.64*	*THW*	*1.92*	*4.91*	*353*
Calibration #4	11-Jan-2015 14:30	*21.4*	*290*	*2.45*	*THW*	*5.37*	*8.91*	*324*
Validation/Test Case #1	25-Jan-2015 02:30	*7.9*	*310*	*1.91*	*THW*	*2.08*	*6.29*	*308*
Validation/Test Case #2	02-Feb-2015 10:00	*9.1*	*313*	*1.47*	*THW*	*2.30*	*6.09*	*303*
Validation/Test Case #3	05-Sep-2015 16:00	*17.9*	*307*	*1.97*	*THW*	*3.91*	*7.39*	*324*
Validation/Test Case #4	16-Dec-2005 23:00	*18.8*	*330*	*2.62*	*THW*	*4.42*	*8.19*	*333*
Validation/Test Case #5	20-Jan-2005 06:00	*18.3*	*277*	*1.70*	*THW*	*3.98*	*7.76*	*324*

The $\beta - kd$ proportionality coefficient α was adjusted individually to the different wind speeds and directions in the range of 0.05–0.25 for both the calibration and validation data, which provided the best results (Fig. 2.6). Comparing the measured and modeled wave parameters after calibration, good results were obtained, since mean absolute percentage errors (MAPE) of 2.07% for H_{m_0}, 9.19% for $T_{m_{0,1}}$, 6.00% for $T_{m_{-1,0}}$, and 8.25% for T_P were achieved for the calibration data. While case A underestimated the spectral energy density peak, it was slightly overestimated in case B. In case C, the wind sea was underestimated and the swell overestimated, which could be caused by the western wind direction. In total, good correspondence of the wave parameters is shown. The calibration to the northwesterly storm event D showed a good fit to the wave parameters, but overestimated the spectral peak and did not fit the wind sea properly. SWAN also showed an overall satisfactory modeling accuracy of the mean wave direction (Fig. 2.6).

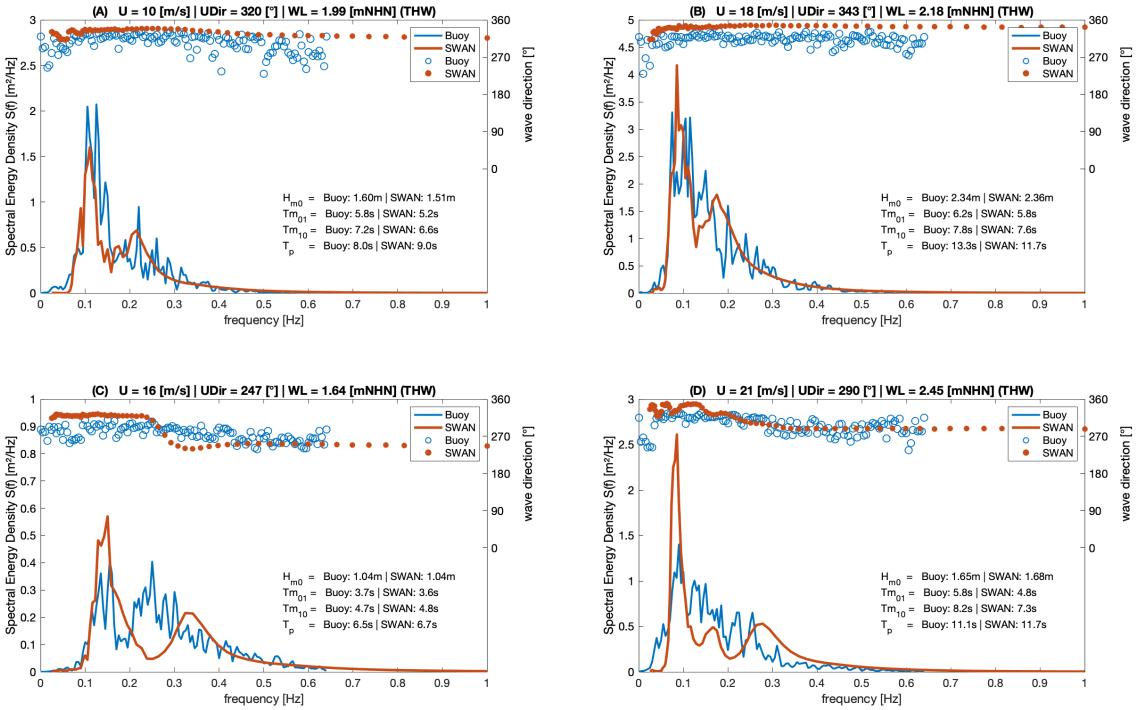

Figure 2.6: *Variance density spectra of the calibration cases. The observed (blue) and modeled (red) spectra and directions as well as the wave parameters are shown for A: 15-Nov-2005 10:00 with $\beta - kd$ proportionality coefficient $\alpha = 0.05$, B: 16-Dec-2005 10:00 with $\alpha = 0.05$, C: 07-Nov-2015 21:30 with $\alpha = 0.05$, and D: 11-Jan-2015 14:30 with $\alpha = 0.25$.*

2.4.4 Model validation and transfer to test cases

After calibration, the model settings were transferred to the five validation and test cases of Tab. 2.4 as shown in Fig. 2.7. The validation reached good results with MAPE of 1.47% for H_{m_0}, 10.80% for $T_{m_{0,1}}$, 7.24% for $T_{m_{-1,0}}$, and 18.76% for T_P. An overall satisfactory accuracy of the mean wave direction was achieved (Fig. 2.7).

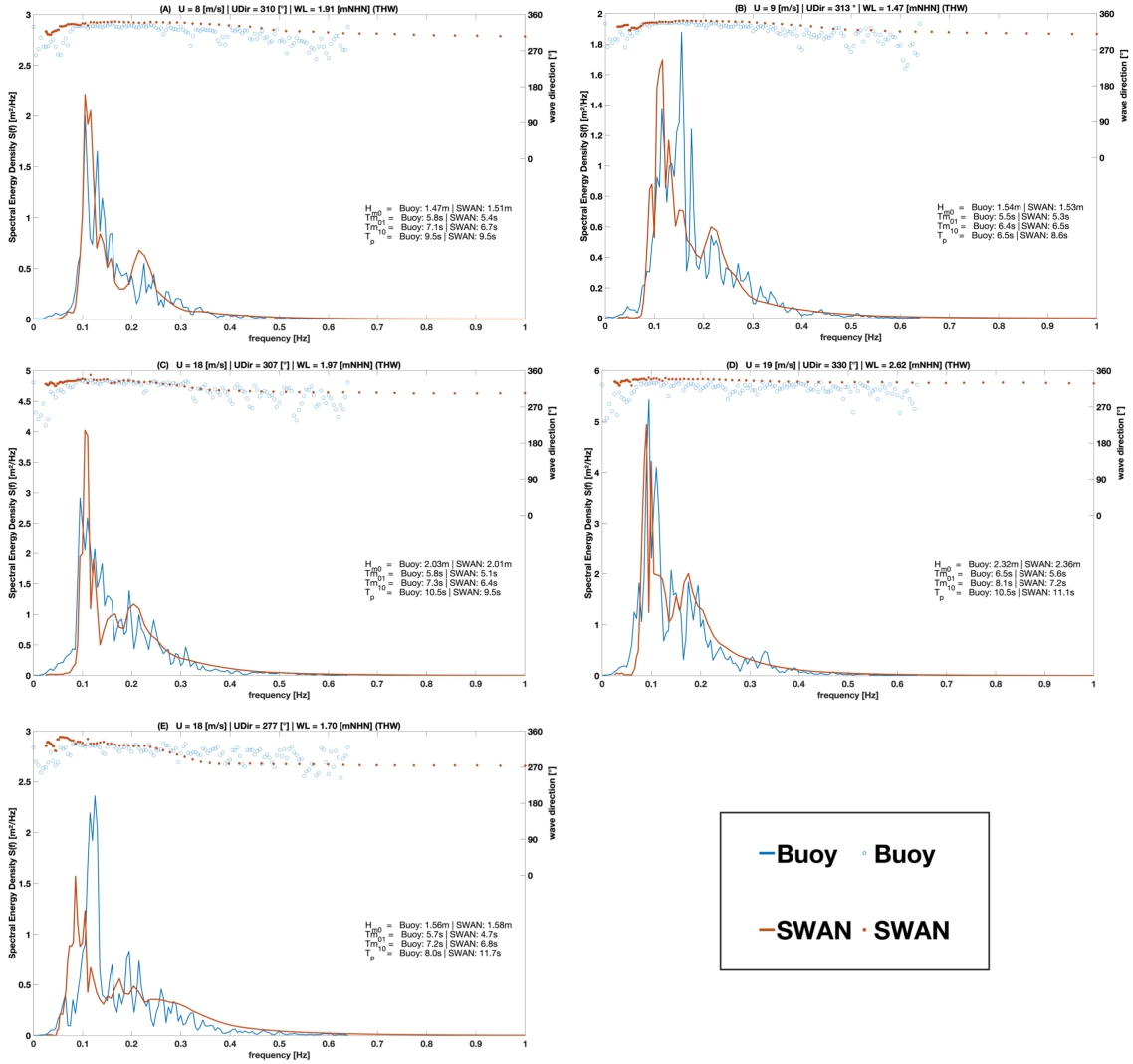

Figure 2.7: *Variance density spectra of the validation and test cases. The observed (blue) and modeled (red) spectra and directions as well as the wave parameters are shown for A: 25-Jan-2015 02:30, B: 02-Feb-2015 10:00, C: 05-Sep-2015 16:00, D: 16-Dec-2005 23:00, and E: 20-Jan-2005 06:00.*

Therefore, the validated cases were suitable to model the isolated changes of the ETD sandbanks on the sea state damping effect and were also used as the test cases of this study with the above described boundary conditions (Tab. 2.4). The boundary conditions were selected for different wind speeds and wind directions as well

as different related bathymetries, since the aim was to focus on representative sea states during storm events at THW. Because only the bathymetries differed in the simulations of each test case, the model results were able to describe the isolated ETD sandbanks influence on the wave height by the sea state damping effect as defined in Eq. 2.6.
For each of the five test cases an ensemble of three simulations with five different bathymetries each were performed. Differences within the three respective computational runs were neglectable, indicating a good numerical convergence.

2.4.5 Sea state damping effect

According to Niemeyer (1979), we define the relative and absolute sea state damping effect of the ETD sandbanks as

$$\Delta_{rel}\ H_{m_0} = 1 - \frac{H^{II}_{m_0}}{H^{I}_{m_0}} \quad and \quad \Delta_{abs}\ H_{m_0} = H^{I}_{m_0} - H^{II}_{m_0} \tag{2.6}$$

where $H^{I}_{m_0}$ and $H^{II}_{m_0}$ are the significant wave heights H_{m_0} measured at position Coast I and Coast II, respectively.
To compare buoy-measured damping effects with SWAN simulations, a filtering of the in-situ data according to the model boundary conditions is mandatory before calculating the damping effects as defined in Eq. 2.6. Simulations were applied to expand measurement data, thus relevant events were only sparsely covered by the buoy data. Since many factors influence the buoy-measured damping effects and SWAN represents only one specific condition, the sea state has to be filtered according to: (A) significant wave height H_{m_0} and mean wave direction, (B) wind direction and wind speed, (C) water level and tidal phase, divided into THW and TLW, whereby a buffer of 1.5 h was placed around the events and the currents at THW and TLW tend to zero, (D) swell and wind sea, as swell behaves differently than wind sea when crossing the ETD sandbanks, and (E) summer and winter period, defined as the period from April to September and from October to March, respectively (Holthuijsen, 2007; Young, 1999).

2.5 Results and discussion of ETD sea state damping

2.5.1 Natural wave climate variability

For a first overview, detailed descriptive statistics of the different buoy positions are given in Tab. 2.5. With increasing coastal proximity, the mean significant

wave height H_{m_0} decreased from 1.20–0.63 m, mainly caused by depth-induced wave breaking. Concurrently, the mean wave period $T_{m_{-1,0}}$ and the quotient of its standard deviation and mean value increased slightly with increasing coastal proximity. This could indicate a variability of different sea states and wave transformation processes.

Table 2.5: *Descriptive statistics of all buoys and buoy clusters of the research area.*

	Coast Ref	**Coast I**	**Coast II**	**Sea Area**	**Nearshore Area**
			H_{m_0} $[m]$		
mean	*1.20*	*1.07*	*0.63*	*1.14*	*0.63*
sd	*0.76*	*0.69*	*0.39*	*0.73*	*0.39*
range	*0.09–8.40*	*0.09–6.91*	*0.04–3.01*	*0.09–8.40*	*0.04–3.01*
			$T_{m_{-1,0}}$ $[s]$		
mean	*5.62*	*5.77*	*6.04*	*5.70*	*6.04*
sd	*1.37*	*1.47*	*1.75*	*1.42*	*1.75*
range	*2.48–24.76*	*2.58–30.87*	*2.29–21.56*	*2.48–30.87*	*2.29–21.56*

Based on the absolute variability (standard deviation) of H_{m_0}, the following two groups were identified by k-means clustering: Sea Area and Nearshore Area (Fig. 2.1). The groups reflect the coastal proximity of the buoys. The absolute variability of H_{m_0} also decreased with shorter distance to the coast, since it was related to the decrease of H_{m_0} (Tab. 2.5). $T_{m_{-1,0}}$ and its variability increased from the Sea Area to the Nearshore Area (Tab. 2.5). While the two buoys of the Sea Area (Coast Ref and Coast I) are unaffected by the ETD sandbanks, buoy Coast II (Nearshore Area) must have been affected by the ETD sandbanks embossing, since the ETD sandbanks are located between buoy Coast I and Coast II (Fig. 2.1). As position Coast II forms its own cluster (Nearshore Area), the particular influences at this position became evident (Tab. 2.5).

In order to identify and analyze small-scale periodical variability and its amplitude, the periodogram of all buoy time series was performed (App. I.2). Since a strong and dominant periodic pattern of the tidal signal was observed for the Nearshore Area (period of 12.5 h) and a weaker signal for the Sea Area, the same two groups were identified as by clustering. Therefore, the Sea Area was relatively less affected by the tides than the Nearshore Area. Periods of about 6.25 h and 4 h showed also small peaks for the Nearshore Area, as they could be artifacts of the tidal peak. The tidal currents (tidal inlets and outlets) cause this effect, as they increase towards the coast and the Wadden Sea. Compared to the North Atlantic wave climate, where

the periodogram is mainly dominated by a one-year period (Vanem and Walker, 2013), the nearshore zone of Norderney shows a more complex structure.
As the correlation of H_{m_0} between the buoys Coast Ref and Coast I was very high ($r = 0.96$), non-linear influences were neglectable. The correlations between Coast I and Coast II and Coast Ref and Coast II were significantly weaker at $r = 0.83$ and $r = 0.82$, respectively. Thus, non-linear influences could also occur between Coast I and Coast II. This suggested that the ETD sandbanks may exert a non-linear impact on the sea state, but the shallower water depth, currents, and interactions with the beach and breakwaters may also have an effect on the correlation between those positions.
In conclusion, to accurately quantify the influence of the ETD sandbanks on the sea state variability and damping effect, the sea state at buoys Coast I and Coast II was suitable. Fig. 2.8 shows the time series of buoys Coast I and Coast II. Since H_{m_0} was much higher at both buoys in the winter month, a clear annual periodicity became evident. This was caused by rougher weather, storms, and storm surges with more frequent wind directions from north and northwest. Due to its position in deeper waters, Coast I showed higher waves than Coast II. In contrast to previous studies from Vanem and Walker (2013), a statistically significant long-term linear trend could not be observed in the two datasets, since the R^2 were 0.0379 and 0.0002 for annual mean values, respectively (Fig. 2.8).

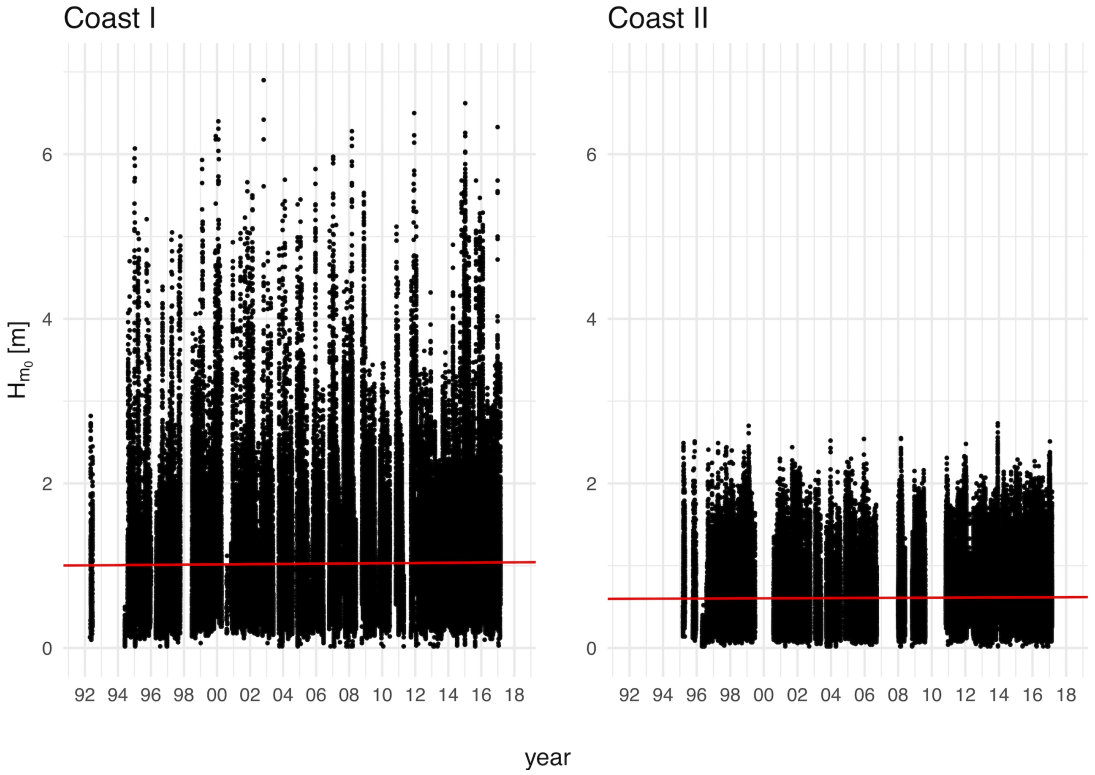

Figure 2.8: *Time series of H_{m_0} from 1992–2017 for the measurement of Coast I and Coast II. The red line indicates a linear regression.*

2.5.2 Sea state damping effect – impact of ETD sandbanks on sea state

In addition to the analysis of the morphodynamic variability of the ETD sandbanks as described in Section 2.3, this section presents the quantification of the corresponding influences on the sea state damping effect in the nearshore area of Norderney, which is of main interest in this study. To get a better overview, the results of H_{m_0} and the wave directions were evaluated graphically for the entire research area. For test case #3, the northwest storm, the parameters $T_{m_{-1,0}}$, H_{m_0} *swell*, and Dir_{Spr} were shown exemplarily for the whole research area (App. I.3). Since all test cases were modeled to THW with neglectable currents, no generalizations were feasible beyond this. The modeled sea state parameters and damping effects due to the different bathymetries were analyzed for every one of the five test cases separately. Afterwards some general findings of the results were presented.

2.5.2.1 Test case #1

Test case #1 reflects a moderate wind event from the northwest at THW (Tab. 2.4). In 1995, H_{m_0} decreased slightly from 2.04–1.64 m between Coast I and Coast II (Tab. 2.6). Although the relative damping effect was only locally defined, the waves propagated almost undamped over the entire width of the coastline. Compared to all other bathymetries, H_{m_0} was highest along the entire coastline in 1995 (Fig. 2.9A). $T_{m_{-1,0}}$ also decreased least in 1995 from 7.26–7.06 s (Tab. 2.6). The waves propagated almost straight from Coast I to Coast II without being deflected by the topography (Fig. 2.9A). From 1995–2005 the greatest morphodynamic change in the northern gap of the ETD sandbanks was modeled, resulting in an increase in the absolute damping effect by 0.28–0.68 m. (Fig. 2.9A, Tab. 2.6). In 2005, the lowest H_{m_0} and $T_{m_{-1,0}}$ values were modeled with 1.36 m and 6.42 s, respectively, since the area of the ETD sandbanks showed a maximum relative damping (Fig. 2.9A, Tab. 2.6). The respective relative swell damping increased from 1995–2005 by 0.17 (Tab. 2.6). In 2006, H_{m_0} increased again around Coast II compared to 2005 (Fig. 2.9A, Tab. 2.6), as well as $T_{m_{-1,0}}$ and H_{m_0} of the swell (Tab. 2.6). This was likely caused by the gap between NWG and NG, where the waves reaching position Coast II were passing through (Fig. 2.9A). In 2008, the increasing trend of H_{m_0} continued for the area between Coast I and Coast II (Fig. 2.9A). The northern gap between NWG and NG moved east and got larger, thus more wave energy could pass through to Coast II (Fig. 2.9A). In 2015, no significant changes of H_{m_0} between Coast I and Coast II were simulated compared to 2008. H_{m_0} and $T_{m_{-1,0}}$ changed only slightly at Coast II and the mean wave direction moved north by 3° in 2008 (Fig. 2.9A, Tab. 2.6).

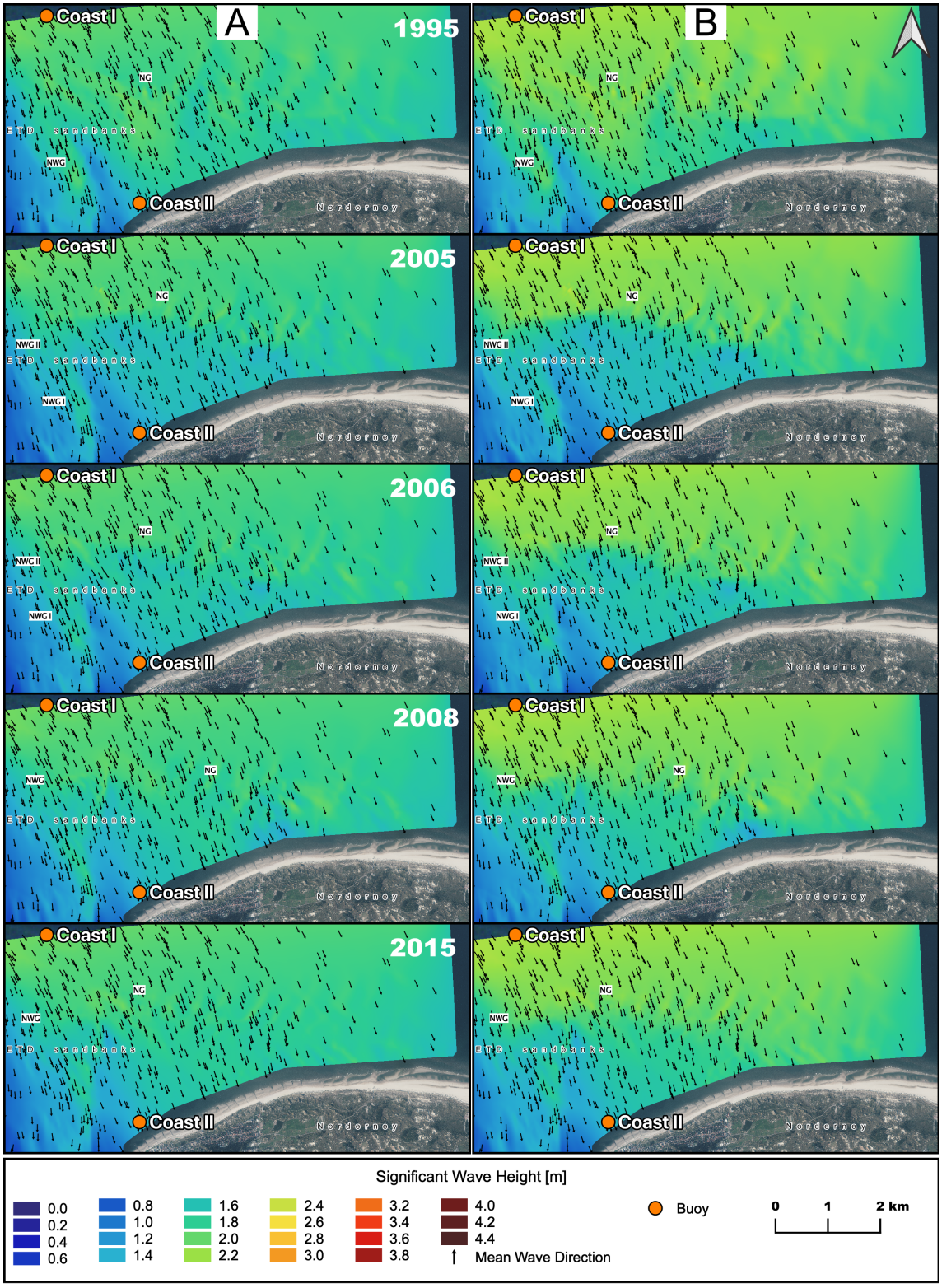

Figure 2.9: *Modeling results of H_{m_0} for the years 1995, 2005, 2006, 2008, and 2015 for test case #1 (A) and test case #2 (B). The arrows indicate the mean wave direction at the respective grid points.*

Table 2.6: *Modeled sea state parameters and damping effects of the considered years for test case #1 with damping range 0.13 and 0.17 for swell, respectively.*

Parameter	Buoy Coast	Topography				
		1995	**2005**	**2006**	**2008**	**2015**
H_{m_o} $[m]$	I	*2.04*	*2.04*	*2.04*	*2.04*	*2.04*
	II	***1.64***	*1.36*	*1.42*	*1.50*	*1.52*
$T_{m_{-1,0}}$ $[s]$	I	*7.26*	*7.26*	*7.26*	*7.26*	*7.26*
	II	***7.06***	*6.42*	*6.51*	*6.67*	*6.65*
wave direction [°]	I	*324*	*324*	*324*	*324*	*324*
	II	*337*	*336*	*337*	*336*	*339*
directional spreading [°]	I	*35*	*35*	*35*	*35*	*35*
	II	*25*	*25*	***27***	*25*	***27***
H_{m_0} *swell* $[m]$	I	*0.51*	*0.51*	*0.51*	*0.51*	*0.51*
	II	***0.44***	*0.35*	*0.37*	*0.39*	*0.39*
water depth [m]	I	*14.01*	*14.01*	*14.01*	*14.01*	*14.01*
	II	***9.03***	*7.18*	*7.49*	*7.44*	*8.02*
absolute damping effect		***0.40***	*0.68*	*0.62*	*0.54*	*0.52*
relative damping effect		***0.20***	*0.33*	*0.30*	*0.26*	*0.25*
absolute swell damping effect		***0.07***	*0.16*	*0.14*	*0.12*	*0.12*
relative swell damping effect		***0.14***	*0.31*	*0.27*	*0.24*	*0.24*

In summary, the mean relative damping effect at position Coast II of all modeled years for this test case was 0.27 ± 0.05. The relative damping effect increased from 0.20 in 1995 (minimum) to 0.33 in 2005 (maximum). The range of 0.13 or 13% indicated an order of magnitude for the influence of the ETD sandbanks of 1995–2015 on the damping effect for such an event. The mean relative damping effect for swell was 0.24 ± 0.06, thus, the swell was less affected and damped than the entire sea state for this event. With a range of 0.17 (0.14 in 1995 and 0.31 in 2005), swell was percentagewise more influenced by the ETD sandbank dynamics (Tab. 2.6). The mean wave directions between Coast I and Coast II did not change significantly due to the morphological changes of the different years (Fig. 2.9A, Tab. 2.6).

2.5.2.2 Test case #2

Test case #2 also reflects a moderate wind event from northwest at THW, but with a slightly higher wind speed and incoming wave height compared to #1 (Tab. 2.4). The results were comparable with those of test case #1. The change of H_{m_0} due to the changed damping effect from 1995–2005 was 0.41 m or 0.18 of the relative damping, respectively (Tab. 2.7). Again, in 1995, when the highest value of H_{m_0}

was modeled, the waves propagated almost undamped and homogeneously over the entire width of the coastline (Fig. 2.9B). $T_{m_{-1,0}}$ also decreased least in 1995 7.02–6.84 s (Tab. 2.7). In 2005, the lowest H_{m_0} and $T_{m_{-1,0}}$ values were simulated with 1.34 m and 6.27 s, respectively (Fig. 2.9B, Tab. 2.7). The respective relative swell damping increased from 1995–2005 by 0.19 (Tab. 2.7). In 2006, H_{m_0} increased around Coast II compared to 2005 (Fig. 2.9B, Tab. 2.7). In 2008, the decreasing trend in the relative damping effect continued for the area between Coast I and Coast II (Fig. 2.9B). The damping effect of the small plate in the east got significant (Fig. 2.9B). In 2015, again, no significant changes of H_{m_0} between Coast I and Coast II were found compared to 2008 (Fig. 2.9B, Tab. 2.7).

Table 2.7: *Modeled sea state parameters and damping effects of the considered years for test case #2 with damping range 0.18 and 0.19 for swell, respectively.*

Parameter	**Buoy Coast**	**Topography**				
		1995	**2005**	**2006**	**2008**	**2015**
H_{m_o} $[m]$	I	*2.27*	*2.27*	*2.27*	*2.27*	*2.27*
	II	***1.75***	*1.34*	*1.42*	*1.52*	*1.53*
$T_{m_{-1,0}}$ $[s]$	I	*7.02*	*7.02*	*7.02*	*7.02*	*7.02*
	II	***6.84***	*6.27*	*6.38*	*6.53*	*6.51*
wave direction [°]	I	*328*	*328*	*328*	*328*	*328*
	II	*337*	*335*	*336*	*336*	*339*
directional spreading [°]	I	*36*	*36*	*36*	*36*	*36*
	II	*26*	*27*	***28***	*25*	*27*
H_{m_0} *swell* $[m]$	I	*0.64*	*0.64*	*0.64*	*0.64*	*0.64*
	II	***0.53***	*0.41*	*0.44*	*0.47*	*0.47*
water depth [m]	I	*13.57*	*13.57*	*13.57*	*13.57*	*13.57*
	II	***8.59***	*6.74*	*7.05*	*7.00*	*7.58*
absolute damping effect		***0.52***	*0.93*	*0.85*	*0.75*	*0.74*
relative damping effect		***0.23***	*0.41*	*0.37*	*0.33*	*0.33*
absolute swell damping effect		***0.11***	*0.23*	*0.20*	*0.17*	*0.17*
relative swell damping effect		***0.17***	*0.36*	*0.31*	*0.27*	*0.27*

In summary, the mean relative damping effect at position Coast II of all modeled years for this test case was 0.33 ± 0.07, with a range of 0.18 indicating an overall stronger damping effect than in test case #1 for this event. The mean relative damping effect for swell was 0.28 ± 0.07. Thus, the swell was again less affected and damped than the entire sea state for this event. The stronger damping in

test case #2 can be explained by the higher incoming wave height at the northern model boundary (see also Section 2.5.3). The mean wave directions between Coast I and Coast II did not change significantly due to the morphological changes of the different years (Fig. 2.9B, Tab. 2.7).

2.5.2.3 Test case #3

Test case #3 represents a strong northwest storm event at THW with an incoming H_{m_0} of 3.91 m and a wind speed of 18 ms^{-1} (Tab. 2.4). For this test case, the parameters $T_{m_{-1,0}}$, H_{m_0} *swell*, and Dir_{Spr} were mapped exemplarily for the whole research area, since such an event is representative for many storm events relevant for coastal protection in this area. The effect of small-scale morphodynamic changes in form of small gaps in the ETD sandbanks was also analyzed with this test case.
In 1995, H_{m_0} decreased slightly from 3.91–2.30 m between Coast I and Coast II, where H_{m_0} again reached its maximum (Tab. 2.8). The plates in the south, NG, and the accumulating plate in the east were the only ones damping H_{m_0} significantly, while the small gaps in the southwest and northeast indicated higher values of H_{m_0} (Fig. 2.10A). H_{m_0} was a rough approximation of the topographical conditions, since the shallower water depths were associated with lower wave heights, due to wave breaking (Niemeyer and Kaiser, 1999; Salmon and Holthuijsen, 2015) (Fig. 2.10A). Compared to all other bathymetries, the waves again propagated almost homogeneously over the entire width of the Norderney coastline and H_{m_0} was highest in 1995 in the entire area around Coast II (Fig. 2.10A). $T_{m_{-1,0}}$ decreased least in 1995 from 8.19 to 6.80 s (Tab. 2.8). The waves propagated almost straight from Coast I to Coast II (Fig. 2.10A) and the directional spreading was comparable low between the two buoys (App. I.3).
In 2005, the lowest H_{m_0} and $T_{m_{-1,0}}$ values were modeled with 1.79 m and 6.01 s, respectively (Fig. 2.10A, Tab. 2.8). The area around the ETD sandbanks showed a maximum relative damping (Fig. 2.10A). The respective swell values decreased in the entire area between Coast I and Coast II from 1995–2005 (App. I.3). Compared to all other bathymetries, H_{m_0} was lower in 2005, especially on the strip perpendicular to the line Coast I and Coast II (Fig. 2.10A). While the mean wave directions between Coast I and Coast II and the directional spreading at position Coast II did not change significantly, the area of higher directional spreading has increased off the coast due to the morphological changes (App. I.3, Tab. 2.8).
In 2006, H_{m_0} increased around Coast II compared to 2005 (Fig. 2.10A), as well as $T_{m_{-1,0}}$ and H_{m_0} *swell* (App. I.3). This was likely caused by the gap between NWG and NG, where the waves reaching position Coast II were passing through (Fig. 2.10A). In addition, a lot of energy passed through south of NWG I, while

the waves were dampened east of NWG I, likely caused by the high depth of the Buise-Low in the northern parts (Fig. 2.10A). These results only apply to THW events where the currents can be neglected. Considering currents at other times, a different behavior was assumed. Thus, mainly specific small-scale morphodynamic changes of the ETD sandbanks between NWG II and NG (320°–332°) may have led to the decrease of 0.02 of the relative damping effect and the increase in H_{m_0} of 0.07 m compared to 2005 for this test case. Since the directional spreading in the east increased, an interaction with the accumulating plate was presumed (App. I.3).

Table 2.8: *Modeled sea state parameters and damping effects of the considered years for test case #3 with damping range 0.13 and 0.17 for swell, respectively.*

Parameter	**Buoy Coast**	**Topography**				
		1995	**2005**	**2006**	**2008**	**2015**
H_{m_o} $[m]$	I	*3.91*	*3.91*	*3.91*	*3.91*	*3.91*
	II	***2.30***	*1.79*	*1.86*	*2.00*	*2.01*
$T_{m_{-1,0}}$ $[s]$	I	*8.19*	*8.19*	*8.19*	*8.19*	*8.19*
	II	***6.80***	*6.01*	*6.14*	*6.35*	*6.37*
wave direction [°]	I	*325*	*325*	*325*	*325*	*325*
	II	*329*	*328*	*328*	*330*	*332*
directional spreading [°]	I	*33*	*33*	*33*	*33*	*33*
	II	*26*	*28*	***29***	*27*	***29***
H_{m_0} *swell* $[m]$	I	*1.16*	*1.16*	*1.16*	*1.16*	*1.16*
	II	***0.68***	*0.49*	*0.52*	*0.56*	*0.57*
water depth [m]	I	*14.07*	*14.07*	*14.07*	*14.07*	*14.07*
	II	***9.10***	*7.24*	*7.55*	*7.50*	*8.08*
absolute damping effect		***1.61***	*2.12*	*2.05*	*1.91*	*1.90*
relative damping effect		***0.41***	*0.54*	*0.52*	*0.49*	*0.49*
absolute swell damping effect		***0.48***	*0.67*	*0.64*	*0.60*	*0.59*
relative swell damping effect		***0.41***	*0.58*	*0.55*	*0.52*	*0.51*

In 2008, the increasing trend of H_{m_0} continued for the area between Coast I and Coast II (Fig. 2.10A). Although the northern gap between NWG and NG moved east, waves could pass likely by refraction (Fig. 2.10A). Since the gap was 13° larger (345°–10°) compared to 2006, more wave energy passed through to Coast II and the relative damping effect decreased by 0.03 (Tab. 2.8). The accumulating plate in the east caused a strong damping, indicated by low H_{m_0} and $T_{m_{-1,0}}$ and a high directional spreading (Fig. 2.10A, App. I.3).

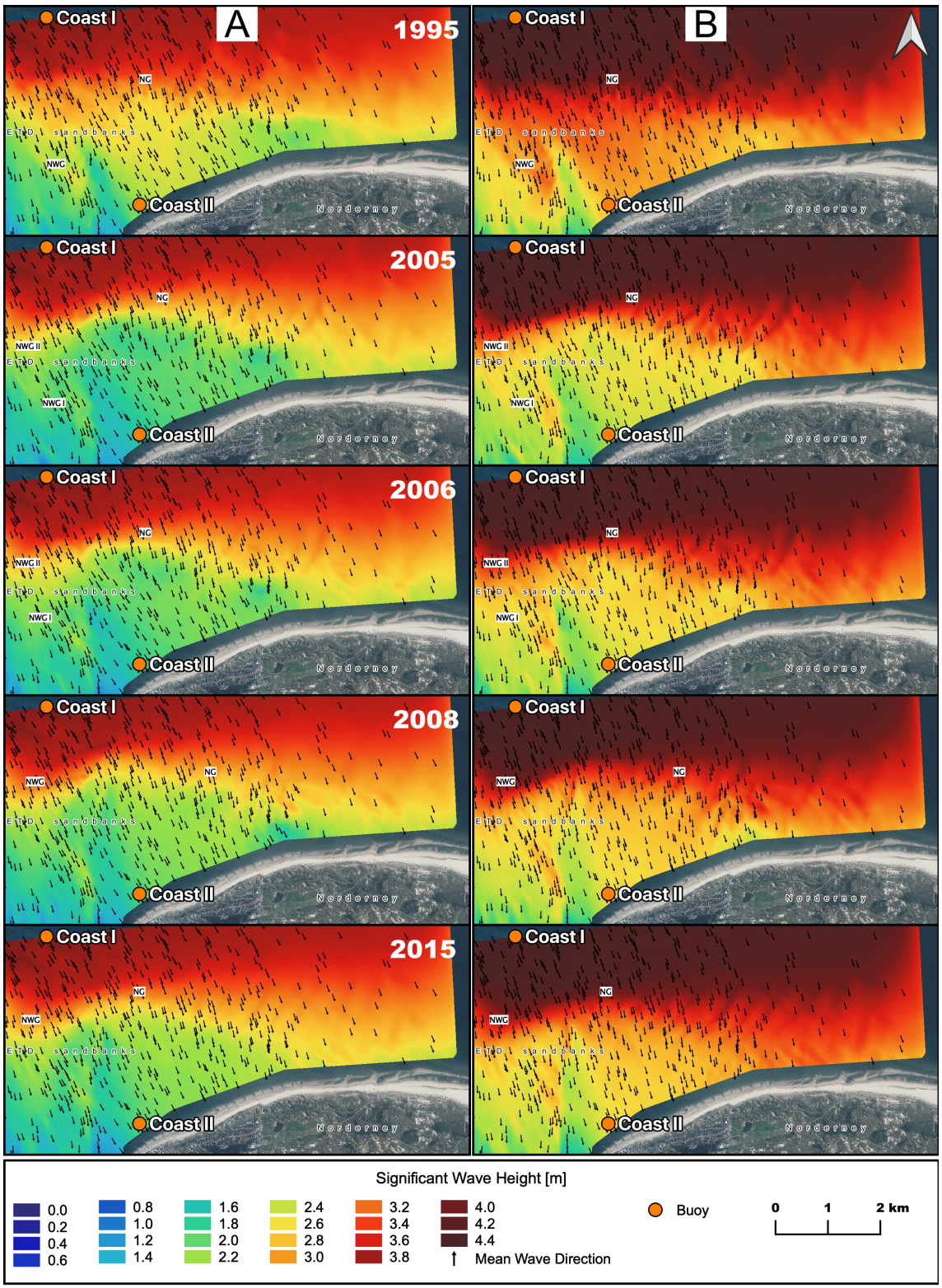

Figure 2.10: *Modeling results of H_{m_0} for the years 1995, 2005, 2006, 2008, and 2015 for test case #3 (A) and test case #4 (B). The arrows indicate the mean wave direction at the respective grid points.*

In 2015, no significant changes of H_{m_0} between Coast I and Coast II were found compared to 2008. H_{m_0} and $T_{m_{-1,0}}$ and the mean wave direction remained nearly unchanged at Coast II (Fig. 2.10A, Tab. 2.8). H_{m_0} was 0.15 m higher and, thus, the relative damping effect still 0.03 lower than in 2006, likely caused by the northern gap between NWG and NG (325°–345°) as well as refraction processes, turning incoming waves to the south. The small plates in the east accumulated at the Norderney beach, causing a H_{m_0} of about 2.5 m in 2015 compared to about 1.5 m in 2008 (Fig. 2.10A). Likewise, $T_{m_{-1,0}}$ increased at this area in 2015, while the directional spreading decreased against 2008 (App. I.3).
Small-scale morphodynamic changes in form of small gaps in the ETD sandbanks locally caused a maximum influence of 0.22 m or 0.05 of the relative damping effect on the sea state, comparing the years 2005, 2006, 2008, and 2015 for this test case (Fig. 2.10A). A generalization of the impact of small-scale morphodynamic changes is therefore not feasible. In summary, the mean relative damping effect at position Coast II of all modeled years for this test case was 0.49 ± 0.05 with a range of 0.13 indicating a range for the influence of the morphodynamic changes of the ETD sandbanks from 1995–2015 on the damping effect for this test case (Fig. 2.10A, Tab. 2.8). The significant higher mean compared to the previous test cases could be caused by the higher input wave height of the storm event. The mean relative damping effect for swell was 0.51 ± 0.06, thus, the swell was more affected and damped than the entire sea state and showed a greater interaction with the ETD sandbanks. With a range of 0.17, swell was also percentagewise more influenced by the ETD sandbank dynamics (Tab. 2.8, App. I.3). A greater effect on swell, compared to the entire sea state, was therefore only found in more severe events, likely caused by depth-induced breaking, leading to a higher variability of the swell damping effect.

2.5.2.4 Test case #4

Test case #4 represents a strong northern storm event at THW with an incoming H_{m_0} of 4.42 m (Tab. 2.4). From 1995–2005, H_{m_0} decreased by 0.58 m at Coast II resulting in an increase of 0.13 in the relative damping effect (Tab. 2.9). In this test case, the waves propagated almost homogeneously over the entire width of the coastline in 1995, when the highest value of H_{m_0} was modeled (Fig. 2.10B). $T_{m_{-1,0}}$ decreased least in 1995 from 9.13 to 7.71 s (Tab. 2.9). In 2005, the lowest H_{m_0} and $T_{m_{-1,0}}$ values were modeled with 2.36 m and 7.21 s, respectively (Fig. 2.10B, Tab. 2.9). The respective relative swell damping also increased from 1995–2005 by 0.13 (Tab. 2.9). In 2006, H_{m_0} increased along the coastline compared to 2005 (Fig.

2.10B, Tab. 2.9). In 2008, the relative damping effect continued to decrease by 0.02 for Coast II (Tab. 2.9). In 2015 again, no significant changes of H_{m_0} between Coast I and Coast II were simulated compared to 2008 (Fig. 2.10B, Tab. 2.9).
In summary, the mean relative damping effect at position Coast II of all modeled years for this test case was 0.42 ± 0.05, with a range of 0.13 indicating an overall weaker damping effect than in test case #3. The reduced damping effect could be explained by the more northern storm direction, as the water level increased and the waves tended to run over the ETD sandbanks. Compared to test case #3, the mean wave directions south of the ETD sandbanks, as well as the mean wave direction at Coast II, were further north (Fig. 2.10B, Tab. 2.9). The mean relative damping effect for swell was 0.56 ± 0.05, thus, the swell was again more affected and damped than the entire sea state for this event. Compared to test case #3, swell was damped significantly more in absolute and relative terms.

Table 2.9: *Modeled sea state parameters and damping effects of the considered years for test case #4 with damping range 0.13 and 0.13 for swell, respectively.*

Parameter	**Buoy Coast**	**Topography**				
		1995	**2005**	**2006**	**2008**	**2015**
H_{m_o} $[m]$	I	*4.46*	*4.46*	*4.46*	*4.46*	*4.46*
	II	***2.94***	*2.36*	*2.48*	*2.60*	*2.59*
$T_{m_{-1,0}}$ $[s]$	I	*9.13*	*9.13*	*9.13*	*9.13*	*9.13*
	II	***7.71***	*7.21*	*7.31*	*7.43*	*7.37*
wave direction [°]	I	*334*	*334*	*334*	*334*	*334*
	II	*338*	*338*	*339*	*339*	*341*
directional spreading [°]	I	*36*	*36*	*36*	*36*	*36*
	II	*26*	*28*	***29***	*27*	*28*
H_{m_0} *swell* $[m]$	I	*2.93*	*2.93*	*2.93*	*2.93*	*2.93*
	II	***1.53***	*1.15*	*1.22*	*1.31*	*1.29*
water depth [m]	I	*14.72*	*14.72*	*14.72*	*14.72*	*14.72*
	II	***9.74***	*7.89*	*8.20*	*8.15*	*8.73*
absolute damping effect		***1.52***	*2.10*	*1.98*	*1.86*	*1.87*
relative damping effect		***0.34***	*0.47*	*0.44*	*0.42*	*0.42*
absolute swell damping effect		***1.40***	*1.78*	*1.71*	*1.62*	*1.64*
relative swell damping effect		***0.48***	*0.61*	*0.58*	*0.55*	*0.56*

2.5.2.5 Test case #5

Test case #5 is the only one that represents a strong storm event from the west at THW with an incoming H_{m_0} of 3.98 m (Tab. 2.4). Apart from the wind direction,

the boundary conditions of this test case were comparable to test case #3. From 1995–2005 H_{m_0} decreased by 0.49 m at Coast II resulting in an increase of 0.12 of the relative damping effect (Tab. 2.10). Again, in 1995, the waves propagated almost homogeneously over the entire width of the coastline (Fig. 2.11A). $T_{m_{-1,0}}$ decreased least in 1995 from 8.62 to 7.66 s (Tab. 2.10). In 2005, again the lowest H_{m_0} and $T_{m_{-1,0}}$ values were modeled for Coast II with 1.58 m and 6.76 s, respectively (Tab. 2.10). From 1995–2005 the respective relative swell damping increased by 0.16 (Tab. 2.10). In 2006, the relative damping effect again decreased along the coastline compared to 2005 (Fig. 2.11A, Tab. 2.10). In 2008, it continued to decrease by 0.03 for Coast II (Tab. 2.10). In 2015 again, no significant changes of H_{m_0} and the damping effects between Coast I and Coast II were modeled compared to 2008 (Fig. 2.11A, Tab. 2.10). A significant damping by the small plate in the east was modeled in 2015.

Table 2.10: *Modelled sea state parameters and damping effects of the considered years for test case #5 with damping range 0.12 and 0.16 for swell, respectively.*

Parameter	**Buoy Coast**	**Topography**				
		1995	**2005**	**2006**	**2008**	**2015**
H_{m_o} $[m]$	**I**	*3.89*	*3.89*	*3.89*	*3.89*	*3.89*
	II	***2.07***	*1.58*	*1.66*	*1.77*	*1.76*
$T_{m_{-1,0}}$ $[s]$	I	*8.62*	*8.62*	*8.62*	*8.62*	*8.62*
	II	***7.66***	*6.76*	*6.98*	*7.23*	*7.21*
wave direction [°]	I	*324*	*324*	*324*	*324*	*324*
	II	*322*	*318*	*318*	*322*	*322*
directional spreading [°]	I	*36*	*36*	*36*	*36*	*36*
	II	*29*	*34*	*34*	*33*	***35***
H_{m_0} *swell* $[m]$	I	*2.13*	*2.13*	*2.13*	*2.13*	*2.13*
	II	***1.11***	*0.77*	*0.83*	*0.92*	*0.91*
water depth [m]	I	*13.80*	*13.80*	*13.80*	*13.80*	*13.80*
	II	***8.82***	*6.97*	*7.28*	*7.23*	*7.81*
absolute damping effect		***1.82***	*2.31*	*2.23*	*2.12*	*2.13*
relative damping effect		***0.47***	*0.59*	*0.57*	*0.54*	*0.55*
absolute swell damping effect		***1.02***	*1.36*	*1.30*	*1.21*	*1.22*
relative swell damping effect		***0.48***	*0.64*	*0.61*	*0.57*	*0.57*

In summary, the mean relative damping effect at position Coast II for all modeled years for this test case was 0.54 ± 0.05, with a range of 0.12. It was the highest mean damping effect of all test cases. The high damping effect could be explained by

the more westerly wind direction, which led to a lower water level and an approach of the waves from the west to the ETD sandbanks. Compared to test case #3, the mean wave directions south of the ETD sandbanks, as well as the mean wave direction at Coast II, were further west, since wind sea had a greater influence on the directions (Fig. 2.11A, Tab. 2.10). The alignment by refraction to the beach was visible. The mean relative damping effect for swell was 0.57 ± 0.06, i.e., this event had not the highest damping effect of all five cases.

2.5.2.6 Test cases – general findings

The interactions of the waves with the ETD sandbanks of the different bathymetries were clearly visible for all test cases, since the direction of wave propagation showed an arched movement west and east around the reef at particularly high sections. At the westside of the island the waves propagated southwards along the Norderney Seegat for all bathymetries and all test cases. Throughout the test cases, propagation of H_{m_0} showed lowest values at the southside of the plates. In the Buise-Low area, where the water depth is comparatively deep, H_{m_0} was lowest (Fig. 2.9–2.11).
In general, the ETD sandbanks caused a mean variation of the local damping effect of 0.14 ± 0.02 and 0.16 ± 0.02 for swell with the lowest damping in 1995 and the highest damping in 2005. These values indicate the mean range of the relative damping effects for the different bathymetries of the five test cases and quantify the influence of the ETD sandbanks on the relative damping effect at position Coast II. A minimum range of 0.12 for test case #5 (0.13 for test case #4 for swell) and a maximum range of 0.18 for test case #2 (0.19 for test case #2 for swell) were modeled. Thus, the range of the damping effect remained rather constant between the different bathymetries, regardless of the boundary conditions.
In addition to the variability caused by the bathymetries, the variability of the different boundary conditions showed a range of 0.27 (0.33 for swell) with means of 0.41 ± 0.11 (0.43 ± 0.16 for swell) for the mean relative damping effects of all test cases, i.e. the various boundary conditions. Thus, the influence of the boundary conditions in the simulations was much higher than the influence of the different bathymetries.
It was also found, that the damping effect increased with increasing incoming wave height (see also Section 2.5.3). With westerly winds and related lower water levels, stronger damping effects were modeled for particular storm events than with more northerly winds and related higher water levels. This could be explained by a new growth of the wind sea beyond the ETD sandbanks at position Coast II at northern wind events and an overall higher water level. The range of the damping effect, however, was independent of the water level at the buoy positions.

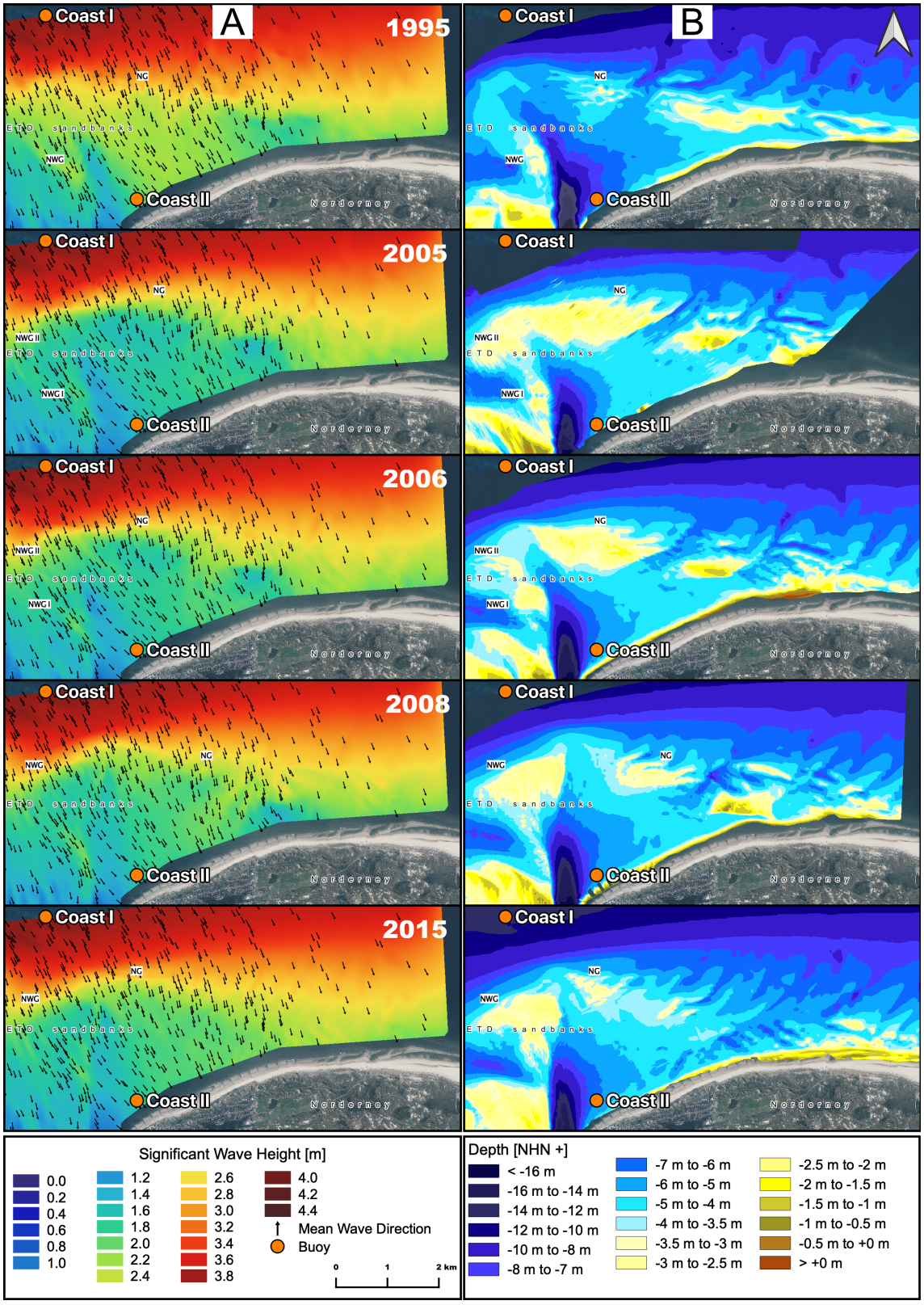

Figure 2.11: *Modeling results of H_{m_0} for test case #5 (A) and the respective bathymetries for a better overview (B). The years 1995, 2005, 2006, 2008, and 2015 are shown. The arrows (A) indicate the mean wave direction at the respective grid points.*

Compared to the entire sea state, swell was relatively more damped for high than for low wave events, likely caused by depth-induced wave breaking. Higher damping effects of the swell compared to the entire sea state were found in northern winds than in western winds. Since the wind sea could increase again behind the ETD sandbanks, especially for northern winds, it had a stronger influence on the wave height and direction of the entire sea state behind the plates. Swell on the other hand was breaking at the ETD sandbanks.
The SWAN simulations of the different years within the test cases differed only in the bathymetries, so the results were interpreted as differences caused only by topographic changes of the ETD sandbanks. The mean correlation between the area of NG and the relative damping effects of the five test cases is significant at $r = 0.94$ and $r = 0.89$ for the DTMs and aerial photos, respectively. Since the waves approach from northwest, they cross NG before reaching Coast II. Between the area of NWG and the relative damping effect, however, no significant correlation was found.

2.5.3 General findings of the buoy-measured damping effect

While the validated SWAN simulations provided detailed spatial information on the influence of the ETD sandbanks on the sea state damping effect for events with few available measurements, this section focuses on the general findings of the buoy-measured damping effect. The overall buoy-measured damping effects were described statistically and the general model results were exemplarily transferred to the in-situ buoy data by filtering.
Extreme events are of particular interest, nevertheless only 9% of all measurements have $H^I_{m_0} \geq 2m$. For the year 1995 no data were available after preprocessing. Therefore, a comparison between the modeled and measured damping effects of small-scale morphodynamic changes of the ETD sandbanks is not suitable, since an isolation of the topographic changes from the hydraulic boundary conditions could not be performed due to not enough data after filtering. However, the general findings of SWAN modeling could be confirmed by buoy measurements, since the model was validated before.
The mean relative damping effect of all measurements from all years and for all hydraulic conditions was about 0.37, distinctly skewed to the left (Tab. 2.11, App. I.4). On average about 37% of the wave height was dissipated by the ETD sandbanks and other influences between Coast I and Coast II. A standard deviation of about 0.18 indicated a high variability in the measured damping effect, which was not caused by the dynamic of the ETD sandbanks alone, but mainly by the parameters described in Section 2.4.5.

Table 2.11: *Various filterings of the measured data. The top value in each box denotes the mean damping, the middle value the standard deviation, and the bottom value the number of observations. Denote: 95% of all wave heights $H^I_{m_0}$ are smaller than 2.38 m and only 9% of all wave heights are greater than 2 m.*

filtering method	**mean 2004–2017**		**filtering method**	**mean 2004–2017**	
	Δ_{rel}	Δ_{abs}		Δ_{rel}	Δ_{abs}
no filtering	*0.37* *0.18* *65 809*	*0.42* *0.42* *65 809*	–	–	–
$2.5\ m \leq H^I_{m_0} \leq 3.5\ m$ $U_{See} \geq 15ms^{-1}$, THW $240° \leq U_\theta \leq 300°$	*0.47* *0.07* *118*	*1.41* *0.30* *118*	$2.5\ m \leq H^I_{m_0} \leq 3.5\ m$ $U_{See} \geq 15ms^{-1}$, THW $330° \leq U_\theta \leq 30°$	*0.40* *0.09* *35*	*1.23* *0.35* *35*
$1.5\ m \leq H^I_{m_0} \leq 2.5\ m$	*0.41* *0.17* *9 713*	*0.76* *0.34* *9 713*	$1.5\ m \leq H^I_{m_0} \leq 2.5\ m$ *swell*	*0.26* *0.27* *6 795*	*0.36* *0.17* *9 713*
$H^I_{m_0} \geq 3\ m$	*0.59* *0.11* *1 282*	*2.25* *0.69* *1 282*	$H^I_{m_0} \geq 3\ m$ *swell*	*0.63* *0.41* *596*	*0.65* *1.05* *1 282*

Since SWAN showed a dependency of the sea state damping effect from the incoming wave height at Coast I and the prevailing water level, their impact on the buoy-measured damping effect is shown in Fig. 2.12. According to the model results, the buoy-measured damping effect increased with increasing wave height at Coast I and decreased with an increasing water level at gauge Norderney Riffgat (Niemeyer, 1979; Niemeyer and Kaiser, 1999). An exponential relationship between these factors was derived (Fig. 2.12). The events of the five boundary conditions of the SWAN modeled test cases were highlighted in the graph and showed a wide range of damping effects and water levels (Fig. 2.12, Tab. 2.4). During storm surges, the water level did not increase as fast as the growing of waves, so that the damping effect was huge at storm surges (Niemeyer, 1979). Previous results showed, that there was no linear increase in the mean water level for the years analyzed in this study (see Section 2.4.3.3), nevertheless the tides caused a high variability of the water level. As high waves are usually linked to high wind speeds, the damping effect was greater with stronger winds. Furthermore, other general findings of the simulations were reflected in the buoy-measured damping effect; Tab. 2.11 shows some filtering.

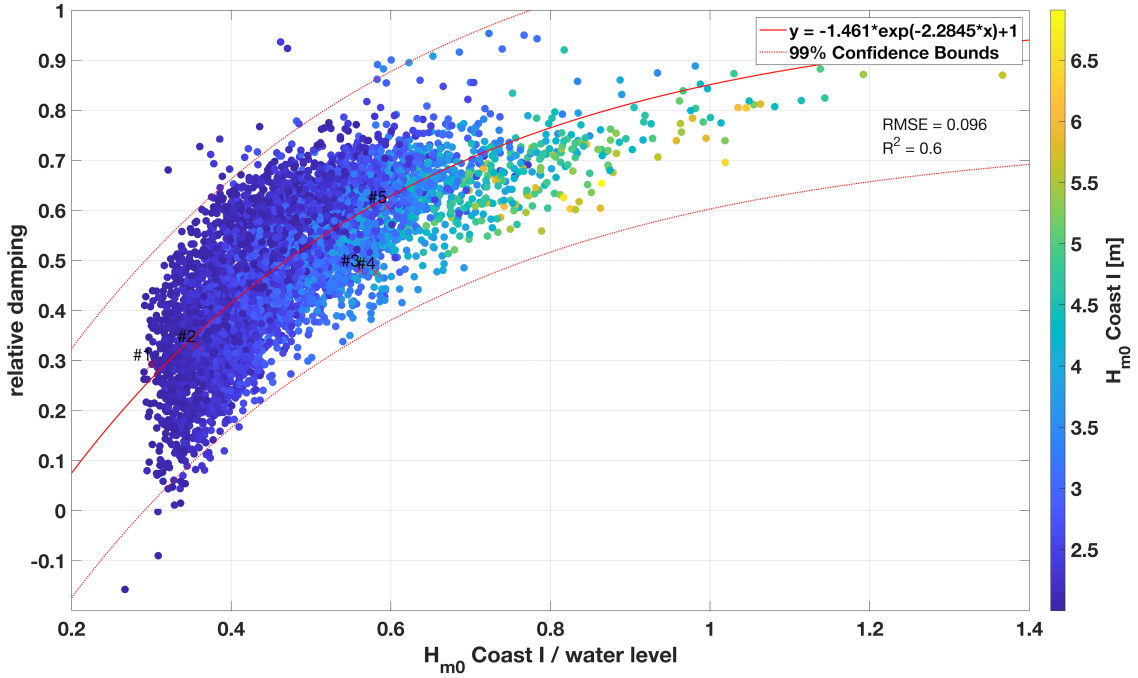

Figure 2.12: *Relative sea state damping in relation to the quotient of incoming H_{m_0} at position Coast I and the water level. Only wave events with $H^I_{m_0} \geq 2m$ are used. A total of $n = 5252$ data points are given and a correlation coefficient of $r = 0.73$ is obtained. An exponential fit with its 0.99-confidence boundaries is plotted on the data. The events of the five boundary conditions of the test cases for SWAN modeling are highlighted in the graph.*

Although only few data remained, the buoy data of all years combined showed a stronger damping at west winds (240°–300°) compared to north winds (330°–30°) for storm events with $U_{See} \geq 15ms^{-1}$ and 2.5 $m \leq H^I_{m_0} \leq 3.5$ m at THW (Tab. 2.11). This was in line with the results of SWAN modeling. Higher relative damping effects were exemplarily obtained for swell at $H^I_{m_0} \geq 3$ m ($\Delta_{rel} = 0.63$) as for 1.5 $m \leq H^I_{m_0} \leq 2.5$ m ($\Delta_{rel} = 0.26$) in absolute terms and relatively to the overall measured damping effect (Tab. 2.11). However, since only $H^I_{m_0}$ was filtered, the standard deviation was high.

Filtering, which constrained the data to the exact boundary conditions of the five test cases, was not feasible due to missing values after filtering, caused by the rare occurrence of such relevant events. Differences between SWAN-modeled and buoy-measured damping effects can occur due to the temporal resolution of the ETD sandbank bathymetries. While storm surges and extreme events can lead to rapid changes in ETD morphology (Bennett et al., 2019; Phillips et al., 2017; Wang et al., 2012), buoy data of an entire year had to be assigned to a single bathymetry and wave climate. This may cause inaccuracies, requiring the development and use of high-resolution measurement methods for the ETD sandbanks, even though approaches by coastal radar already exist (Bell et al., 2016; Bird et al., 2017).

In complex sea state systems such as the southern North Sea, an exact separation between wind sea and swell is difficult to determine. Thus, the applied PM-peak method tends to overestimate wind sea in some cases, which can also lead to a high standard deviation (Portilla et al., 2009).

2.6 Conclusions

This study quantitatively determined a significant influence of the ETD sandbank dynamics on the sea state damping at the nearshore area of Norderney. Non-linear influences, mainly caused by the ETD sandbanks, clearly characterized the wave climate in the nearshore area. Sizes and shapes of single ETD plates showed high dynamics, with mean movement velocities of 435–491 my^{-1}, depending on the position of the plates.

By SWAN modeling, significant differences of the relative damping effects between the bathymetries of the highly different sized ETD sandbanks of 1995, 2005, 2006, 2008, and 2015 were shown for five different test cases. Those morphodynamical changes had a mean impact of 14 $\pm$ 2% on the variation of the relative damping effect in the validated SWAN simulations of various severe storm events at THW. The events had a mean damping effect of 0.41 $\pm$ 0.11 over all bathymetries and all of the five test cases, while the effect was highest in 2005 and lowest in 1995. A mean variation of the damping effect for the test cases of 16 $\pm$ 2% was modeled for swell waves.

It was also found, that the damping effect increased with increasing incoming wave height. With westerly winds and lower water levels, stronger damping effects were modeled for particular storm events than with more northerly winds and higher water levels. Compared to the entire sea state, swell was relatively more damped for high than for low wave events. Since the nearshore wave climate is strongly influenced by the tides, only THW events with negligible currents were considered in this study. Nevertheless, it is still of interest to evaluate the influence of tidal currents on the sea state damping effect more precisely.

The overall buoy-measured relative damping effect showed a mean of 0.37 $\pm$ 0.18, while no reliable conclusions on the specific morphodynamic influences of the ETD sandbanks on the sea state damping could be provided by the in-situ data, since an isolation of the topographic changes from the hydraulic boundary conditions was not applicable due to insufficient data availability. Therefore, SWAN modeling was an appropriate method to fill measurement gaps for relevant rare events.

Due to measurement and buoy failures, the reconstruction of missing data seems

to be essential. Especially during extreme events, which are of special interest for coastal protection, measurement failures often occur. A further problem arises from the assignment of a single ETD sandbank topography to the wave data of a whole year, caused by the bearing measurement intervals. The development and use of high-temporal resolution measurement methods, e.g., a coastal radar to record sandbank topography at a high temporal resolution is therefore essential. We propose that the found dynamics of the ETD sandbanks and the quantification of the damping effect have to be considered by the responsible authorities for coastal protection, whereby further investigations on the specific influences of the boundary conditions on the damping effect are of great interest.

3 Research Study II

Prediction and reconstruction of ocean wave heights based on bathymetric data using LSTM neural networks

Co-Authors: Cordula Berkenbrink & Britta Stumpe

Published in Ocean Engineering (2021), 232 (accepted 18.04.2021)

Abstract

Since climate change impacts threaten the coastal regions of the North Sea, consistent sea state time series are essential for building coastal protection or offshore structures. Vast gaps in buoy data, caused by false measurements or maintenance periods, require reconstruction and forecasting of ocean waves. Morphodynamic changes of the Ebb-Tidal Delta (ETD) sandbanks exert a huge impact on the local wave climate off the island of Norderney, Germany. Therefore, the objective of this paper is to develop a machine learning model based on Long Short-Term Memory (LSTM) neural networks for reconstruction and short- and long-term prediction of nearshore significant wave height (SWH), integrating bathymetric data for the first time. Time series of sea state and weather data of adjacent buoys as well as bathymetric data of the ETD sandbanks from 2004 to 2017 were used and the networks were tuned with Bayesian hyperparameter optimization. Including the bathymetries improved the performance of the LSTM for SWH reconstruction, short-, and long-term predictions by 16.7%, 7.4–11.7%, and 8.8–9.1% in terms of RMSE, respectively. The LSTM outperformed deep feed-forward neural networks and other state-of-the-art machine learning algorithms. With an RMSE of 0.069 m, a parallel LSTM structure (P-LSTM) is the proposed method for SWH reconstruction.

Keywords: Long short-term memory (LSTM) neural network, Machine learning, Wave height prediction, Forecasting ocean waves, Bayesian optimization, Dynamic bathymetries

3.1 Introduction

Climate change impacts, such as sea level rise, a possible increase in storm frequency and intensity, a possible trend in increasing ocean wave heights, and global warming in general, will become emerging challenges for coastal regions of the North Sea (Grabemann and Weisse, 2008; Mori et al., 2013; Vanem and Walker, 2013). Furthermore, storm surges are a frequently recurring and well-known danger (Bitner-Gregersen et al., 2018; Weisse et al., 2012). Substantial knowledge of the sea state is therefore essential for building coastal protection or offshore structures and marine operations, as well as for feeding numerical and climate models (Ti et al., 2018). Wave height is consequently one of the most important wave characteristics for such purposes (Teich et al., 2018) and commonly measured by networks of moored wave buoys. Time series of such wave buoys typically have gaps due to erroneous measurements or maintenance operations (Londhe, 2008). Especially extreme events are often recorded erroneously, although they are of great interest in ocean science (Cavaleri, 2009). Thus, sophisticated tools for reconstructing and predicting ocean wave heights are required.

Especially in coastal areas, several complex processes influence the propagation of waves and wave height. The morphodynamic change of Ebb-Tidal Delta (ETD) sandbanks of the coastal area of Norderney, Germany exerts a huge impact on the local wave climate (Jörges et al., 2021a). Caused by ebb-tidal currents, the ETD sandbanks are dynamically changing by migration, erosion, and accretion (Castelle et al., 2007; Dallas and Barnard, 2011) and thus influence the reconstruction and forecast of coastal wave heights by increasing the complexity of the prediction (Spicer et al., 2019). The ETD sandbanks are also a natural coastal protection by damping the incoming wave height, especially for extreme events (Jörges et al., 2021a; Niemeyer, 1979; Niemeyer and Kaiser, 1999).

Usually, numerical models such as Simulating WAves Nearshore (SWAN) are used for predicting irregular, random waves. However, they are often expensive in computational costs but precise due to exact physical equations. Soft-computing frameworks as machine learning algorithms can be an alternative in some cases (Abed-Elmdoust and Kerachian, 2012; Fernández et al., 2015; Mahjoobi et al., 2008; Malekmohamadi et al., 2011; Mandal and Prabaharan, 2010; Nitsure et al., 2012). Combinations of numerical modeling and soft-computing techniques also exist (Malekmohamadi et al., 2008; Puscasu, 2014; Reikard et al., 2011). Artificial neural networks (ANN) in particular are widely used nowadays for wave height reconstruction and prediction, namely the feed-forward neural networks (FFNN) as a common example (Deo and Naidu, 1999; Deo et al., 2001; Law et al., 2020; Londhe and Panchang, 2007; Peres

et al., 2015; Zamani et al., 2008). In the project MOSES (Herman et al., 2007), time series of spatial wave climate data of the nearshore area of Norderney were reconstructed with FFNNs for the period 1962–2002. As a result, numerical modeling including SWAN was enhanced by an efficient soft-computing method (Herman et al., 2009). Berkenbrink and Niemeyer (2017) analyzed a changing salinity intrusion due to estuarine waterway deepening in the German Bight with FFNN. James et al. (2018) used FFNN with numerically modeled SWAN data of Monterey Bay and reduced the computational time by a factor >1000 compared to the physics-based model. ANN approaches with a genetic algorithm (Altunkaynak, 2013) and evolutionary product unit neural networks (EPUNN) (Durán-Rosal et al., 2016) were also used. Kumar et al. (2017) predicted ocean wave heights with a Radial Basis Function (RBF) called Minimal Resource Allocation Network (MRAN), while Stefanakos (2016) used the Adaptive Network-based Fuzzy Inference System (ANFIS). Furthermore, approaches with Extreme Learning Machine (ELM) (Ali and Prasad, 2019; Kumar et al., 2018) and Discrete Wavelet Transforms Neural Networks (Deka and Prahlada, 2012; Dixit and Londhe, 2016) were successfully implemented for wave height forecasting. However, for sequential data, recurrent neural networks (RNN), ANNs with feedback connections to neurons from previous time steps (LeCun et al., 2015), are suitable and still used for wave height prediction (Balas et al., 2004). The Long Short-Term Memory (LSTM) neural network is a specific RNN and was designed by Hochreiter and Schmidhuber (1997) to handle long-term dependencies. An LSTM decides on its own weather to remember or to forget information (Gers et al., 2000). Besides the use for speech and handwriting recognition (Graves and Jaitly, 2014; Graves et al., 2009), machine translation (Wu et al., 2016), image captioning (Aggarwal, 2018), and stock price prediction (Kim and Won, 2018), LSTMs are already used for reconstruction and prediction of time series in various related fields such as wind speed forecasting (Li et al., 2018), sea surface temperature prediction (Xiao et al., 2019), solar irradiance prediction (Qing and Niu, 2018), and weather forecasting (Salman et al., 2018). First experiments with LSTM neural networks for the reconstruction and prediction of nearshore wave heights with data of the Norderney coastal area were quite promising (Jörges, 2018; Jörges et al., 2020). Pirhooshyaran et al. (2020) first applied a sequence-to-sequence neural network with LSTM encoder and decoder to reconstruct and predict offshore and deep ocean wave heights of the North Pacific and North Atlantic Ocean. They achieved a reconstruction R^2 of up to 0.81 at a deep ocean buoy position near Los Angeles. Fan et al. (2020) used LSTM neural networks for one and 6-h ahead predictions of offshore and deep ocean significant wave height and combined it with the third-generation

numerical wave model SWAN to improve the predictions. 6-h ahead predictions were performed with an R^2 of up to 0.97 for a deep ocean buoy east of Florida. Kagemoto (2020) satisfactorily predicted water-surface wave trains and the motion of a floating body in it with LSTM neural networks, while Ni and Ma (2020) applied a LSTM neural network for the prediction of polar westerlies wave height with only few measurements available. However, to the best of our knowledge, none of the previous studies have yet considered complex morphodynamic conditions or integrated bathymetric data for the reconstruction or prediction of ocean wave heights with LSTM neural networks. In addition, the LSTM has not yet been applied to shallow waters in nearshore areas, where morphodynamic influences typically are greater.
Thus, the objective of this study was to design an LSTM based reconstruction and short- and long-term prediction model for significant wave height including bathymetric data to overcome the challenge of morphodynamic changes in the nearshore area. The reconstruction and short- and long-term predictions were performed on buoy measurements of neighboring buoys off the Norderney coastline – separated by the ETD sandbanks – and wind and water level parameters as well as bathymetric data of the ETD sandbanks.
In a first step, a model comparison of several machine learning methods was performed on the nearshore wave data to benchmark the LSTM neural network, while the proposed method was also validated on well-studied offshore buoy measurements off the U.S. west coast. In a second step, the impact of including bathymetric data on model performance was investigated by experiments, and further feature selections were tested. The main contributions of this study are as follows:

- A LSTM neural network and an ensemble of parallel LSTM networks are presented for reconstruction and short- and long-term prediction of nearshore wave heights
- Bathymetric data are presented as features to improve wave height reconstruction and prediction for the first time

3.2 Methods

3.2.1 Research area and dataset

3.2.1.1 Research area

The research area is located in the nearshore area of the East Frisian island of Norderney in the German North Sea (Fig. 3.1). The islands Juist (west) and Baltrum (east) are separated from Norderney by the tidal inlets Norderneyer Seegat

and Wichter Ee, respectively. Between the barrier islands high tidal currents are present, which activate the sediment transport of the ETD and the tidal flats. The long-term Mean High Water (MHW) at gauge Norderney Riffgat is 1.24 m above the German reference height (NHN) or the German mean sea level with a mean tidal range of 2.45 m. The main wind direction of the research area is southwest (Fig. 3.1).

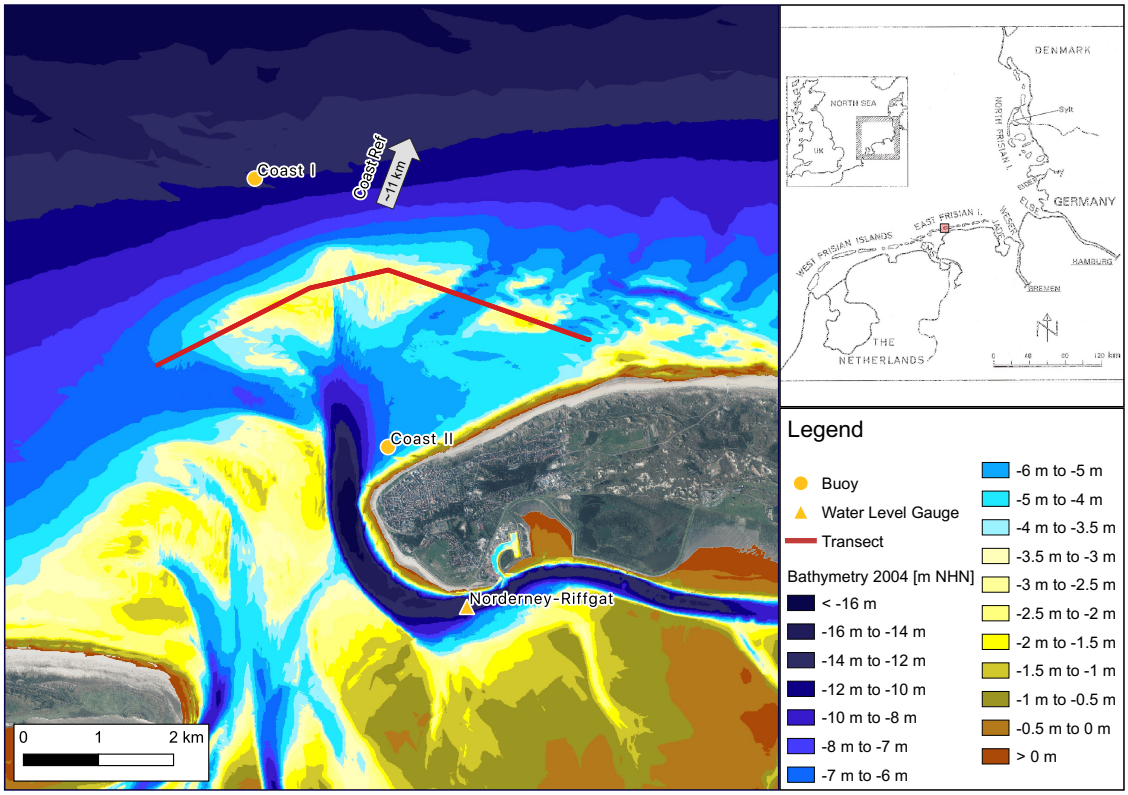

Figure 3.1: *Map of the research area and the positions of the buoys. The transect of all bathymetries (red line) and the bathymetry of 2004 (extend of the ETD sandbanks) are shown on the map.*

3.2.1.2 Dataset

Buoy data of three Waverider (Datawell, Haarlem, The Netherlands) buoys were provided by the Coastal Research Station of the Lower Saxony Water Management, Coastal Defence and Nature Conservation Agency (NLWKN), Norderney (Tab. 3.1). While the two buoys Coast Ref and Coast I are located north of the ETD sandbanks – the reef arch west and north of Norderney – the predicted position Coast II is located south of it and directly in front of the Norderney beach (Fig. 3.1). Nearshore wave height at buoy Coast II thus is greatly influenced by morphodynamic processes of the ETD sandbanks (Jörges et al., 2021a; Niemeyer, 1979). The parameters sig-

nificant wave height (SWH), mean wave period, energy period, and mean wave direction were calculated from the energy density spectra of the raw buoy data. These data were available in half-hourly resolution for the years 2004–2017 without the years 2012–2014 and with many missing values throughout the dataset (Tab. 3.1). Fig. 3.2 shows a sample portion of the SWH at position Coast I. Besides several data gaps in the annual record (top), the influence of the tides was significant in the daily record (bottom).

Table 3.1: *Basic information of the different buoy positions.*

Buoy Position	SWH Mean [m]	SWH Std. [m]	SWH Range [m]	Mean energy wave period [s]	# Data	Location	Water depth 2015 [m NHN]
Coast Ref	1.20	0.76	[0.09; 8.40]	5.62	84 428	N 53.836653 E 7.193904	23.39
Coast I	1.07	0.69	[0.09; 6.91]	5.77	95 887	N 53.744781 E 7.112101	12.11
Coast II	0.63	0.39	[0.04; 3.01]	6.04	88 804	N 53.715653 E 7.141280	6.15

Wind data were provided by the German Meteorological Service (Deutscher Wetterdienst, DWD) for the station Norderney at a resolution of 10 min and water level data by the Waterways and Shipping Administration (WSA) for the gauge level Norderney-Riffgat at a resolution of 1 min. All parameters were averaged to a half-hourly resolution.

Morphodynamic changes of the ETD sandbanks were obtained by the bathymetries of yearly surveys, which were provided by the NLWKN, Coastal Research Station, and EasyGSH-DB. Due to the huge financial effort and time required, additional survey missions by ship were not possible in this challenging area. The EasyGSH-DB bathymetries and metadata were obtained from the EasyGSH-DB portal `www.easygsh-db.org` (Hagen et al., 2020) and only used after a validation process for years where no measurement data of the NLWKN Coastal Research Station were available.

Fig. 3.3 shows the correlogram of relevant metric parameters. Significant correlations between the SWH of Coast II and the other positions were found ($r = 0.82$ and $r = 0.83$). Wind speed and water level also showed correlations to the SWH at Coast II with $r = 0.56$ and $r = 0.48$, respectively. Peak wave periods did not show any relevant correlations to the SWH at Coast II (Fig. 3.3). The feature selection

was conducted according to these correlations by selecting parameters with $r > 0.3$ to SWH at Coast II. Parameters with low correlation coefficients also can have a high non-linear influence. The correlogram also provides a better insight into the structure of the dataset.

The feature parameters significant wave height H_{m_0} $[m]$ of the buoys Coast Ref and Coast I, wind speed at sea $[ms^{-1}]$, wind direction $[°]$, water level $[m + NHN]$, the tidal phase {ebb, low water, flood, high water}, and the morphodynamic shape of the ETD sandbanks (see 3.2.2) were used for the reconstruction and prediction of the significant wave height H_{m_0} at buoy Coast II.

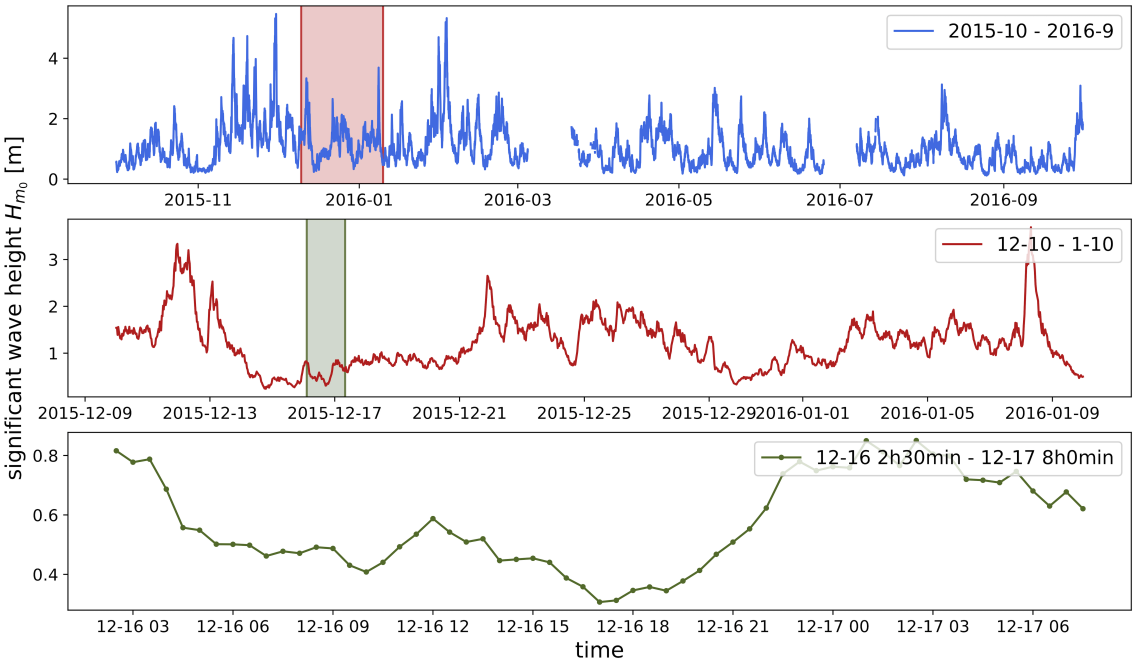

Figure 3.2: *Time series of the SWH at position Coast I from October 2015 to September 2016 with zoom in to one month and one day.*

3.2.1.3 Data preprocessing

As a preprocessing, all measurement rows or time steps with at least one missing value (NaN) of any buoy or water level and wind measurement were deleted completely. Then, all features were min-max-normalized to the interval $[0, 1]$ by

$$x_N = \frac{x - min(x)}{max(x) - min(x)} \tag{3.1}$$

where x_N is the normalized feature value and x the measured feature value. The normalization was performed to prevent certain features from having a greater in-

fluence on the neurons input by their range alone. No further data preprocessing beyond that was done in order to not manipulate the prediction results in any way (Ali and Prasad, 2019).

The dataset was randomly divided into train (70%) and test data (30%) for every model setting or model run to obtain robust results by forcing the algorithms to work fully independent of the specific training data. All error metrics given in this study (see 3.2.5) refer to the rescaled predictions, thus did not underlie the normalization and can easily be interpreted. The normalization only refers to the min and max of the training data, to ensure the models did not see any information (e.g., the range) of the test data during the training process.

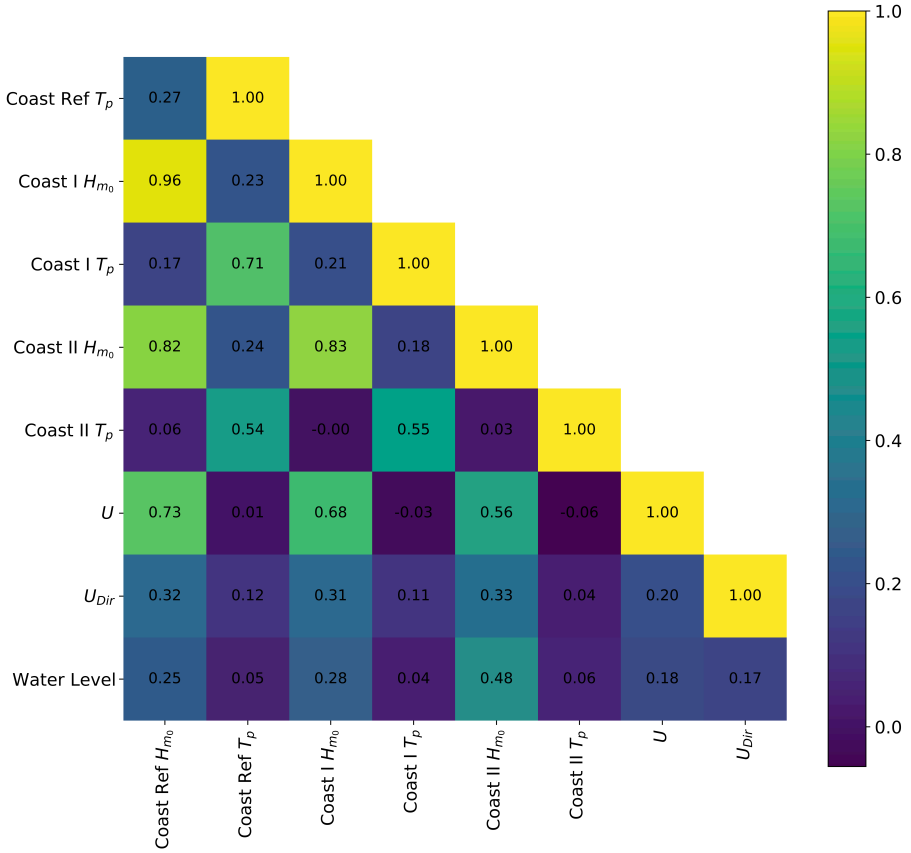

Figure 3.3: *Correlogram of the moored wave buoy parameters of Coast Ref, Coast I, and Coast II, wind speed and direction, and water level.*

3.2.2 Morphodynamic influences and bathymetry preprocessing

Since the ETD sandbanks exert a huge and dynamical influence on the sea state damping in the nearshore area of Norderney – depending on their shape and size

the damping effect varies on average by 14% – better reconstruction and prediction performances were expected by including the bathymetric data (Jörges et al., 2021a). In order to avoid overfitting and to achieve a better generalization of the models, a dimension reduction of the bathymetries was carried out. For this purpose, a one-dimensional transect covering the mean ETD sandbank line over all years was defined from the two-dimensional measurement data (Fig. 3.1). In this way, both a good representation of the yearly ETD sandbanks and a lower dimensional dataset by omitting unimportant information were achieved and thus used for data selection. Due to the different resolutions of the bathymetries, the transect was divided into six equidistant sections and the mean values of each sections bottom heights were used as input for the models (Tab. 3.2). This feature engineering process results in a strong dimension reduction and generalization of the bathymetries already with the input data. Thus, the error caused by the bathymetries with only annual resolution was addressed. Substantial variations in the characteristics of the ETD sandbank within the respective years were observed, while the mean bottom height of the transect over all bathymetries was -3.68 ± 0.20 m (Tab. 3.2).

Table 3.2: *ETD sandbank heights, mean, and standard deviation [m] of the different bathymetries along the transect from west (Sec. 1) to east (Sec. 6) in relation to NHN-0m.*

year	Sec. 1	Sec. 2	Sec. 3	Sec. 4	Sec. 5	Sec. 6	Mean	Std.
2004	-4.82	-2.62	-2.85	-2.58	-4.44	-3.46	-3.46	0.88
2005	-4.46	-3.03	-2.36	-2.80	-3.94	-2.69	-3.21	0.74
2006	-4.80	-3.54	-3.01	-3.31	-4.09	-2.64	-3.57	0.71
2007	-5.79	-3.09	-4.11	-3.42	-3.46	-3.53	-3.90	0.90
2008	-5.04	-2.94	-3.21	-3.82	-3.03	-4.13	-3.70	0.74
2009	-5.00	-3.02	-2.97	-3.43	-3.52	-4.50	-3.74	0.76
2010	-4.97	-2.86	-2.96	-3.47	-3.68	-4.42	-3.73	0.76
2011	-5.07	-3.01	-3.15	-3.51	-3.67	-4.49	-3.82	0.73
2015	-5.22	-3.10	-3.29	-3.41	-3.68	-4.23	-3.82	0.72
2016	-5.56	-3.56	-2.78	-3.54	-3.86	-4.05	-3.89	0.84
2017	-5.72	-2.86	-2.90	-2.74	-3.75	-4.07	-3.67	1.04

Caused by the huge difference of the measurement intervals between the bathymetries and the sea state data, we matched a wave event to the bathymetry corresponding to the same measurement year. This led to the greatest accuracy, as the ship measurements were usually carried out during the summer month of a year. Nevertheless, strong storm surges can lead to rapid changes of ETD sandbanks and inaccuracies in matching the bathymetries to the buoy data (Wang et al., 2012).

3.2.3 Long Short-Term Memory neural networks

LSTM neural networks are special RNN able to learn long term dependencies and coping with the vanishing gradient problem (Gers et al., 2000; Hochreiter and Schmidhuber, 1997; Pascanu et al., 2013).
Each LSTM cell (Fig. 3.4) consists of four units, the forget gate f_t, input gate i_t, memory cell state C_t, and output gate o_t at time t, defined as:

$$f_t = \sigma(W_f h_{t-1} + W_f x_t) \tag{3.2}$$

$$i_t = \sigma(W_i h_{t-1} + W_i x_t) \tag{3.3}$$

$$\tilde{C}_t = tanh(W_C h_{t-1} + W_C x_t) \tag{3.4}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{3.5}$$

$$o_t = \sigma(W_o h_{t-1} + W_o x_t) \tag{3.6}$$

$$h_t = o_t \cdot tanh(C_t) \tag{3.7}$$

where x_t, h_t, $\tilde{C}_t$, W, and σ are the input data, hidden state, temporary cell state, weight matrices, and sigmoid activation function, respectively. The operation $\cdot$ indicates the Hadamard matrix product for matrices with equal dimensions. Biases are not added explicitly, since they can be integrated to the weight matrices. The final output $\hat{y}_t = W_y h_t$ is used in the loss function (Eq. 3.8) (Gers et al., 2000; Hochreiter and Schmidhuber, 1997; Qing and Niu, 2018).
To translate the abstract learned features into the prediction output, a fully-connected layer was added after the last LSTM layer.
Additionally, a parallel LSTM structure (P-LSTM) was designed. Therefore, an additive resampling method similar to Pirhooshyaran et al. (2020) was performed to split the data for the respective parallel ensemble members: The entire dataset was chronologically split into s subsets of equal size. Then, each subset was randomly divided into a training and a test dataset as described above. The proposed LSTM model was trained independently on each of these training subsets, concatenated with the entire training dataset to learn additional functional relations hidden in the entire dataset (Pirhooshyaran et al., 2020). P-LSTM performance was evaluated on the union of all test subsets by the mean prediction of the parallel ensemble models. None of the ensemble models has ever seen this test data during training. For $s = 1$, the P-LSTM structure is the same as the standard LSTM structure of this study, but trained on twice the set of training data. Since this leads to no new information, we set $s > 1$, with the best results obtained for $s = 2$. The P-LSTM ensemble

method with additive resampling in the background results in the same number of test samples as the other algorithms, hence it is comparable to them.

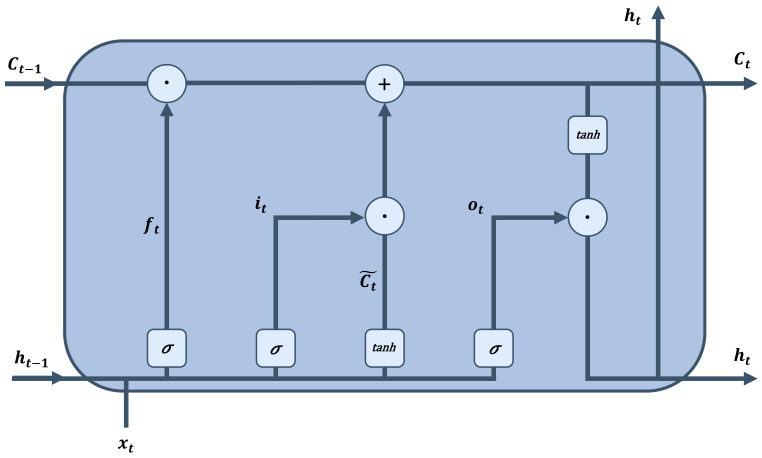

Figure 3.4: *LSTM cell with forget gate f_t, input gate i_t, memory cell state C_t, and output gate o_t at time t, and the sigmoid and tangent hyperbolic activation functions σ and tanh, respectively.*

3.2.4 Model settings

In all computational runs, the proposed LSTM and P-LSTM models were always compared to a standard FFNN (Hastie et al., 2009), a random forest (RF) (Breiman, 2001), a support vector regression (SVR) (Vapnik, 1998), and a multiple linear regression (MLR) as benchmark methods found in the literature (Fan et al., 2020; Pirhooshyaran et al., 2020; Qing and Niu, 2018). While MLR with least-squares fitting is a simple linear model that allows both a good interpretability of how the inputs affect the outputs by considering the β-coefficients and low computational costs, RF belongs among the more complex supervised machine learning algorithms. Based on several random decision trees fitted to a dataset, the RF algorithm outputs a mean prediction of the ensemble. Each tree is grown from a random subset of the features which are used as candidates for the best binary split at each node at random. By repeated selection with replacement from the training data (bagging), uncorrelated random trees are created. The bootstrap samples left out by selection – out-of-bag samples – are used for the calculation of the prediction error (Breiman, 2001; Hastie et al., 2009).

SVR is based on the idea of a support vector machine (SVM) applied on regression problems. Using an ϵ-sensitive loss function makes the SVR robust and sparse since it is less affected by noisy inputs (Vapnik, 2000). Thus, only prediction errors outside a defined ϵ-tube with $\epsilon > 0$ are penalized by the L2-objective function of

the coefficient vector. Data points around the ϵ-tube are called the support vectors. Another parameter $C > 0$ controls the regularization of the loss function for predictions outside the ϵ-tube. Both ϵ and C were set to $\epsilon = 0.08$ and $C = 1.5$ by Bayesian hyperparameter optimization (see 3.2.6). Non-linear functions are approximated with SVR by mapping the data points to a higher dimensional feature space using a kernel function (Awad and Khanna, 2015). In this study, the radial basis function kernel and the Python LIBSVM implementation were used for calculations (Chang and Lin, 2021).

A model input sequence $X = (x_1, \ldots, x_k)$ in this study consists of the last k values of all input variables $x_i = (x^{(1)}, \ldots, x^{(s)})$, where s is the number of input features. Since the measurements were available in half-hourly resolution, the $k/2$ last hours were used as input. $Y = (y_1, \ldots, y_l)$ is the output sequence for the prediction until time step $t+l$, where t denotes the current time step. In this study, predictions were made on the SWH of the current time step ($l = 0$), the next 3 h ($l = 6$), 6 h ($l = 12$), 9 h ($l = 18$), and 12 h ($l = 24$), which according to the literature is an appropriate interval for short-term predictions (Berbić et al., 2017; Fan et al., 2020). Long-term predictions were made for an usual range of 24 h-ahead ($l = 48$) and 48 h-ahead ($l = 96$) (Dixit and Londhe, 2016; Pirhooshyaran et al., 2020). The lookback k on the input sequence was set to 6 for all models, thus the last 3 h of all input features were used for the predictions. Especially the past wind speed and direction, the fetch, is crucial for the further propagation of the sea state, since the development of ocean waves takes a certain time (Holthuijsen, 2007; Nitsure et al., 2012). Every setting consisting of a specific statistical model and the predicted time step l were calculated in an ensemble of three independent runs to cope with random variations in the performance of the training process and the split in training and testing data. A robust model can only be assumed if the variation of the error metrics around the mean prediction error is small. The neural networks were trained for 100 epochs each.

All calculations were performed with Python 3.8 and TensorFlow 2 on an Intel Xeon Gold 6146 CPU 12 Core 3.2Ghz, 64 GB RAM, and a Quadro P1000 GPU.

3.2.5 Loss function, optimizer, and error metrics

The loss function L used in the model is the mean squared error (MSE) with an L2 regularization term:

$$L(X, W) = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 + \frac{1}{2} \lambda \|W\|_2^2 \tag{3.8}$$

where n is the number of data, $\hat{y}$ the predicted value, y the measured value and λ the tunable regularization parameter for the 2-norm $\|W\|_2$ of the weight matrices W. L2 regularization was used to reduce overfitting in the models (Goodfellow et al., 2016; Ng, 2004; Nie et al., 2010).

The gradient of the loss function is calculated by backpropagation with the adaptive moment estimation (Adam) optimization algorithm (Kingma and Ba, 2015). The weights are updated for mini batches of samples X with a tunable size $b \leq n$ of a power of 2. A tunable decay of the tunable learning rate α is used to achieve a more robust convergence of the optimization process.

As error metrics the root mean squared error (RMSE), the mean absolute error (MAE), the mean arctangent absolute percentage error (MAAPE) (Kim and Kim, 2016), the Pearson correlation coefficient r, and the coefficient of determination R^2 were used in this study:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{3.9}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \mid \hat{y}_i - y_i \mid \tag{3.10}$$

$$MAAPE = \frac{1}{n}\sum_{i=1}^{n} arctan\left(\left|\frac{\hat{y}_i - y_i}{y_i}\right|\right) \tag{3.11}$$

$$r = \frac{cov(\hat{y}, y)}{\sqrt{(var(\hat{y})var(y)}} \tag{3.12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3.13}$$

where $cov(x, y)$ is the covariance of x, y, $var(x)$ the variance of x, and $\bar{y}$ the mean of the measured values.

The RMSE (Eq. 3.9) is equal to the square root of the MSE and scale-dependent as it has the same dimension as the predicted parameter (here [m]). Thus, an RMSE of 0 indicates a perfect fit, while the prediction error increases with increasing RMSE. MAE (Eq. 3.10) is scale-dependent, but does not penalize large errors as much as MSE and RMSE. r on the other hand is dimensionless with $-1 \leq r \leq 1$ and can only be used for describing linear correlations (Kuhn and Johnson, 2013). MAAPE (Eq. 3.11) is scale-independent and dimensionless, overcoming the problem of division by zero or values close to zero (Kim and Kim, 2016).

3.2.6 Hyperparameter tuning

In order to find the best hyperparameters of the proposed model, a Bayesian hyperparameter optimization was performed on a training dataset (Bergstra et al., 2011; Snoek et al., 2012). The number of LSTM layers, number of LSTM nodes in each layer, activation function, mini batch size b, learning rate α, decay of learning rate α, dropout rate between the LSTM layers, and regularization parameter λ were tuned. A number of 50 iterations were performed maximizing the scale-independent R^2 objective function during the tuning process.

The Bayesian optimization handles the LSTM model as a random black box function on the hyperparameters, which is expensive to evaluate. Therefore, the algorithm estimates the best hyperparameters with a minimum of objective function evaluations (Brochu et al., 2010). Prior, a Gaussian Process is built over the objective functions and best hyperparameters are selected according to their likelihood. After evaluating the objective function with those hyperparameters, the posterior is built and the next hyperparameters for function evaluation can be determined again (Cabrera et al., 2020; Cornejo-Bueno et al., 2018; Snoek et al., 2012).

The partly discrete search domain and the results of the Bayesian optimization are shown in Tab. 3.3. The results of the Bayesian hyperparameter search were used for all LSTM and P-LSTM models as well as for all FFNN models to have a comparable architecture.

Table 3.3: *Hyperparameter tuning values and results.*

Parameter	Search interval	Final value
# layers	[1, 5]	2
# nodes per layer	[1, 128]	50
activation function	{relu, sigmoid, tanh}	relu
mini batch size b	[1, 128]	64
learning rate α	[1e-4, 1e-2]	0.0025
decay of learning rate α	[1e-6, 1e-2]	0.00025
dropout rate between layers	[0, 0.5]	0.025
regularization parameter λ	[1e-7, 1e-3]	0.000001

3.3 Results and discussion

In a first step, the performance of our proposed model for significant wave height reconstruction ($l = 0$), short-term ($l = 6,\ 12,\ 18,\ 24$), and long-term ($l = 48,\ 96$) prediction was compared to other common statistical and machine learning models, both in the nearshore area and on a commonly studied deep ocean buoy dataset. In a next step, we investigated the improvements of the reconstruction and short- and long-term prediction models by adding the bathymetries to the training data. The test data were not shown to the networks during training in any way. Also, the influence of other feature selections on the model performance was investigated.

3.3.1 Model comparison – nearshore excluding bathymetry features

To evaluate the performance of the proposed LSTM and P-LSTM architecture for the different reconstruction and short- and long-term prediction settings excluding the bathymetry, the models were compared to a two-layered deep fully connected FFNN with the same settings for learning rate, optimization algorithm, batch size, and number of neurons as the LSTM models as well as a simple single-layered fully connected FFNN (SL-FFNN), a SVR, a RF, and a common MLR.

3.3.1.1 Reconstruction

Tab. 3.4 gives an overview of the results for the reconstruction task excluding and including bathymetries. The unknown SWH of buoy Coast II was reconstructed by the known SWH of the other buoys and the wind and water level data only. By excluding the bathymetries, the influence of the ETD sandbanks was implicitly considered with the buoy data of Coast II, since this position is influenced by those. Best reconstruction results excluding bathymetry features were achieved with the P-LSTM and LSTM neural networks with RMSEs of 0.079 and 0.084 m, respectively (Tab. 3.4). Because the hyperparameter tuning of the FFNN was initialized with the optimal tuning parameters of the LSTM and the initial values were among the best results, the same hyperparameters allowed a proper comparison of both models. The performances of the FFNN and RF and the SL-FFNN and SVR were quite similar (Tab. 3.4). For all models, the scale-independent metric r was higher than the scale-independent metric R^2 and the scale-dependent metric RMSE higher than the scale-dependent metric MAE. Fig. 3.5 shows the results of the reconstruction excluding the bathymetries by comparing measured and predicted values. The data points of the two LSTMs were closest around the bisector. Although extreme values tended to be slightly underestimated, the reconstruction error did not increase

further at higher wave heights. The FFNN, SVR, and RF showed deviations of the reconstruction with increasing wave height, while extreme wave heights were underestimated especially by the RF (Fig. 3.5). Since the regression line (solid red line) was consistently below the bisector (dashed red line), all models underestimated the SWH on average.

The MLR distinctly showed the greatest spreading around the bisector and the regression line with increasing wave height and tended to overestimate extreme values (Fig. 3.5). Nevertheless, the MLR is easy to interpret compared to the other methods. The β-coefficients of the MLR provide information about the directional influence of the different input parameters on the prediction. For the reconstruction excluding bathymetric data, the β-coefficient for time step t of H_{m_0} Coast I was highest with 0.457 ± 0.007 and 0.131 ± 0.034 for H_{m_0} Coast Ref. This corresponds to the correlation of the input parameters, which also showed the highest value for H_{m_0} Coast I (Fig. 3.3). With $r = 0.96$, the correlation between H_{m_0} Coast Ref and H_{m_0} Coast I was very high, which could explain the low influence of this variable in the MLR. A mean β-coefficient of 0.322 ± 0.032 was found for the water level of time step t, while the general correlation between this variable and H_{m_0} Coast II was $r = 0.48$. Although the correlation between the wind speed and H_{m_0} Coast II was $r = 0.56$, the β-coefficient at time step t for this variable was low at 0.054 ± 0.009. This may be due to the high correlation of $r = 0.73$ between H_{m_0} Coast Ref and the wind speed (Fig. 3.3).

A comparison of the reconstructed time series is given in Fig. 3.6. The LSTM and P-LSTM (dashed green lines) showed an overall good fit to the measured data (solid grey line). A high volatility of the other methods, except the FFNN, were found for local extreme values.

Fig. 3.7 shows the error of the standardized training and test data during the training process of the LSTM and FFNN over the 100 epochs for the reconstruction excluded the bathymetries. While the performance on the training and test data were nearly equal for the LSTM, the FFNN had a significantly worse performance on the test data. It was assumed that the LSTM learned a better generalization of the time series data due to its recurrent units and various gates, while the FFNN showed a slight overfitting (Gers et al., 2000). Hyperparameter tuning showed that deeper networks did not increase the test performance, but led to an overfitting of the training data.

Table 3.4: *Reconstruction of significant wave height H_{m0} performance results excluding and including bathymetric data. The mean value and standard deviation of the three independent runs are given, respectively.*

Model	Bathymetry Features	RMSE [m]	R^2	r	MAAPE	MAE [m]
LSTM	Excluded	0.084 ± 0.003	0.948 ± 0.003	0.976 ± 0.000	0.109 ± 0.002	0.061 ± 0.002
	Included	0.070 ± 0.001	0.964 ± 0.002	0.983 ± 0.001	0.096 ± 0.001	0.052 ± 0.001
P-LSTM	Excluded	0.079 ± 0.001	0.953 ± 0.001	0.977 ± 0.001	0.105 ± 0.000	0.058 ± 0.000
	Included	**0.069** ± 0.001	**0.965** ± 0.002	**0.984** ± 0.000	**0.093 ± 0.001**	**0.051** ± 0.001
FFNN	Excluded	0.107 ± 0.001	0.916 ± 0.002	0.959 ± 0.000	0.137 ± 0.000	0.077 ± 0.001
	Included	0.085 ± 0.001	0.947 ± 0.001	0.974 ± 0.001	0.114 ± 0.003	0.062 ± 0.000
SL-FFNN	Excluded	0.125 ± 0.000	0.884 ± 0.001	0.942 ± 0.002	0.171 ± 0.002	0.093 ± 0.000
	Included	0.101 ± 0.005	0.925 ± 0.008	0.965 ± 0.002	0.145 ± 0.010	0.076 ± 0.004
SVR	Excluded	0.122 ± 0.001	0.890 ± 0.002	0.949 ± 0.001	0.201 ± 0.002	0.097 ± 0.001
	Included	0.114 ± 0.000	0.904 ± 0.001	0.953 ± 0.001	0.189 ± 0.002	0.092 ± 0.000
RF	Excluded	0.110 ± 0.000	0.911 ± 0.000	0.955 ± 0.000	0.147 ± 0.001	0.081 ± 0.000
	Included	0.104 ± 0.001	0.921 ± 0.002	0.960 ± 0.001	0.142 ± 0.001	0.077 ± 0.000
MLR	Excluded	0.152 ± 0.000	0.830 ± 0.002	0.911 ± 0.001	0.210 ± 0.001	0.113 ± 0.000
	Included	0.149 ± 0.000	0.837 ± 0.001	0.915 ± 0.001	0.208 ± 0.002	0.111 ± 0.000

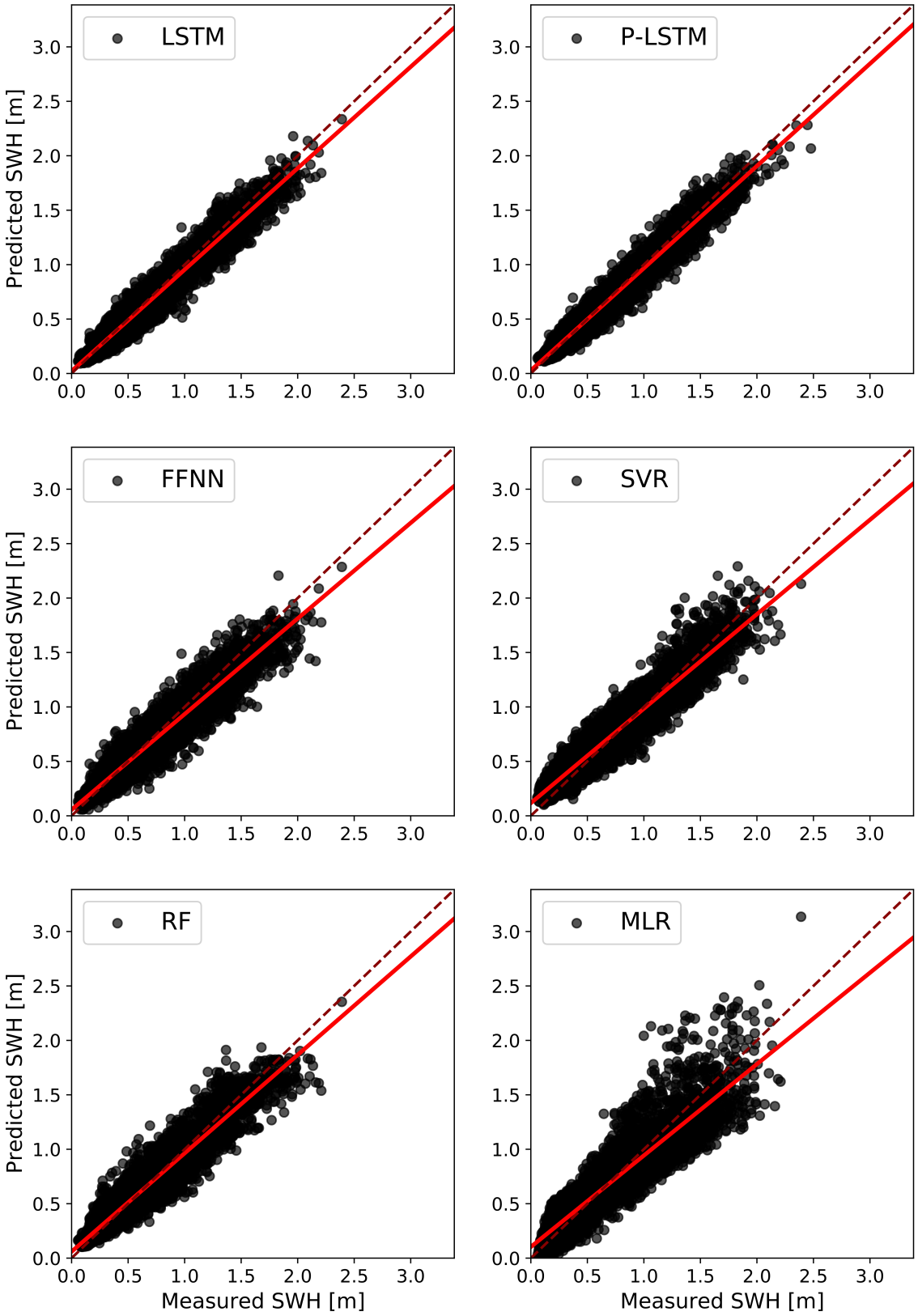

Figure 3.5: *Comparison between the measured and predicted SWH for the reconstruction excluding bathymetries for the LSTM, P-LSTM, FFNN, SVR, RF, and MLR on the test data.*

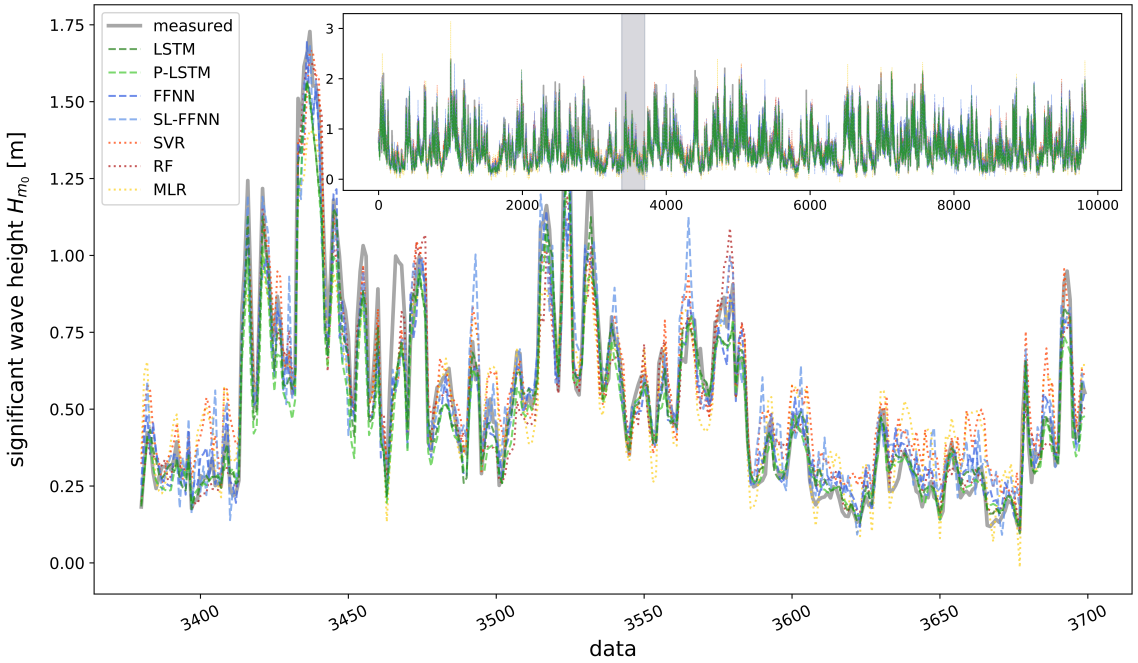

Figure 3.6: *Time series of SWH reconstruction excluded the bathymetries. The inset shows the entire test dataset.*

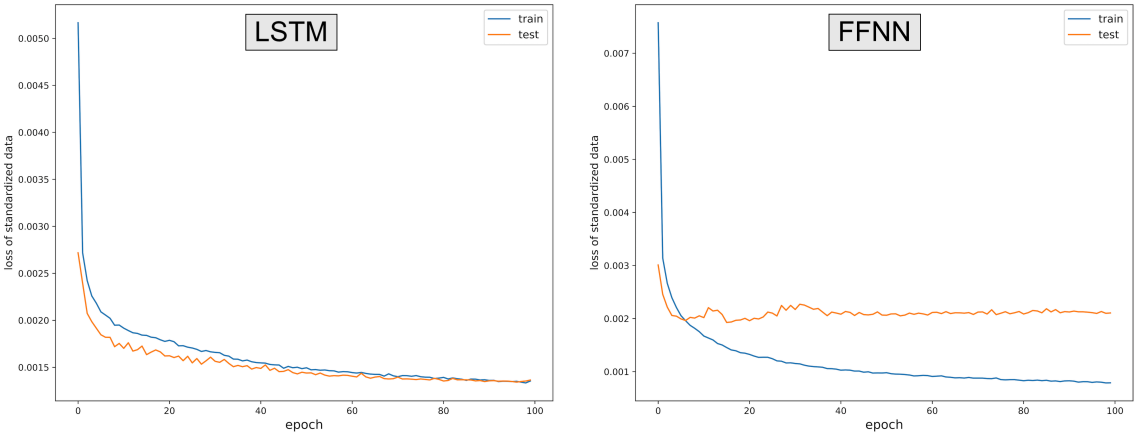

Figure 3.7: *Loss of standardized train and test data during the 100 training epochs of the LSTM (left) and FFNN (right) for the reconstruction excluded the bathymetries.*

3.3.1.2 Short-term prediction

Not only for the reconstruction but also for the multi-step forecasts, the performance of the P-LSTM and LSTM were better than those of the other methods. The P-LSTM parallel structure with additive resampling consequently showed a slightly better performance than the non-parallelized LSTM structure for all short-term prediction intervals. The performance for the 3 h, 6 h, 9 h, and 12 h-ahead forecasts of the significant wave height for buoy Coast II by historical data of the surrounding buoys is given in Tab. 3.5. As expected, the performance of all models became increasingly worse, as the prediction interval increased. Fig. 3.8 shows the logarithmically increasing RMSE of the LSTM for the different prediction intervals. For the prediction task, the performance of the RF was better than the SVR and, except for the 3 h-ahead prediction, it was better than the FFNN (Tab. 3.5). Nevertheless, differences between the models could be identified. While the R^2 of the LSTM for the 3 h-ahead prediction increased by 4.7% compared to the FFNN, by 7.7% compared to the SVR, and by 5.3% compared to the RF, it increased by 10.2%, 11.2%, and 7.8% for the 6 h-ahead prediction, 13.5%, 15.8%, and 10.9% for the 9 h-ahead prediction, and by 20.3%, 21.8%, and 13.6% for the 12 h-ahead prediction, respectively (Tab. 3.5). The error metric of r showed a similar behavior (Tab. 3.5). This indicates that the LSTM and P-LSTM performed better compared to the other machine learning algorithms with increasing forecasting horizon. On the other hand, the RMSE of the FFNN, SVR, and RF was 20.7%, 31.5%, and 23.4% higher than the RMSE of the LSTM for the 3 h-ahead prediction, 24.3%, 25.7%, and 19.3% higher for the 6 h-ahead prediction, 22.7%, 25.2%, and 18.4% for the 9 h-ahead prediction, and 23.6%, 25.3%, and 17.6% higher for the 12 h-ahead prediction, respectively (Tab. 3.5). MAE showed a similar pattern. Fig. 3.9 shows the differences between the measured and modeled values for the 12 h-ahead prediction excluding bathymetric data. The P-LSTM and LSTM showed the narrowest line here as well. The FFNN showed a higher variability for extreme values than the RF and a similar scatter plot to the SVR. Differences between the several models were most significant for this forecasting period.

Thus, the improvement of the LSTM and P-LSTM for the different predictions in terms of R^2 and r increased with increasing forecasting period, while the improvement in terms of RMSE and MAE was equal or decreased with an increasing forecasting period. Nevertheless, it corresponds to the expectation that the LSTM and P-LSTM can bring their power to bear, especially with larger forecasting periods. This highlights the broad potential of LSTMs for time series data (Fan et al., 2020).

Table 3.5: *Short-term (3–12h) and long-term (24–48h) forecasting significant wave height H_{m_0} performance results excluding and including bathymetric data. The mean value and standard deviation of the three independent runs are given, respectively.*

Prediction	Model	Excluded Bathymetry					Included Bathymetry				
		RMSE [m]	R^2	r	MAAPE	MAE [m]	RMSE [m]	R^2	r	MAAPE	MAE [m]
3h-ahead	LSTM	0.111 ± 0.002	0.910 ± 0.003	0.957 ± 0.001	0.142 ± 0.004	0.080 ± 0.001	0.098 ± 0.001	0.928 ± 0.002	**0.966** ± 0.000	0.131 ± 0.002	0.072 ± 0.000
	P-LSTM	0.106 ± 0.000	0.918 ± 0.000	0.959 ± 0.000	0.140 ± 0.000	0.078 ± 0.000	**0.096** ± 0.000	**0.932** ± 0.000	**0.966** ± 0.001	**0.128** ± 0.002	**0.070** ± 0.000
	FFNN	0.134 ± 0.001	0.869 ± 0.001	0.933 ± 0.000	0.173 ± 0.002	0.097 ± 0.000	0.119 ± 0.001	0.895 ± 0.000	0.946 ± 0.000	0.159 ± 0.002	0.087 ± 0.001
	SL-FFNN	0.152 ± 0.003	0.831 ± 0.003	0.916 ± 0.001	0.190 ± 0.002	0.111 ± 0.001	0.135 ± 0.003	0.864 ± 0.004	0.933 ± 0.002	0.180 ± 0.005	0.099 ± 0.002
	SVR	0.146 ± 0.002	0.845 ± 0.005	0.924 ± 0.002	0.232 ± 0.003	0.115 ± 0.002	0.137 ± 0.000	0.860 ± 0.001	0.931 ± 0.001	0.217 ± 0.000	0.108 ± 0.000
	RF	0.137 ± 0.001	0.864 ± 0.004	0.930 ± 0.002	0.181 ± 0.001	0.100 ± 0.001	0.133 ± 0.000	0.868 ± 0.002	0.932 ± 0.001	0.178 ± 0.001	0.098 ± 0.000
	MLR	0.168 ± 0.002	0.794 ± 0.005	0.891 ± 0.003	0.231 ± 0.001	0.126 ± 0.001	0.166 ± 0.001	0.795 ± 0.000	0.892 ± 0.000	0.230 ± 0.001	0.124 ± 0.001
6h-ahead	LSTM	0.140 ± 0.001	0.855 ± 0.002	0.925 ± 0.001	0.190 ± 0.000	0.103 ± 0.000	0.128 ± 0.001	0.881 ± 0.003	0.940 ± 0.001	**0.167** ± 0.002	0.094 ± 0.001
	P-LSTM	0.137 ± 0.002	0.863 ± 0.003	0.930 ± 0.000	0.179 ± 0.001	0.101 ± 0.001	**0.124** ± 0.001	**0.887** ± 0.003	**0.944** ± 0.002	**0.167** ± 0.002	**0.092** ± 0.001
	FFNN	0.174 ± 0.000	0.776 ± 0.002	0.882 ± 0.002	0.223 ± 0.002	0.127 ± 0.001	0.157 ± 0.001	0.822 ± 0.004	0.907 ± 0.002	0.200 ± 0.004	0.113 ± 0.001
	SL-FFNN	0.192 ± 0.004	0.726 ± 0.012	0.859 ± 0.008	0.250 ± 0.009	0.141 ± 0.003	0.174 ± 0.004	0.779 ± 0.009	0.887 ± 0.004	0.221 ± 0.002	0.127 ± 0.002
	SVR	0.176 ± 0.000	0.769 ± 0.001	0.882 ± 0.002	0.264 ± 0.004	0.137 ± 0.001	0.166 ± 0.000	0.799 ± 0.002	0.897 ± 0.001	0.251 ± 0.001	0.130 ± 0.000
	RF	0.167 ± 0.000	0.793 ± 0.002	0.891 ± 0.001	0.227 ± 0.001	0.124 ± 0.000	0.167 ± 0.001	0.798 ± 0.003	0.894 ± 0.002	0.223 ± 0.001	0.123 ± 0.001
	MLR	0.195 ± 0.001	0.718 ± 0.003	0.848 ± 0.002	0.269 ± 0.001	0.148 ± 0.001	0.195 ± 0.002	0.724 ± 0.003	0.851 ± 0.002	0.266 ± 0.002	0.147 ± 0.001
9h-ahead	LSTM	0.163 ± 0.001	0.807 ± 0.003	0.900 ± 0.002	0.215 ± 0.002	0.120 ± 0.000	0.151 ± 0.003	0.833 ± 0.006	0.917 ± 0.001	0.199 ± 0.003	0.111 ± 0.001
	P-LSTM	0.157 ± 0.001	0.817 ± 0.005	0.906 ± 0.003	0.208 ± 0.004	0.116 ± 0.001	**0.145** ± 0.001	**0.846** ± 0.001	**0.922** ± 0.000	**0.191** ± 0.000	**0.106** ± 0.001
	FFNN	0.200 ± 0.001	0.711 ± 0.002	0.845 ± 0.002	0.252 ± 0.001	0.145 ± 0.001	0.182 ± 0.001	0.755 ± 0.006	0.870 ± 0.003	0.236 ± 0.001	0.133 ± 0.001
	SL-FFNN	0.231 ± 0.003	0.612 ± 0.007	0.803 ± 0.003	0.283 ± 0.002	0.167 ± 0.002	0.202 ± 0.004	0.701 ± 0.018	0.840 ± 0.010	0.253 ± 0.009	0.146 ± 0.004
	SVR	0.204 ± 0.000	0.697 ± 0.003	0.839 ± 0.002	0.292 ± 0.003	0.157 ± 0.000	0.189 ± 0.001	0.737 ± 0.006	0.862 ± 0.004	0.279 ± 0.002	0.147 ± 0.001
	RF	0.193 ± 0.001	0.728 ± 0.005	0.854 ± 0.003	0.257 ± 0.002	0.143 ± 0.000	0.191 ± 0.001	0.733 ± 0.006	0.857 ± 0.003	0.256 ± 0.002	0.142 ± 0.000
	MLR	0.219 ± 0.001	0.653 ± 0.005	0.808 ± 0.004	0.295 ± 0.001	0.166 ± 0.001	0.215 ± 0.001	0.659 ± 0.006	0.812 ± 0.004	0.294 ± 0.000	0.164 ± 0.000
12h-ahead	LSTM	0.182 ± 0.002	0.759 ± 0.001	0.873 ± 0.002	0.239 ± 0.004	0.134 ± 0.000	0.167 ± 0.001	0.793 ± 0.004	0.895 ± 0.002	0.220 ± 0.002	0.124 ± 0.001
	P-LSTM	0.179 ± 0.002	0.766 ± 0.007	0.877 ± 0.004	0.232 ± 0.002	0.131 ± 0.001	**0.165** ± 0.006	**0.798** ± 0.010	**0.901** ± 0.004	**0.210** ± 0.002	**0.121** ± 0.003
	FFNN	0.225 ± 0.001	0.631 ± 0.006	0.799 ± 0.003	0.272 ± 0.005	0.162 ± 0.001	0.207 ± 0.003	0.683 ± 0.011	0.829 ± 0.007	0.259 ± 0.005	0.151 ± 0.002
	SL-FFNN	0.262 ± 0.007	0.502 ± 0.034	0.747 ± 0.012	0.315 ± 0.007	0.189 ± 0.004	0.232 ± 0.001	0.605 ± 0.009	0.789 ± 0.009	0.281 ± 0.004	0.167 ± 0.001
	SVR	0.228 ± 0.001	0.623 ± 0.006	0.793 ± 0.004	0.319 ± 0.000	0.175 ± 0.001	0.209 ± 0.001	0.678 ± 0.007	0.825 ± 0.004	0.295 ± 0.003	0.160 ± 0.001
	RF	0.214 ± 0.002	0.668 ± 0.004	0.818 ± 0.002	0.285 ± 0.001	0.160 ± 0.000	0.212 ± 0.002	0.670 ± 0.006	0.819 ± 0.004	0.279 ± 0.003	0.157 ± 0.001
	MLR	0.233 ± 0.002	0.608 ± 0.007	0.779 ± 0.005	0.311 ± 0.002	0.176 ± 0.001	0.231 ± 0.002	0.607 ± 0.004	0.779 ± 0.003	0.309 ± 0.003	0.175 ± 0.001
24h-ahead	LSTM	0.243 ± 0.000	0.578 ± 0.005	0.762 ± 0.004	0.307 ± 0.002	0.178 ± 0.001	0.221 ± 0.002	0.648 ± 0.007	0.808 ± 0.001	0.281 ± 0.008	0.162 ± 0.000
	P-LSTM	0.235 ± 0.003	0.606 ± 0.012	0.783 ± 0.005	0.294 ± 0.001	0.173 ± 0.002	**0.211** ± 0.002	**0.681** ± 0.006	**0.829** ± 0.004	**0.267** ± 0.002	**0.154** ± 0.001
	FFNN	0.302 ± 0.002	0.345 ± 0.011	0.611 ± 0.005	0.355 ± 0.004	0.221 ± 0.003	0.294 ± 0.027	0.446 ± 0.017	0.675 ± 0.011	0.337 ± 0.006	0.203 ± 0.003
	SL-FFNN	0.287 ± 0.001	0.418 ± 0.011	0.653 ± 0.005	0.360 ± 0.008	0.214 ± 0.003	0.283 ± 0.004	0.437 ± 0.006	0.667 ± 0.002	0.361 ± 0.012	0.211 ± 0.004
	SVR	0.295 ± 0.001	0.376 ± 0.003	0.624 ± 0.001	0.381 ± 0.003	0.224 ± 0.000	0.269 ± 0.003	0.477 ± 0.013	0.693 ± 0.009	0.359 ± 0.001	0.205 ± 0.002
	RF	0.276 ± 0.000	0.457 ± 0.005	0.676 ± 0.004	0.359 ± 0.003	0.208 ± 0.001	0.268 ± 0.002	0.481 ± 0.010	0.694 ± 0.008	0.351 ± 0.001	0.202 ± 0.001
	MLR	0.289 ± 0.002	0.401 ± 0.003	0.633 ± 0.002	0.375 ± 0.002	0.221 ± 0.000	0.284 ± 0.003	0.416 ± 0.011	0.645 ± 0.009	0.371 ± 0.001	0.217 ± 0.002
48h-ahead	LSTM	0.283 ± 0.001	0.414 ± 0.007	0.645 ± 0.004	0.368 ± 0.004	0.215 ± 0.000	0.258 ± 0.003	0.514 ± 0.009	0.720 ± 0.006	0.337 ± 0.007	0.194 ± 0.002
	P-LSTM	0.276 ± 0.002	0.454 ± 0.008	0.675 ± 0.006	0.359 ± 0.004	0.209 ± 0.001	**0.249** ± 0.002	**0.552** ± 0.006	**0.746** ± 0.006	**0.325** ± 0.003	**0.186** ± 0.001
	FFNN	0.332 ± 0.004	0.205 ± 0.006	0.468 ± 0.003	0.413 ± 0.009	0.252 ± 0.002	0.316 ± 0.001	0.288 ± 0.008	0.542 ± 0.003	0.399 ± 0.009	0.240 ± 0.002
	SL-FFNN	0.335 ± 0.005	0.185 ± 0.012	0.457 ± 0.015	0.421 ± 0.012	0.256 ± 0.003	0.319 ± 0.004	0.265 ± 0.008	0.533 ± 0.007	0.388 ± 0.008	0.240 ± 0.002
	SVR	0.338 ± 0.003	0.165 ± 0.012	0.447 ± 0.010	0.435 ± 0.003	0.262 ± 0.002	0.311 ± 0.001	0.292 ± 0.005	0.547 ± 0.005	0.409 ± 0.002	0.241 ± 0.001
	RF	0.312 ± 0.000	0.287 ± 0.003	0.537 ± 0.003	0.411 ± 0.003	0.241 ± 0.000	0.302 ± 0.001	0.328 ± 0.005	0.575 ± 0.005	0.401 ± 0.002	0.234 ± 0.000
	MLR	0.325 ± 0.001	0.226 ± 0.006	0.476 ± 0.006	0.425 ± 0.003	0.253 ± 0.000	0.321 ± 0.001	0.242 ± 0.010	0.494 ± 0.009	0.418 ± 0.002	0.249 ± 0.002

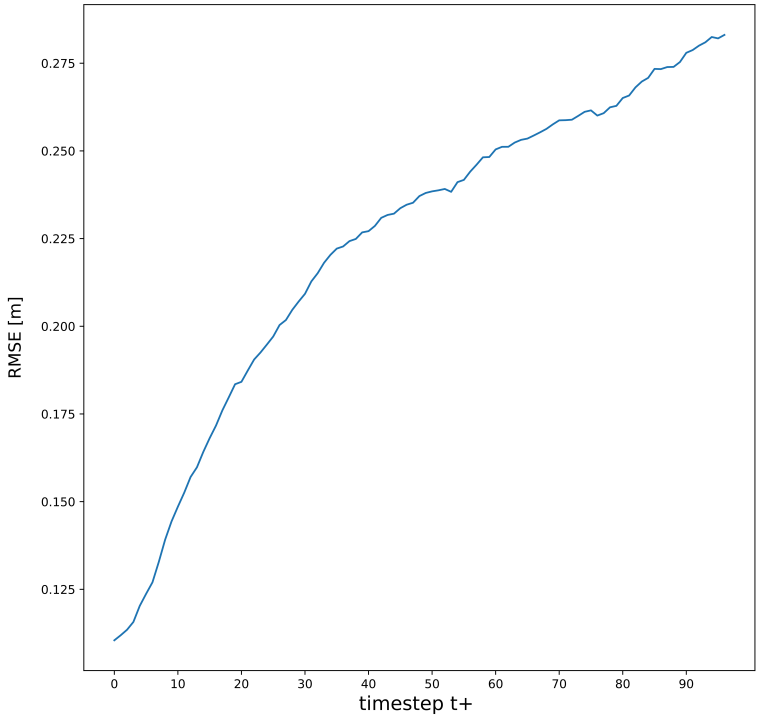

Figure 3.8: *RMSE of the LSTM for the different prediction intervals of the SWH excluded the bathymetries.*

The overall poor performance of the MLR can be explained by considering only linear relations, while the data showed a much more complex and non-linear pattern. Since the ETD sandbank movements and their non-linear influence on the sea state at position Coast II were not considered in this setting, a poor performance of the MLR was expected (Jörges et al., 2021a).
As the variability or standard deviation of the different model runs error metrics was mostly smaller by more than a factor of 100 than the mean value (Tab. 3.4, Tab. 3.5), it is sufficient to consider the mean value only, while variations can be neglected. The SL-FFNN showed the highest variability, especially for longer forecast periods. All models, except the SL-FFNN, were robust against the random train-test split of the dataset as well as the random initialization of particular parameters in the algorithms.

3.3.1.3 Long-term prediction

Since the SL-FFNN showed an extreme overfitting in both long-term tasks and the FFNN in the 48 h-ahead predictions when training over 100 epochs, the number of epochs for the long-term prediction of these methods was reduced to 10. Results of the long-term predictions are given in Tab. 3.5.

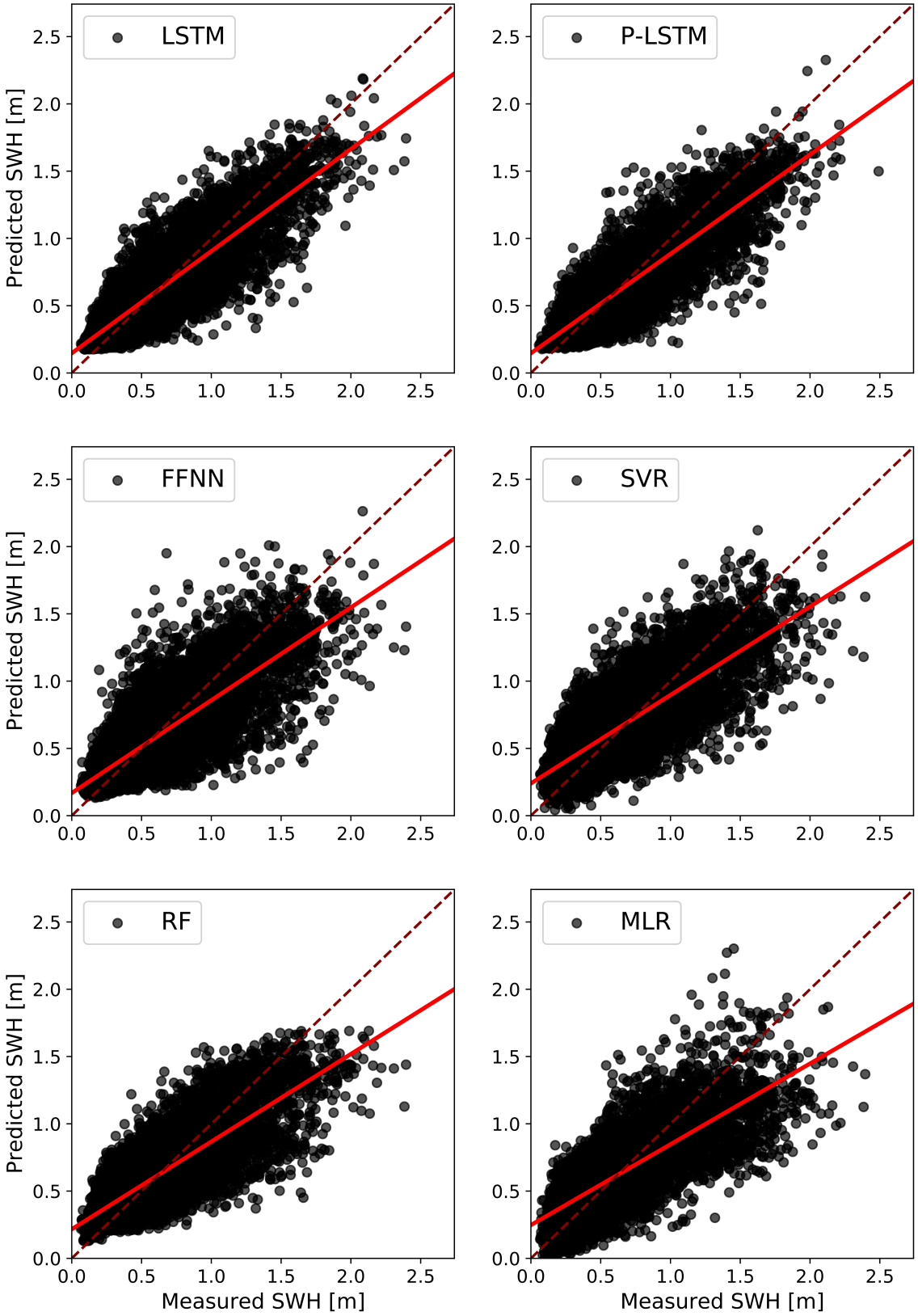

Figure 3.9: *Comparison between the measured and predicted SWH for the 12 h-ahead prediction excluded the bathymetries for the LSTM, P-LSTM, FFNN, SVR, RF, and MLR on the test data.*

Compared to the SWH range of 2.97 m at position Coast II (Tab. 3.1), the MAEs of 0.173 m at 24 h-ahead and 0.209 m at 48 h-ahead using the P-LSTM are still acceptable. The MAAPEs for the LSTM (0.368) and P-LSTM (0.359) for the 48 h-ahead prediction indicate that the LSTMs are capable of long-term prediction. Values are similar to the experimental results on some deep ocean buoy positions in Pirhooshyaran et al. (2020). As seen from Tab. 3.5, the differences between the LSTMs and the benchmark methods are significant, while the differences among the benchmark methods are negligible.

3.3.2 Model comparison – validation on offshore data

3.3.2.1 NOAA dataset

Validation of the proposed LSTM and P-LSTM method was performed on a well-studied deep ocean buoy dataset from the National Oceanic and Atmospheric Administration (NOAA) (`https://www.ndbc.noaa.gov`). The three adjacent offshore buoys, 46042, 46069, and 46025, are located in the Pacific Ocean off the U.S. west coast (supplementary material, App. II.1). Hourly SWH data from 2009 to 2010 were used for training and testing, respectively. For detailed information on the buoys parameters, refer to App. II.1 of the supplementary material. As in previous studies (Cornejo-Bueno et al., 2018; Londhe, 2008; Londhe and Panchang, 2007; Pirhooshyaran et al., 2020), we reconstructed the SWH of buoy 46069 at time t using information of SWH for the last 12 h ($k = 12$) to t from the two nearby buoys 46042 and 46025. The neural networks only received information about the SWH of the two adjacent buoys for the reconstructed time step t, while some of the benchmark methods used additional features such as wind speed, average wave period, air temperature, atmospheric pressure, and longer lookback times. Missing data were deleted completely, resulting in a total of 3 728 data points for training and 4 201 for testing. The architecture of the LSTM and P-LSTM was optimized for the NOAA dataset using Bayesian hyperparameter optimization. One LSTM layer with 125 neurons, relu activation function, and a final fully-connected layer was trained over 30 epochs.

3.3.2.2 Experimental results on NOAA dataset

Results of the reconstruction task with our proposed methods and the performance of benchmark methods from the literature are given in Tab. 3.6. Compared to Cornejo-Bueno et al. (2018) and Pirhooshyaran et al. (2020), our proposed method, original designed for the nearshore area in the East Frisian North Sea, also performed well on deeper and more offshore buoy positions. With a MAE of 0.2327 m, the

proposed LSTM model achieved the smallest error. While the LSTM had a smaller RMSE, the P-LSTM had a slightly higher R^2. Overall, the performance of the LSTM and P-LSTM was very similar to the performance of the Seqtoseq and LSTM models in Pirhooshyaran et al. (2020). In addition to the reconstruction of the SWH, we also performed forecasting for the deep ocean data (Tab. 3.6). While the 3 h-ahead prediction came to a MAE of 0.239 m, the 12 h-ahead prediction had a MAE of 0.3289 m.

3.3.3 Impact of bathymetry features

3.3.3.1 Reconstruction including bathymetry

Including bathymetric data to the training purpose significantly improved the performance of our proposed LSTM frameworks and all other models on the test data. Tab. 3.4 shows the results for the reconstruction with included bathymetric data. The mean ensemble RMSE of the LSTM and P-LSTM was reduced by 16.7% to 0.070 m and by 12.7% to 0.069 m with included bathymetries, respectively. The P-LSTM setting achieved the best accuracy for all experiments and in terms of all error metrics. Therefore, the SWH can be predicted with a MAE of about 5 cm, although the model does not have any historical data on the predicted position south of the ETD sandbanks. The reconstruction error of the FFNN decreased by 20.6% to an RMSE of 0.085 m by including bathymetry features, which was the greatest decrease of all models. Including bathymetry features decreased the RMSE of the SL-FFNN by 19.2% to 0.101 m, of the SVR by 6.6% to 0.114 m, of the RF by 5.5% to 0.104 m, and of the MLR by 2.0% to 0.149 m.

The performance of the FFNN was better or equal than the performance of the SVR and RF for the reconstruction excluding and including bathymetries. Pirhooshyaran et al. (2020) mentioned the superior performance of the LSTM for SWH reconstruction, while they also had a very good performance using RF in contrast to our results. It turned out, that the neural networks could use the additional information on bathymetry best to increase their performance.

The performances of the P-LSTM excluding and including the bathymetries for reconstructing SWH are compared in Fig. 3.10. The data points were significantly closer to the bisector when the bathymetry was included and most extreme values were reconstructed satisfactory. The improvements in all error metrics also demonstrate this. For example, MAAPE decreased by 11.4% and R^2 increased by 1.3% for the P-LSTM. The relative improvements in R^2 and r were always below those of the other error metrics. Even by including bathymetric data, the measured values were underestimated on average – although not as much – as can be seen from the solid red line in Fig. 3.10.

Table 3.6: *Reconstruction and short-term forecasting significant wave height H_{m_0} performance results on NOAA buoy dataset.*

Problem task	Reference	Model	RMSE [m]	R^2	r	MAAPE	MAE [m]
Reconstruction	Cornejo-Bueno et al. (2018)	All features-ELM	0.4653	0.6624	–	–	0.3582
		All features-SVR	0.6519	0.3949	–	–	0.4986
		GGA-ELM-ELM	0.3650	0.7049	–	–	0.2858
		GGA-ELM-SVR	0.3599	0.7056	–	–	0.2727
		BO-GGA-ELM-ELM	0.3324	0.7429	–	–	0.2519
		BO-GGA-ELM-SVR	0.3331	0.7396	–	–	0.2461
		3-layered FCN	0.3960	0.7503	–	–	0.2913
	Pirhooshyaran et al. (2020)	Seqtoseq (Adam)	0.3419	**0.8106**	–	–	0.2528
		LSTM (Adam)	**0.2784**	0.7922	–	–	0.2392
	This study	LSTM	0.3096	0.7979	0.8946	**0.1278**	**0.2327**
		P-LSTM	0.3303	0.8034	**0.8971**	**0.1278**	0.2448
3h-ahead		LSTM	**0.3167**	0.7906	**0.8928**	0.1329	**0.2390**
		P-LSTM	0.3397	**0.7911**	0.8907	**0.1314**	0.2496
6h-ahead		LSTM	**0.3470**	0.7527	0.8723	0.1411	**0.2577**
		P-LSTM	0.3684	**0.7606**	**0.8730**	**0.1372**	0.2721
9h-ahead		LSTM	**0.3989**	**0.6887**	**0.8342**	**0.1583**	**0.2933**
		P-LSTM	0.4302	0.6748	0.8242	0.1622	0.3172
12h-ahead		LSTM	**0.4502**	0.6026	0.7887	0.1758	**0.3289**
		P-LSTM	0.4617	**0.6378**	**0.8024**	**0.1680**	0.3364

Differences in the predicted time series between the P-LSTM excluding and including the bathymetries for the reconstruction are illustrated in Fig. 3.11. Although both lines are generally close to each other, differences were found especially at local extremes, while extreme values were regularly underestimated.

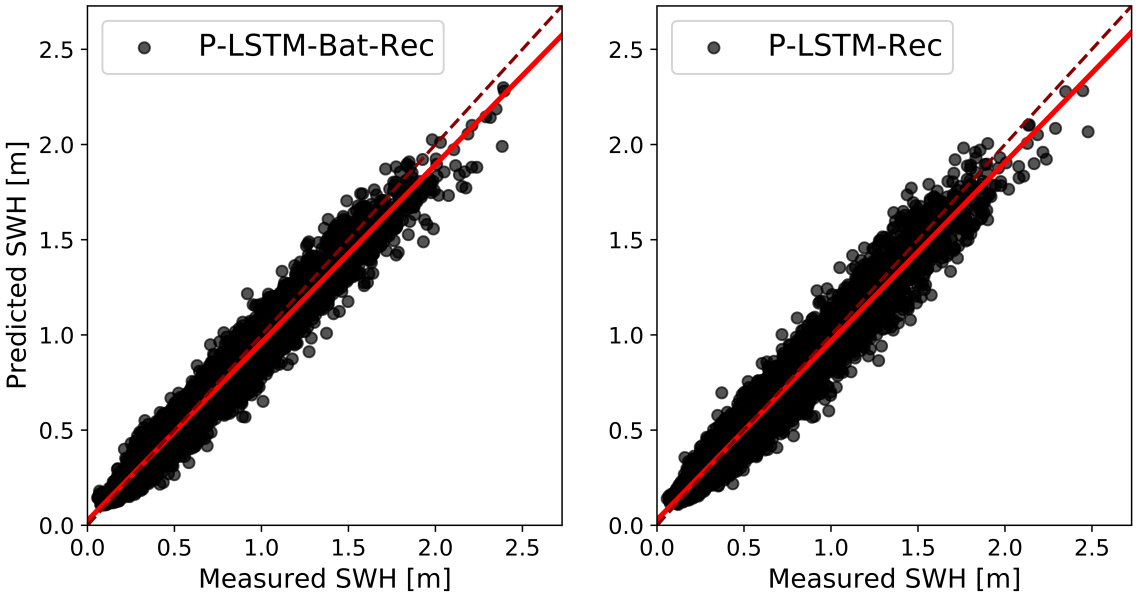

Figure 3.10: *Comparison between the measured and predicted SWH for the reconstruction with included (left) and excluded (right) bathymetries for the P-LSTM on the test data.*

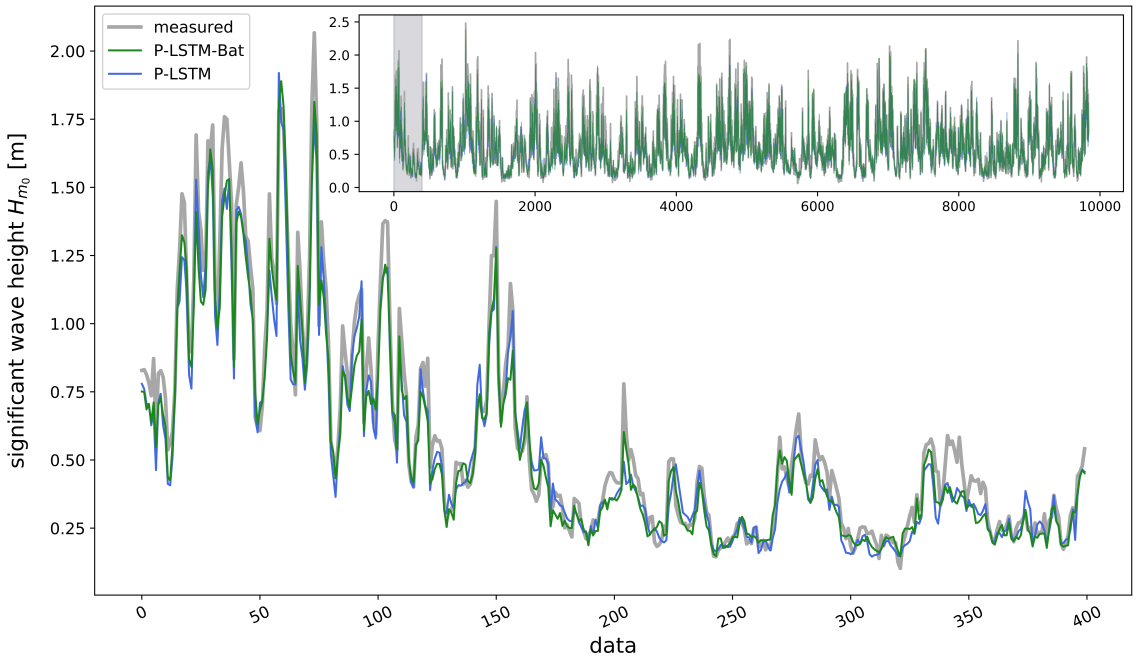

Figure 3.11: *Time series of SWH reconstruction with included (green) and excluded (blue) bathymetries for the P-LSTM. The inset shows the entire test dataset.*

3.3.3.2 Prediction including bathymetry

By including the bathymetric data to the 3, 6, 9, and 12 h-ahead prediction of the LSTM and P-LSTM, the relative reduction of the RMSE was about similar for both methods and all time steps, resulting in a reduction of the RMSEs by 7.6–9.5% for the P-LSTM and 7.4–11.7% for the LSTM. Thus, the RMSE decreased to 0.096 m, 0.124 m, 0.145 m, and 0.165 m for the P-LSTM, respectively (Tab. 3.5). Valid 12 h-ahead predictions were still made with the proposed models. This time span can be sufficient to initiate practical solutions, e.g., for coastal protection. It underlines the good performance of neural networks, especially the recurrent LSTM neural networks, to accurately predict complex time series data with morphodynamic influences.

Among the benchmark methods, the FFNN significantly performed best in terms of RMSE and MAAPE. It improved the RMSE by 11.2% for the 3 h-ahead prediction and by 8% for the 12 h-ahead prediction by including bathymetric data. The SVR, RF, and MLR could reduce the RMSE by only up to 8.3%, 2.9%, and 1.8% for the short-term predictions, respectively. While the performance of all error metrics of the FFNN was similar or worse than that of the RF by excluding the bathymetries, it was better by including the bathymetries for the short- term predictions (Tab. 3.5).

As seen in Fig. 3.12, the data points of the 3 h-ahead prediction with included bathymetries were closer around the bisector. Even with the prediction of extreme events there was no increase in the prediction error of the P-LSTM. With the inclusion of the bathymetries, the prediction error of extreme events was lower in the 3 and 6 h-ahead forecast of the P-LSTM. By increasing the forecasting interval to 6 h, the prediction error became larger, as indicated by the more developed bulge around the bisector.

Fig. 3.13 illustrates the comparison between the measured and predicted values of the P-LSTM including and excluding the bathymetry features for the 9 and 12 h-ahead prediction. There is a distinctly greater bulge around the bisector as well as a stronger spread of certain predicted values. Nevertheless, with bathymetry included, the data points were closer to the bisector even at these prediction intervals.

The LSTM gate units allow the recurrent model to remember long periods while selecting only essential information (Xiao et al., 2019). Without these morphodynamic information, the models tended to underestimate extreme values (Fig. 3.12 and Fig. 3.13).

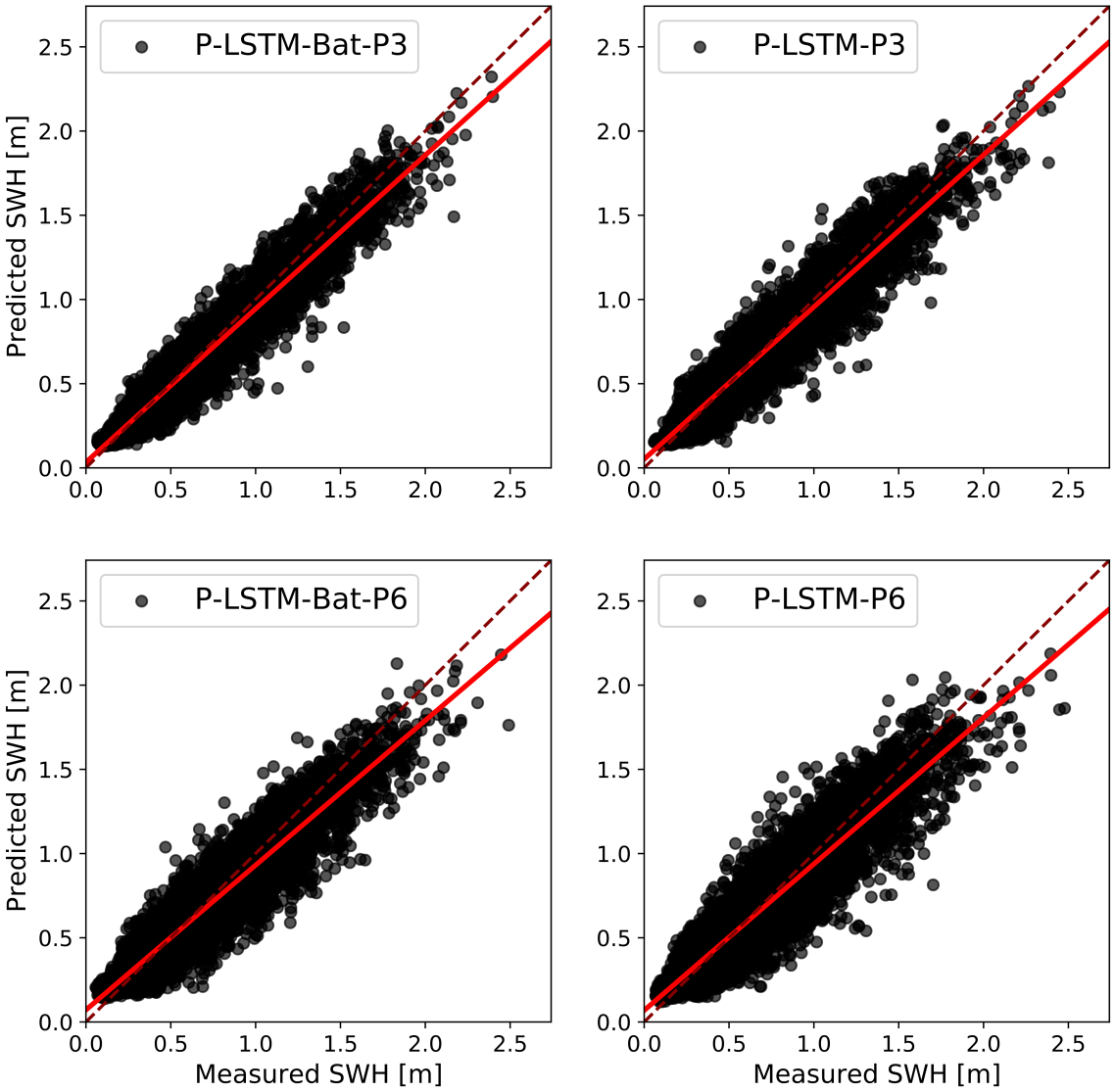

Figure 3.12: *Comparison between the measured and predicted SWH for the 3 h-ahead (top) and 6 h-ahead (bottom) prediction with included (left) and excluded (right) bathymetries for the P-LSTM on the test data.*

The 24 and 48 h-ahead long-term predictions of SWH were also improved by including bathymetric data (Tab. 3.5). RMSE and MAAPE of the P-LSTM were reduced by 10.2 and 9.2% for the 24 h-ahead prediction and by 9.8 and 9.5% for the 48 h-ahead prediction, respectively. As a result, the long-term predictions were improved by a greater percentage in terms of RMSE compared to the short-term predictions by the P-LSTM with bathymetry features.

In general, all models in the nearshore area improved their performance by including bathymetric features, hence they are highly relevant in reconstructing and predicting SWH in the study area.

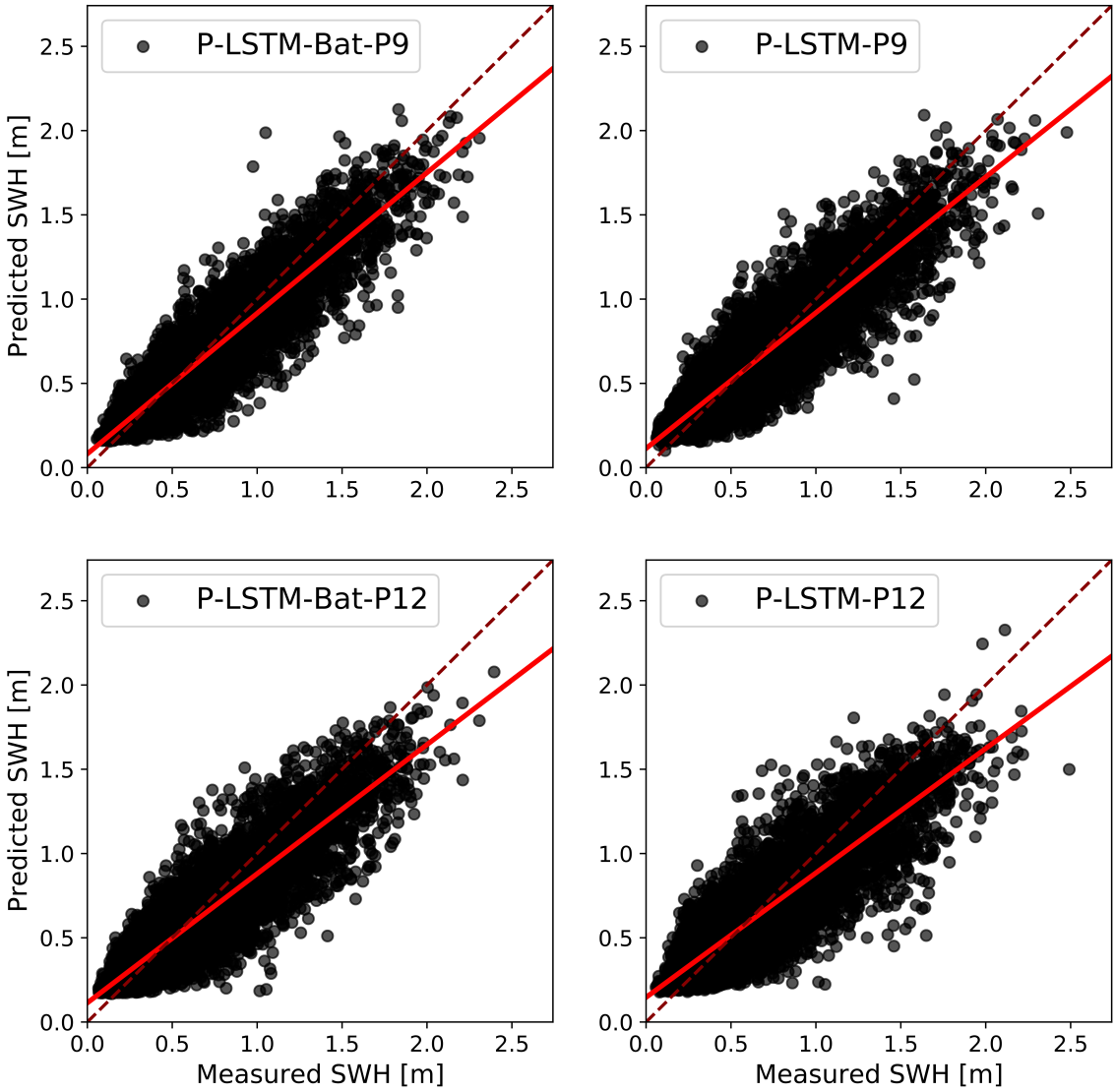

Figure 3.13: *Comparison between the measured and predicted SWH for the 6 h-ahead (top) and 12 h-ahead (bottom) prediction with included (left) and excluded (right) bathymetries for the P-LSTM on the test data.*

3.3.3.3 Limitations of including bathymetric data

The train and test split was randomly selected, thus the models and their performances on the test data cannot be transferred or generalized to other areas, especially other bathymetries. For a better generalization, more bathymetric data are necessary for training and learning the ETD influence. However, it can be assumed that the LSTM is a suitable method to process even more complex morphodynamic data. In order to process 2-dimensional bathymetric data a combined LSTM and convolutional neural network (CNN) could be promising. Choi et al. (2020) used such a framework in a similar context with ocean images instead of bathymetric data. In addition, a MLR, which tends to underfitting and leads to a greater gener-

alization due to its low number of weights, consistently showed an improvement by adding bathymetric data, indicating a general improvement.
Overall, the bathymetric data have improved the reconstruction and prediction accuracy. The results showed that the ETD sandbanks have a significant influence on the propagation of the waves and the wave height on the island coast, which was also found by numerical modelling (Jörges et al., 2021a). As the size and shape of the ETD sandbanks are constantly changing by morphodynamic effects, their influence on local forecast models must be considered in addition to continuous monitoring.

3.3.4 Further feature selection analysis

While adding bathymetric features significantly improved the performance of all models, one should also investigate the influence of other features. It must be noted that the best performance was observed with the inclusion of all the above features, hence they all contain significant information for the prediction.

3.3.4.1 Wind speed and direction

Even though theoretically the greatest influence is expected from wind speed, by including this feature the RMSE of the reconstruction was reduced by only 11.4% for the LSTM and 14.8% for the P-LSTM (supplementary material, App. II.2). Thus, the influence of bathymetric features (-16.7% for LSTM and -12.7% for P-LSTM in terms of RMSE) was about similar to the influence of wind speed. This may be due to the fact that the SWH data of the two other buoy positions are affected significantly by wind speed and thus a fraction of the wind speed information is redundant. The influence of the bathymetry, however, only becomes effective south of the two other buoy positions, and is not already hidden within them in any way. By reducing the RMSE of the reconstruction by 21.3% for the LSTM and 21.6% for the P-LSTM, the influence of wind direction was slightly higher than for wind speed and bathymetries (supplementary material, App. II.3). The wind direction influences the non-linear wave breaking at the ETD sandbanks as well as fetch length and wave direction.

3.3.4.2 Water level

With feature selection excluding only water level and tidal information, the best RMSE and R^2 were 0.132 m and 0.873 for reconstruction with the P-LSTM, respectively (supplementary material, App. II.4). Since the reduction in RMSE for the reconstruction was 47.8% for the LSTM and 47.7% for the P-LSTM, the influence of

the water level was greater than any other single feature, which can be explained by the complex, nearshore location of the buoys where tides dominate the wave climate.

3.3.4.3 SWH only – nearshore vs. offshore performance

Considering only SWH for reconstruction and forecast resulted in the worst prediction performance with a reconstruction RMSE of 0.175 m and 0.173 m for the LSTM and P-LSTM, respectively (supplementary material, App. II.5). While the differences between the models in terms of all error metrics are huge in some cases, the performances in this case, except for MLR, were very similar. As the regression problem became simpler with less parameters, the LSTM methods, which are optimized for complex problems, may have been overfitted. Thus, not only LSTM and P-LSTM had the best performance, but also (SL-)FFNN and RF performed quite well in this case. The best reconstruction performance on the Norderney nearshore dataset was $R^2 = 0.781$ for the FFNN and MAAPE = 0.2 for the P-LSTM.
Compared to the performance on the deep ocean NOAA dataset, the scale-independent error metrics were worse in the nearshore area, since the best performance on the offshore dataset was $R^2 = 0.8106$. Although this is a small difference, it must be noted that the distance between the buoy positions is much smaller in the nearshore area. This emphasizes that considering SWH only in the nearshore area is not a proper method and other features must be considered since the sea state is much more complex. Bathymetry – mainly represented by the ETD sandbanks for the research area of this study – is a main feature in the nearshore area.

3.4 Conclusions

This study aimed to reconstruct and multi-step forecast SWH of the nearshore area of Norderney, Germany based on neighboring moored buoy measurements, with the focus set on the model improvements by including bathymetric data on a local scale. Exact knowledge of nearshore wave conditions – considering natural sea state damping by the morphodynamically changing ETD sandbanks – is essential for coastal protection, planning offshore missions, and running sea state and climate forecasting models.
It was found in our experiments, that the LSTM and P-LSTM neural networks performed best among other existing benchmark methods such as FFNN, SVR, RF, and MLR both for reconstruction and prediction. Compared to the LSTM, the RMSE of the FFNN was at least 21%, the SVR 25%, the RF 18%, and the MLR

28% higher for the reconstruction and short-term prediction task. To achieve reliable and fair results, hyperparameters were tuned using Bayesian optimization. For all models and reconstruction or prediction intervals an ensemble of three independent model runs was performed. All models showed robust results on the test data with respect to the various training sets.

By including bathymetry features, the performance of the proposed LSTM and P-LSTM method in terms of RMSE increased by 16.7% to an RMSE of 0.07 m and by 12.7% to an RMSE of 0.069 m for the reconstruction, respectively. The short-term predictions 3, 6, 9, and 12 h- ahead showed significantly better performance in terms of all error metrics used – namely RMSE, R^2, r, MAAPE, and MAE – by including the bathymetries. Compared to other feature selections, the improvements were considerable. Thus, our findings highlight the broad potential of LSTM neural networks for SWH reconstruction and prediction as well as the need to consider bathymetries in morphodynamic changing coastal areas, such as the East Frisian Islands.

Further experiments on well-studied deep ocean wave data also showed satisfactory results, so the proposed method is not limited to the nearshore area.

The LSTM and P-LSTM neural networks could satisfactorily predict even extreme values of SWH using the bathymetric data, those events that are of special interest for coastal protection and marine operations. Limitations of the presented study are the insufficient availability of bathymetries for training the models. Nevertheless, the presented methods can be used for other data and regions as well. Using high-temporal-resolution measurements of a coastal radar may be a promising future way. Due to the accurate reconstruction and prediction of the proposed model it might also be possible to place the expensive ocean buoys more efficiently or even to drop positions completely. Results can also be used to verify numerical models and vice versa. Using additional information regarding the local wave climate such as wave period and wave direction might also be considered in future studies.

4 Research Study III

Spatial ocean wave height prediction with CNN mixed-data deep neural networks using random field simulated bathymetry

Co-Authors: Cordula Berkenbrink, Hanno Gottschalk & Britta Stumpe

Published in Ocean Engineering (2023), 271 (accepted 10.01.2023)

Abstract

With climate change impacts like sea level rise and changing storms, proper prediction of significant wave height (SWH) becomes increasingly important for coastal protection and marine disaster prevention. In the coastal areas of the North Sea, the morphodynamically changing ebb-tidal delta (ETD) sandbanks cause non-linear wave propagation. Therefore, consideration of spatial dependencies using bathymetric data is essential for accurate machine learning predictions. We developed a novel two-dimensional mixed-data deep convolutional neural network (CNN) for spatial SWH prediction in the nearshore area of Norderney, Germany. To overcome the problem of limited bathymetry data, dynamic ETD sandbank morphologies were simulated using random fields and used as model input for the first time. The regional mixed-data CNN was also trained and tested with in-situ metocean input data from 2004–2017 and SWAN-modeled ground truth wave fields as output. The proposed CNN architecture outperformed other benchmark models on the test data ($RMSE = 0.097m$, $R^2 = 0.977$, $MAAPE = 6.7\%$). Further validation on 59 buoy measurements revealed very similar accuracy of the CNN and SWAN. Compared to the commonly used numerical SWAN model, the trained CNN reduced the computational cost by a factor of more than 300 000, making it an efficient surrogate predictive model.

Keywords: convolutional neural network (CNN), machine learning, wave height forecasting, dynamic bathymetry, SWAN surrogate modeling, random fields simulation

4.1 Introduction

Proper predictions of nearshore significant wave height (SWH) are crucial for coastal protection, ocean engineering, offshore operations, and marine disaster prevention (Deo and Naidu, 1999; Teich et al., 2018; Ti et al., 2018). Especially under consideration of climate change impacts like sea level rise and changes in storm characteristics, challenging storm surges become an even worse risk for people living in coastal regions like the North Sea (Bitner-Gregersen et al., 2018; Grabemann and Weisse, 2008; Mori et al., 2013; Weisse et al., 2012).

Common methods for predicting wave heights are third-generation numerical wave models based on the spectral action balance equation such as Simulating Waves Nearshore (SWAN) (Booij et al., 1999), WAM (The Wamdi Group, 1988), and WAVEWATCH III (Tolman et al., 2002). Although numerical models are precise due to their physical representation of wave propagation, they require immense computational resources, up to high-performance computers for large computational domains. Caused by the high computational cost and time requirement, they are not suitable for rapid investigations like sea rescue or storm surge impacts (Hu et al., 2021).

Alternatively to numerical modeling, numerous recent studies highlighted promising results in wave height forecasting using machine learning techniques (Abed-Elmdoust and Kerachian, 2012; Mahjoobi et al., 2008; Malekmohamadi et al., 2011; Mandal and Prabaharan, 2010). Artificial neural networks (ANN) are widely used as a soft-computing method for wave height prediction and reconstruction (Ali and Prasad, 2019; Deo and Naidu, 1999; Deo et al., 2001; Dixit and Londhe, 2016; Durán-Rosal et al., 2016; Kumar et al., 2017; Law et al., 2020; Londhe, 2008; Londhe and Panchang, 2007; Peres et al., 2015; Zamani et al., 2008). Pirhooshyaran et al. (2020) applied a recurrent long short-term memory (LSTM) neural network to forecast time series of wave heights. Fan et al. (2020) combined the LSTM neural network with the numerical wave model SWAN and improved the accuracy of the standard SWAN model by more than 65% for single-point predictions at a wave buoy position. In other studies like Malekmohamadi et al. (2008) and Puscasu (2014), coupling of ANN and numerical modeling for wave prediction has also been successfully performed and yielded a significant speed-up in the prediction computation. O'Donncha et al. (2018) integrated a physics-based SWAN model with machine learning for an ensemble prediction, which outperformed either technique in isolation.

Despite the current popularity of machine learning methods for ocean wave height prediction and reconstruction, most studies focused on single-point predictions, resulting in insufficient consideration of spatial dependencies of surrounding waves and

the underlying bathymetry, compared to two-dimensional SWAN numerical modeling. Only a few studies have already developed two-dimensional machine learning surrogate models for the SWAN numerical wave model. James et al. (2018) designed a feed-forward neural network (FFNN) trained with ocean currents, winds, and SWAN-modeled SWH and wave period of the Monterey Bay area. An RMSE of 0.09 m was achieved for the spatial SWH prediction. Their model required less than 1/1000 of the computational time of the physics-based SWAN model. Feng et al. (2020) predicted SWH and peak period of Lake Michigan with an FFNN using SWAN-modeled input data from 2005–2014 and not only achieved comparable performance to SWAN, but also reduced computation time by a factor of up to 20 000. Chen et al. (2021b) designed a random forest surrogate model to predict nearshore spatial wave fields modeled by SWAN based on buoy data at single locations. Validation against in-situ buoy data showed better performance of the surrogate model ($R^2 = 0.91$) than the corresponding SWAN model ($R^2 = 0.85$) and reduced computation time by a factor of 100. Zhou et al. (2021) presented a two-dimensional convolutional LSTM neural network for forecasting SWH in the South and East China Seas with lead times up to 24 h, trained on WAVEWATCH III data. For the 24 h forecast, a mean absolute percentage error of 61% was achieved. Convolutional neural networks (CNN) are very efficient on two-dimensional data like images since they can learn local patterns in the data (LeCun et al., 1989). Besides applications in image classification, face and speech recognition, and natural language processing, Choi et al. (2020) used a CNN for real-time ocean wave height prediction based on raw ocean images in-situ captured on wave buoys. Wei and Davison (2022) created a CNN model to predict waves and hydrodynamics in the nearshore area, using SWASH-modeled data as input. They found that the CNN accurately predicted wave propagation and wave breaking over a sandbar. Bai et al. (2022) developed a two-dimensional wave field prediction model based on CNN for SWH forecasts in the South China Sea. Forecasts using ERA5 reanalysis data with lead times from 12–72 h were performed, achieving mean absolute percentage errors between 8.55% and 19.48%, respectively.

Although some ANN surrogate models for two-dimensional prediction of wave height already exist, none of the above mentioned studies have yet considered bathymetry as input. Particularly in coastal areas, SWH forecasting can be very challenging due to complex non-linear nearshore processes of wave propagation and interaction. Thus, it was found that the morphodynamic changes of the ebb-tidal delta (ETD) sandbanks off the coast of Norderney, Germany significantly affect the local wave climate (Jörges et al., 2021a). Driven by the ebb-tidal currents or even single strong

storm surges, the ETD sandbanks are subject to constant and rapid morphological changes by migration, erosion, and accumulation (Castelle et al., 2007; Dallas and Barnard, 2011; Wang et al., 2012). Since the ETD sandbanks provide natural coastal protection through sea state damping, especially for extreme events, their extent is essential for predicting nearshore SWH (Jörges et al., 2021a; Niemeyer, 1979; Niemeyer and Kaiser, 1999). Therefore, Jörges et al. (2021b) improved the single-point prediction and reconstruction of nearshore SWH by including a transect of ETD sandbank bathymetry features to LSTM neural networks for the first time. Since high temporal and spatial resolution bathymetry data are limited and hydrographic surveys are very cost-intensive (Spicer et al., 2019), no two-dimensional predictions or a generalization of the morphodynamic evolution of the ETD sandbanks was yet feasible.

Thus, to the best of our knowledge, none of the previous studies have yet integrated dynamic bathymetric data for spatial SWH predictions. Therefore, the objective of this study was to develop a two-dimensional spatial machine learning prediction model capable of integrating the constantly morphodynamically changing ETD sandbanks in the nearshore area of Norderney. Due to the limited availability of high temporal resolution bathymetry data, a new methodological approach was required to create a training dataset of simulated bathymetries. For the first time, synthetic ETD sandbank morphologies were therefore generated using a geostatistical approach based on variogram analysis and random field simulation, allowing real-time exploration of the effects of different bathymetries. Two mixed-data CNN deep neural networks were designed for spatial SWH prediction. The simulated synthetic ETD sandbank morphologies and metocean measurement data were used as input, and corresponding numerical SWAN-modeled wave fields as spatial ground truth. The model setup, calibration, and validation for the nearshore area of Norderney are mostly adopted from Jörges et al. (2021a). Special focus was set on the model performance at the Norderney coastline and a climate change scenario including massive sea level rise. The main contributions of this study were as follows:

- Spatial ETD sandbank bathymetric data were simulated with random fields as features to improve wave height prediction
- A mixed-data CNN deep neural network was designed to predict spatial nearshore wave heights

4.2 Material and methods

4.2.1 Study area

This study focuses on the nearshore area of the East Frisian North Sea island of Norderney, Germany (Fig. 4.1). Separated by the tidal inlets Norderneyer Seegat and Wichter Ee, the islands of Juist and Baltrum are located to the west and east of the barrier island of Norderney, respectively. Strong tidal currents in the tidal inlets drive the sediment transport of the ETD and the tidal flats. The wave climate in the study area is dominated by the morphodynamic of the ETD sandbanks. The study area was therefore selected to fully cover the variation of random plate displacements. Further details on ETD sandbank dynamics are given in Section 4.2.3. While the long-term Mean High Water (MHW) at gauge Norderney Riffgat is 1.23 m above the German mean sea level (NHN), its mean tidal range is 2.47 m, with the study area affected by sea level rise. The main wind direction is southwest.
Fig. 4.2 summarizes the structure of the methodological framework of this study concluding with the results of the performance analysis. Detailed information about each step is given in the following subsections.

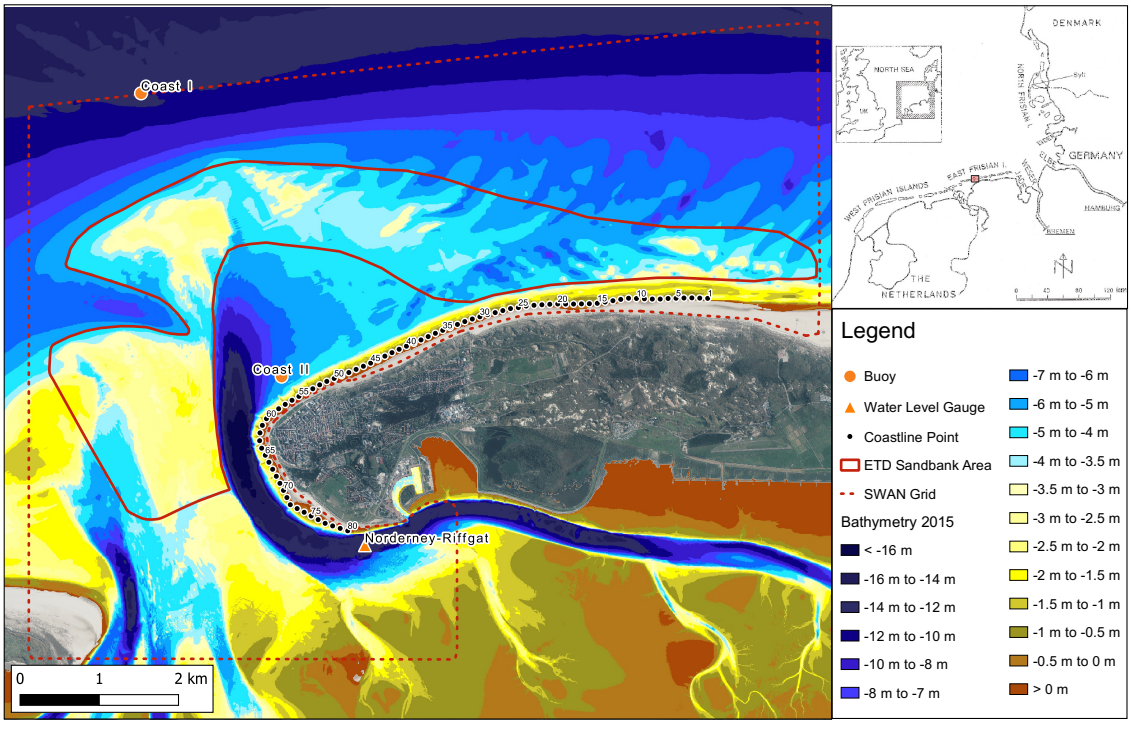

Figure 4.1: *Map of the research area, SWAN grid boundary, and buoy positions. The boundary of the simulated bathymetries (red line) and the output locations on the coastline (black dots) are presented on the map.*

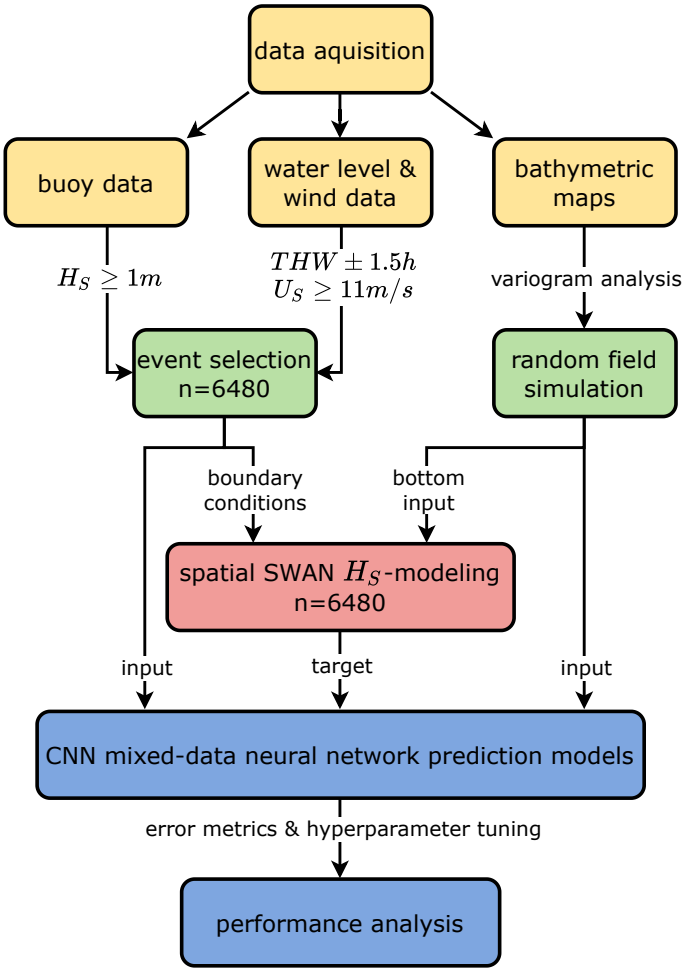

Figure 4.2: *Structure of the methodological framework of this study.*

4.2.2 Metocean data

The different plates of the ETD sandbank area north and west of the Norderney coastline are located between the northern buoy Coast I and the southern buoy Coast II (Fig. 4.1). Therefore, nearshore wave heights and periods at position Coast II are significantly influenced by morphodynamic changes of the ETD sandbanks (Jörges et al., 2021a; Niemeyer, 1979), why including the bathymetry of the sandbanks improves nearshore wave height prediction at position Coast II (Jörges et al., 2021b).

Buoy measurements of half-hourly resolution for the years 2004–2017 (without 2012–2014) – measured by two Waverider (Datawell, Haarlem, The Netherlands) buoys – were provided by the Coastal Research Station of the Lower Saxony Water Management, Coastal Defence, and Nature Conservation Agency (NLWKN), Norderney (Tab. 4.1). The buoy-measured feature parameters SWH H_{m_0} $[m]$, mean wave periods $T_{m_{0,1}}$, energy period $T_{m_{-1,0}}$, and peak period T_P of buoy Coast I were used as inputs for the spatial machine learning prediction of H_{m_0}. All parameters were

calculated from the energy density spectra of the buoy measurements (Tab. 4.1), while the same spectra were also used as model input for SWAN. Wave direction at Position Coast I was not considered as necessary model input since not all buoys are able to measure it, thus providing greater independence of the deployed buoy system. In addition, there is a correlation between wave direction and wind direction, currents, and coastal refraction, so wave direction was indirectly taken into account. Buoy Coast II was only used for model validation.

Table 4.1: *Main statistics of the two buoy measurement positions Coast I and Coast II.*

Buoy Position	SWH Mean [m]	SWH Std. [m]	SWH Range [m]	Mean energy wave period [s]	Location	Water depth 2015 [m NHN]
Coast I	1.07	0.69	[0.09; 6.91]	5.77	N 53.744781 E 7.112101	12.11
Coast II	0.63	0.39	[0.04; 3.01]	6.04	N 53.715653 E 7.141280	6.15

Wind data of the station Norderney (53.71° N, 7.15° E, 16 m height) were provided by the German Meteorological Service (Deutscher Wetterdienst, DWD) with a temporal resolution of 10 min. Water level data of the gauge Norderney-Riffgat (Fig. 4.1) were provided by the Waterways and Shipping Administration (WSA) with a temporal resolution of 1 min. An averaging to a common half-hourly resolution of water level and wind parameters was performed, thus wind speed at sea U_S $[ms^{-1}]$, wind direction U_θ [°], and water level WL $[m + NHN]$ were also used for spatial predictions of H_{m_0}.

Since storm surge events are of great interest for coastal regions due to their special strength, features were selected accordingly. The data were filtered by $H_{m_0} \geq 1\ m$, $U_S \geq 11\ ms^{-1}$, and events within 1.5 h of the tidal high water (THW), so THW $\pm$ 1.5 h. All measurement time steps with at least one missing value (NaN) of the above-described parameters were deleted completely. The feature selection procedure resulted in 6 480 data points which were used for further analysis.

4.2.3 ETD sandbank characterization and geostatistical bathymetry simulation

As stated in the introduction, the morphodynamic influence of the ETD sandbanks must be considered for spatial wave height prediction, since they exert a huge and variable influence on sea state damping as natural coastal protection in the nearshore area of Norderney (Jörges et al., 2021a,b). Therefore, in addition to wave, water

level, and wind data, bathymetric data were used for further SWAN modeling and spatial CNN mixed-data deep neural networks.

Yearly surveyed bathymetries or digital elevation models (DEM) of the ETD sandbanks were provided by the NLWKN, Coastal Research Station, and EasyGSH-DB for the years 1995, 2004–2011, and 2015–2017. EasyGSH-DB bathymetries were obtained from the EasyGSH-DB portal `www.easygsh-db.org` (Hagen et al., 2020) and considered only after visual validation and for years with no measurements of the NLWKN Coastal Research Station available. Due to immense financial effort and small time windows, more frequent survey missions by ship were not possible in the study area.

Although there is a constant eastward migration of individual plates over a period of about 37 years for a complete cycle through the entire study area due to currents, the morphodynamic changes in general are rather considered to be random (Bremermann and Meyer, 2012; Jörges et al., 2021a; Meyer, 2014; Niemeyer, 1979). Thus, even single severe storm surges can lead to rapid and immense changes of the ETD sandbanks (Wang et al., 2012). These high random dynamics can therefore not be captured at a high temporal resolution using common measurement techniques. To overcome the problem of limited bathymetric training data available, we developed a geostatistical feature engineering method based on variogram analysis and random field simulations to generate synthetic bathymetries for machine learning model training for the first time. Morphodynamic changes and simulations only were conducted in the areas of the ETD sandbanks, defined as all plates above -3.5 m NHN (Bremermann and Meyer, 2012). A smooth and averaged -3.5m-isobath of all years was computed and used as the boundary of the ETD sandbank area (Fig. 4.1).

Based on the empirical variograms (Matheron, 1965) calculated for the defined ETD sandbank areas of the respective years, a mean theoretical variogram model of the Matérn class (Matérn, 1960) was fitted to the bathymetric data variograms by weighted least squares using the gstat package of R (Pebesma, 2014). Matérn isotropic covariance F is defined as (Handcock and Stein, 1993; Stein, 1999):

$$F(h) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \left(\frac{h}{r}\right)^{\nu} K_{\nu}\left(\frac{h}{r}\right) \tag{4.1}$$

where h is the spatial distance between points, K_{ν} a modified Bessel function of the second kind of order ν (Abramowitz and Stegun, 1972), Γ the gamma function, $r > 0$ the range parameter indicating the rate of correlation decay with distance, and $\nu > 0$ a smoothness parameter. Due to the smoothness parameter, the Matérn model is

flexible in modeling spatial covariances of various spatial processes (Minasny and McBratney, 2005). E.g. for $\nu = 0.5$ the Matérn model is equal to the exponential model, while small ν represent rough spatial processes and large ν smooth processes. The semivariance γ of the Matérn model with nugget effect c_0, partial sill variance c_1, and sill variance $c_0 + c_1$ is calculated from Eq. 4.1 by:

$$\gamma(h) = c_0 + c_1(1 - F(h)) \tag{4.2}$$

The fitted mean Matérn variogram of the ETD sandbanks – averaged over all years – had the parameters $c_0 = 0.11$, $c_1 = 1.23$, $r = 262.83$, and $\nu = 2.08$ (Fig. 4.3). For further random field simulations of bathymetries, the semivariance defined by the above model was used.

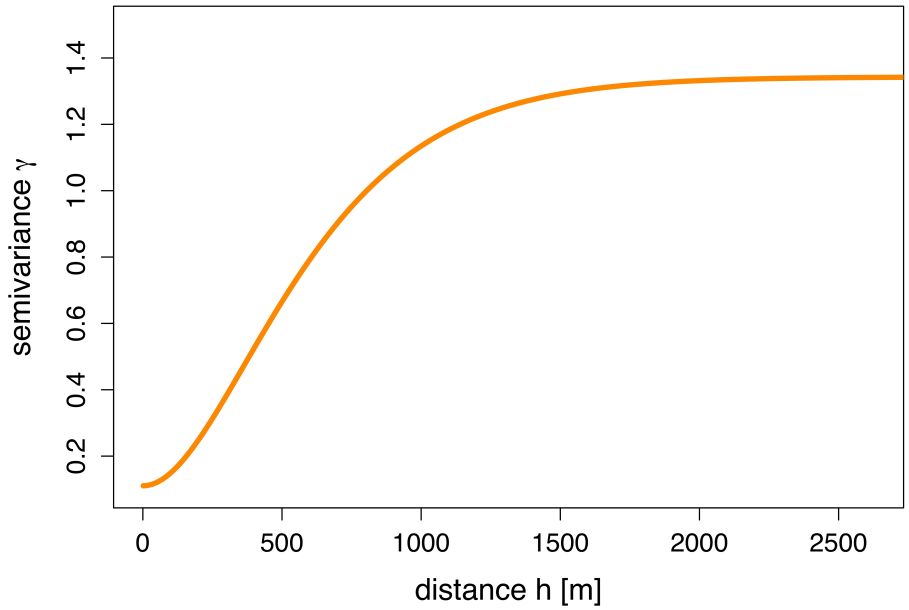

Figure 4.3: *Mean variogram of the ETD sandbanks of the Matérn class, averaged over all years. Parameters are* $c_0 = 0.11$, $c_1 = 1.23$, $r = 262.83$, *and* $\nu = 2.08$, *according to Eq. 4.1 and Eq. 4.2.*

Due to the random behavior of the ETD sandbanks, a random field was appropriate for simulating bathymetrical changes. Gaussian random field (GRF) simulation with zero-mean and a covariance matrix given by the above-defined Matérn model was therefore performed using R statistical computing software (Schlather, 2001, 2012). GRF use a (zero-mean) multivariate normal distribution to generate values and can be used to model natural spatial processes. A two-dimensional GRF is equivalent to a two-dimensional stochastic process (Lantuéjoul, 2002). Circular embedding with computation on a torus based on Fourier transformation was used in this study due

to its fast simulation performance for stationary GRFs (Dietrich and Newsam, 1997). The more surveyed bathymetries are included in the variogram, the more precise the simulation with random fields becomes. Random variations of the ETD sandbanks were achieved by adding the yearly mean bottom height of the ETD sandbank area to the simulation of the respective year and merging the final simulated ETD sandbank area with a measured background bathymetry of the respective year. The background bathymetries for the respective years consist of the corresponding survey data. Since the spatial extent of the yearly bathymetries differed in the study area, sections of missing values were filled by a reference bathymetry from the year 2015. To obtain a smooth boundary between the simulated and measured areas, a two-dimensional Gaussian smoothing with filter size 3×3 was computed for overlapping boundary areas (Chung, 2021). Finally, the accuracy of the synthetic bathymetries was visually verified for plausibility by experts, and single unnatural bathymetries (e.g. sandbanks with large holes) were removed. A representative sample of four simulated bathymetries with artificial ETD sandbanks is given in Fig. 4.4. These cover the varying magnitude of the plates in general as well as the varying spatial extent of accumulations of plates.

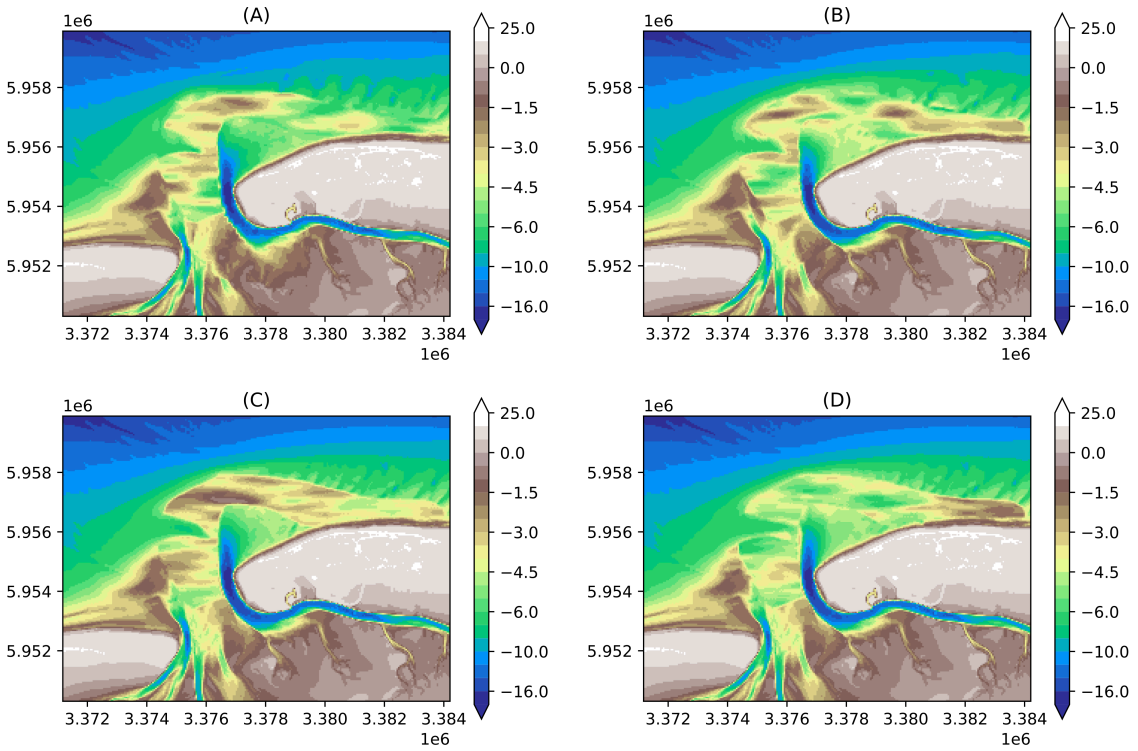

Figure 4.4: *Example of four simulated ETD sandbank bathymetries. Different survey data were used as the mean of the variances, with (A) modified based on the bathymetry of 1995, (B) based on 2005, (C) based on 2015, and (D) based on 2017. All values are given in m above the mean sea level (NHN).*

4.2.4 SWAN modeling

Since simulated bathymetries were used in this study, nearshore spatial wave data first had to be generated by numerical modeling to be used as ground truth for the machine learning models. The third-generation wave model SWAN (Booij et al., 1999), designed for coastal areas, was used with an unstructured grid. Wave propagation is computed in SWAN by solving the spectral action balance equation:

$$\frac{\partial}{\partial t}N + \frac{\partial}{\partial x}c_x N + \frac{\partial}{\partial y}c_y N + \frac{\partial}{\partial f}c_f N + \frac{\partial}{\partial \theta}c_\theta N = \frac{S}{f} \tag{4.3}$$

where $N = N(f, \theta)$ is the action density spectrum with frequency f and wave direction θ, t the time, and x and y the spatial propagation with velocities c_x and c_y, respectively. The fourth and fifth terms of Eq. 4.3 describe the changes in depth and currents, and the changes in direction caused by refraction, respectively (Booij et al., 1999; SWAN, 2019a,b).

The unstructured grid (Fig. 4.1) was adapted to the bathymetries including a mean ETD sandbank shape to achieve both a good fit to the topography and a low number of 29 347 nodes in total. This method reduced the computational cost compared to a regular grid. The highest grid resolution was about 20 m in the area of the morphodynamic ETD sandbanks, while the lowest resolution of about 120 m was set at constant boundary positions. SWAN computations were executed in stationary mode using SWAN version 41.20AB, and the unstructured grid was designed using the Surface-water Modeling System (SMS) with the paving method (Surface-Water Modeling System, 2012).

The SWAN modeling setup and model calibration and verification for this region are described in more detail in Jörges et al. (2021a). All model parameters were adopted from Jörges et al. (2021a) except for the proportionality coefficient α for the rate of dissipation of the $\beta - kd$ parametrization of depth-induced wave breaking (Salmon and Holthuijsen, 2015; Salmon et al., 2015). It was set to 0.15 in this study, which showed the overall best results.

While the northern boundary of the computational grid was controlled by the signals of buoy Coast I, the western boundary was controlled according to eight one-dimensional computations using SWANOne (Verhagen et al., 2008) with the same hydrodynamic boundary conditions applied to the respective northern boundary. Southern and eastern boundaries were set as open boundaries. Wind and water level inputs were defined according to the measured data of the respective event and assumed to be constant over the entire computational grid (Witting et al., 2014).

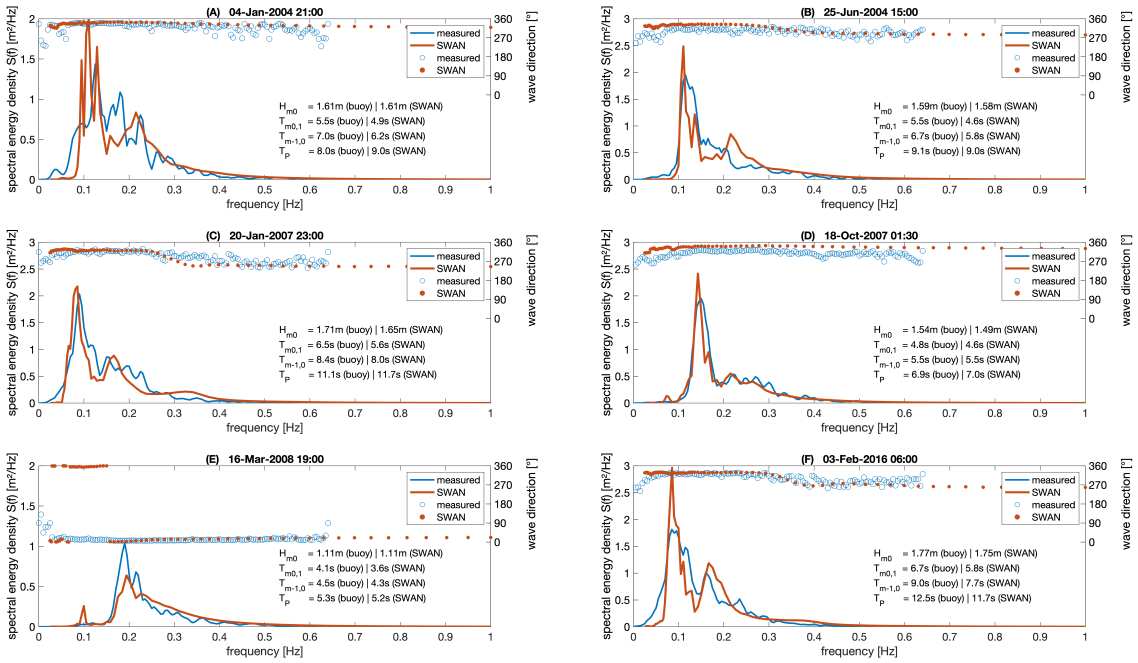

Figure 4.5: *Variance density spectra of six validation cases of the SWAN modeling. The observed (blue) and modeled (red) spectra and directions as well as the basic wave parameters are shown for the events measured on A: 04-Jan-2004 21:00, B: 25-Jun-2004 15:00, C: 20-Jan-2007 23:00, D: 18-Oct-2007 01:30, E: 16-Mar-2008 19:00, and F: 03-Feb-2016 06:00.*

Validation of the SWAN model was performed on 59 measured events with surveyed bathymetries and buoy-measured boundary conditions. Good performance was achieved on the validation data with a MAAPE of 14.3%, a R^2 of 0.725, an RMSE of 0.218 m, and a MAE of 0.18 m (Tab. 4.4, Fig. 4.11). Fig. 4.5 shows the performance of the SWAN model for six sample events. Although the peak values had the highest error, the general fit of the variance density spectra was satisfactory, as well as the accuracy of the wave directions. In conclusion, the SWAN-modeled data were suitable to produce an input dataset for further machine learning algorithms. A total of 6 480 SWAN models were computed. Fig. 4.6 provides examples of four events with different simulated bathymetries covering the range from high (A) to low (D) resulting wave fields.

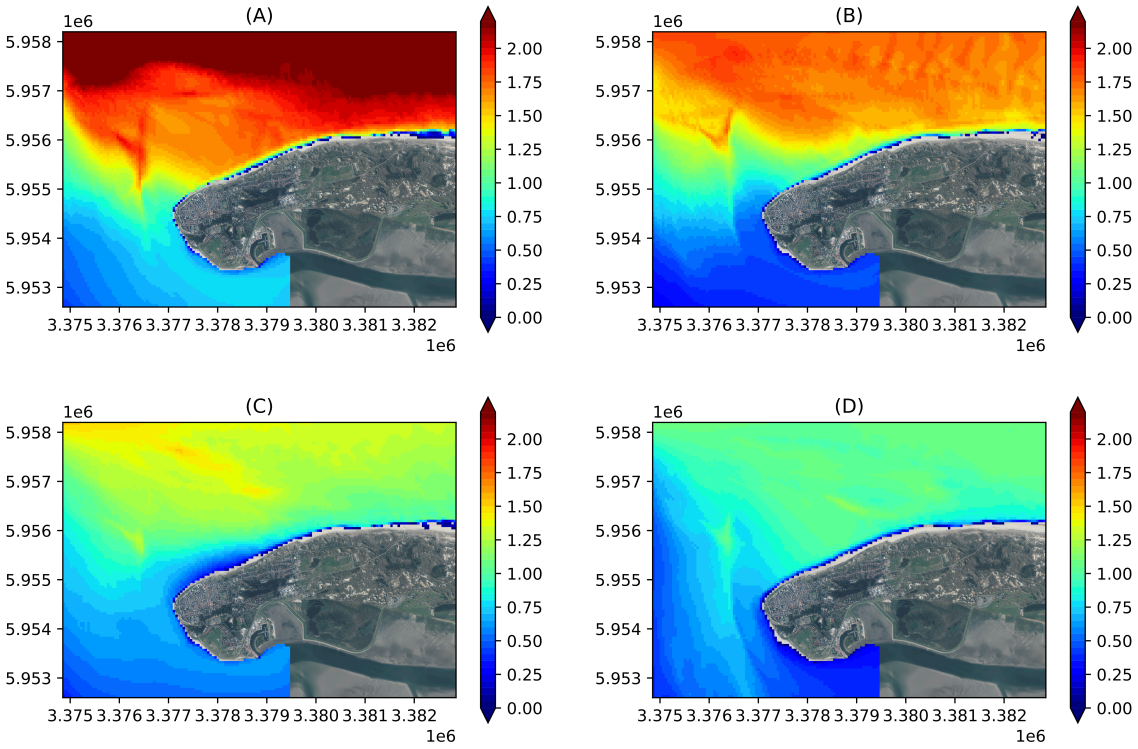

Figure 4.6: *Example of four SWAN-modeled spatial wave fields of various strengths. (A) represents an event with a northern boundary wave height (Coast I) of* $H_{m_0} = 3.28\ m$*, wind speed* $U_S = 17.13\ ms^{-1}$*, and water level* $WL = 1.97\ m\ NHN$*. (B) represents an event with a northern boundary wave height (Coast I) of* $H_{m_0} = 1.91\ m$*, wind speed* $U_S = 12.31\ ms^{-1}$*, and water level* $WL = 1.35\ m\ NHN$*. (C) represents an event with a northern boundary wave height (Coast I) of* $H_{m_0} = 1.83\ m$*, wind speed* $U_S = 16.51\ ms^{-1}$*, and water level* $WL = 1.73\ m\ NHN$*. (D) represents an event with a northern boundary wave height (Coast I) of* $H_{m_0} = 1.22\ m$*, wind speed* $U_S = 11.52\ ms^{-1}$*, and water level* $WL = 0.75\ m\ NHN$*. All* H_{m_0} *values are given in* m*.*

4.2.5 Neural network data preprocessing

All in-situ point-measured and spatial SWAN-modeled features were min-max-normalized to the interval $[0, 1]$ by

$$x_N = \frac{x - min(x)}{max(x) - min(x)} \tag{4.4}$$

where x_N is the normalized and x the original feature value. Feature normalization prevents any feature from having a disproportionately influence on the network training by its range of values.

The 6 480 data points were divided into 75% train data (4 860 data points) covering the simulated bathymetries from 2004–2011 and 25% test data (1 620 data points) covering the simulated bathymetries from 2015–2017 for each machine learning model run, respectively. On this, a random shuffle before each training epoch

was performed to ensure robust results. Feature normalizations only refer to the min and max of the respective training data to hide any information (e.g. the range of values) of the test data during the model training. Thus, test data were not shown to the algorithms during training in any way. Especially the randomly simulated bathymetries of the test data were hidden during the training and therefore used to assess the performance of generalization of the proposed method.
Both, simulated bathymetry maps and SWAN-modeled wave height maps, were scaled to a spatial resolution of 50 × 50 m due to computational costs. This included a down-sampling by a factor of 10 for the simulated bathymetries resulting in 191 × 260 pixels and a down-sampling of the SWAN-modeled wave heights resulting in 113 × 161 pixels. Further data preprocessing was not performed to minimize the manipulation of the prediction results (Ali and Prasad, 2019).

4.2.6 Convolutional neural networks (CNN)

For working with data on grid topologies, the convolutional neural network (CNN) class (LeCun et al., 1989) is very effective due to its convolutional operation in multi-dimensional spaces. This allows CNNs to learn local patterns in the data (e.g. image data like maps), share weights, and shift invariants, making CNNs widely applied to e.g. image classification, face and speech recognition, and natural language processing (Aggarwal, 2018). Computed on GPUs, CNNs can fully deploy their power, especially for large spatial datasets (Li et al., 2016).
The convolution operation extracts latent features from an input image using convolution filters, called kernels. The resulting output of the convolution operation is called a feature map (Goodfellow et al., 2016). Convolution layers are basically defined by the number and size of the convolution kernels, the stride (kernel step size), and the padding (number of rows and columns added to the boundaries of the input feature matrices to adjust the size of the output feature map). All weights of the convolution kernels are learned through backpropagation during the training, while the kernel size is a tunable hyperparameter. A convolution layer operation $*$ for the layer l and the output feature map k with input feature matrix $X_k = (x_{ijk})_{i,j=1}^{n,m} \in \mathbb{R}^{n \times m}, n, m \in \mathbb{N}$ and weight matrix of the convolution kernel $W_k = (w_{i'j'kk'})_{i',j'=-r}^{r} \in \mathbb{R}^{(s \times k) \times (s \times k')}, s = 2r + 1, r \in \mathbb{N}$ is performed by:

$$z_{ijk}^{(l)} = (W * X)_{i,j,k}^{l} = \sum_{k'=1}^{d^{(l-1)}} \sum_{i',j'} x_{i+i',j+j',k'}^{(l-1)} \cdot w_{i',j',k',k}^{(l)} \tag{4.5}$$

$$x_{ijk}^{(l)} = f\left(z_{ijk}^{(l)}\right) \tag{4.6}$$

where $Z_k = (z_{ijk})_{i,j=1}^{n,m} \in \mathbb{R}^{n \times m}$ is the output feature map of the convolution and $d^{(l)}$ the number of feature maps of layer l. Various activation functions f can be applied to the convolution operation, such as sigmoid, hyperbolic tangent (tanh), and rectified linear unit (ReLU), with the latter used throughout this study due to its non-linearity, positivity, and fast computation (Khan et al., 2020). The architecture of a CNN can typically be composed of an input layer, convolution layer, pooling layer, batch normalization layer, dropout layer, fully-connected layer, and an output layer. Pooling layers reduce the spatial size of the feature map, decreasing computational cost by dimensionality reduction and achieving generalization by extracting the main features. Max pooling returns the maximum value of the area covered by the pooling filter, while average pooling calculates the average of all values covered by the pooling filter. Since max pooling also suppresses noise by discarding small noisy values, we used this method throughout this study for all CNNs.

4.2.7 Neural network model settings and architecture

Mixed-data neural networks were applied in this study to combine spatially simulated bathymetries and point-measured metocean data as input features. As mixed-data architectures are more complex, the Keras functional API was used to design flexible networks (Gulli and Pal, 2017). Two branches of the networks, a fully-connected-block for buoy-measured, wind, and water level data, and a CNN-block for bathymetric maps, were concatenated to output SWAN-modeled spatial H_{m_0} ground truth data. The architecture of the proposed CNN model is described in detail in Fig. 4.7 (a). In addition, a kind of encoder-decoder architecture for the CNN neural network (ED-CNN) was developed to overcome the feature map compression by the pooling operation with an up-sampling operation. Hence, the initial size of the feature maps was restored by inverse pooling and the CNN-block of the ED-CNN model has a symmetric shape. Details about the ED-CNN architecture are given in Fig. 4.7 (b). For computational reasons, a fully-connected layer at the end of the respective CNN blocks serves as a bottleneck.

The proposed CNN and ED-CNN mixed-data neural networks were benchmarked to a common deep FFNN (DFFNN) (Hastie et al., 2009), a Resnet-50v2 (He et al., 2016), a VGG-19 (Simonyan and Zissermann, 2015), an Inception-v3 (Szegedy et al., 2015), and a Densenet-121 (Huang et al., 2018). For the DFFNN, the CNN-block was replaced by three fully-connected layers, a batch normalization layer, and a dropout layer. While the proposed CNN and ED-CNN models as well as the DFFNN were designed for spatial regression outputs, the other benchmark models were designed for image classification. For the layers inside the CNN-block, a fixed number of a 2 $\times$ 2 pooling size, a stride of 1, and padding, resulting in an output of the same size

as the input, were defined. The choice of other model parameters (hyperparameters) is described in the next subsection since they were selected by an optimization technique.

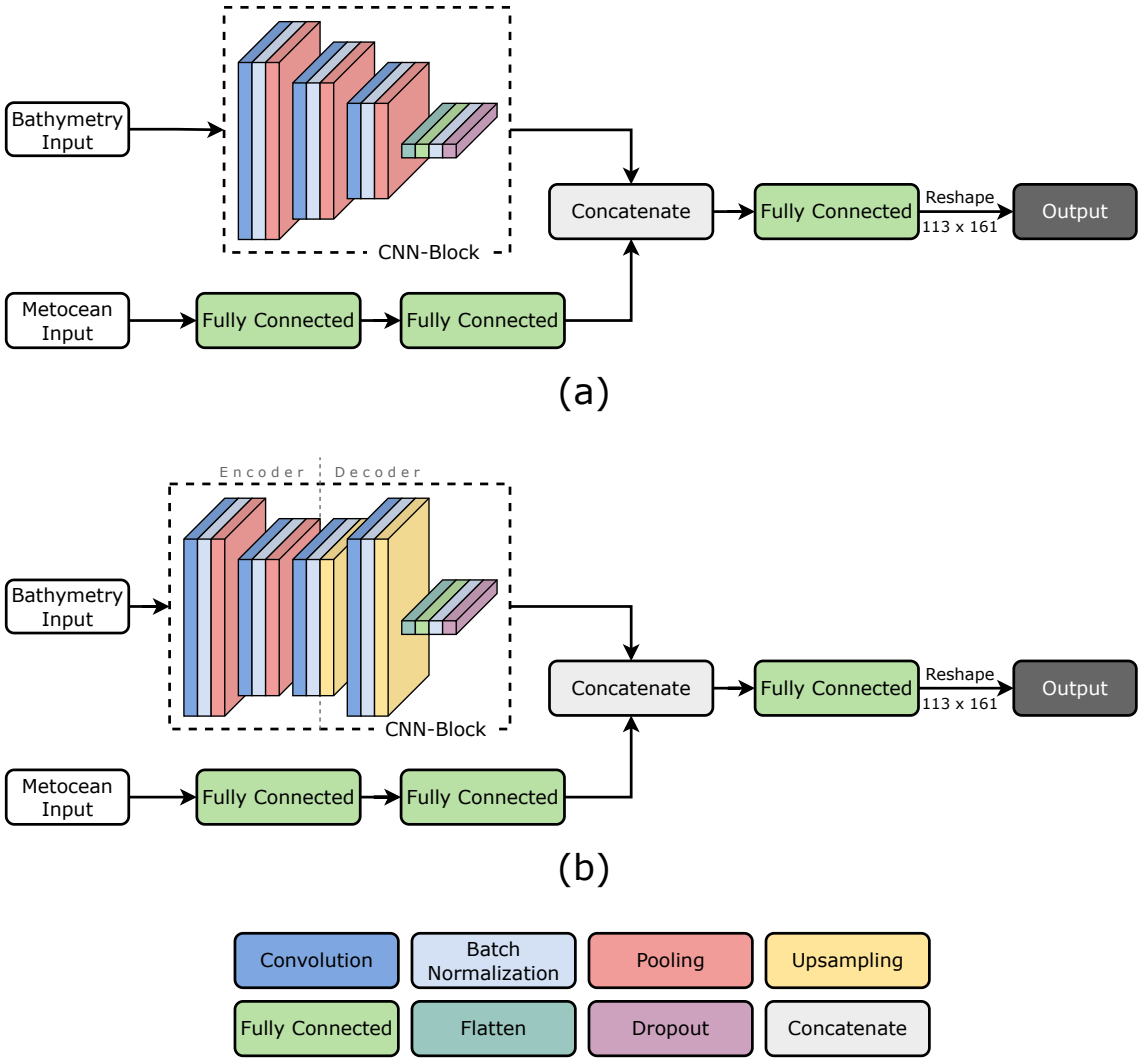

Figure 4.7: *Architecture of the proposed CNN (a) and ED-CNN (b) models for mixed-data with the CNN-block for two-dimensional bathymetry input (size 191 × 260) and a fully-connected part for the metocean input (size 42). The convolution sizes are (a) 191 × 260, 95 × 130, and 47 × 65 for the three respective layers with 32 feature maps each, and (b) 191 × 260, 95 × 130, 47 × 65, and 94 × 130 on the two encoder and two decoder convolution layers with 64, 32, 32, and 64 feature maps, respectively. The output size is 113 × 161 for both (a) and (b).*

A mixed-data network input $I = (I_1, I_2)$ consists of a concatenation of one-dimensional fixed point metocean data I_1 and two-dimensional spatial simulated bathymetric data I_2. The model output Y consists of the two-dimensional SWAN-modeled H_{m_0} data. Y is represented by a matrix containing zeros at the respective geographic

positions above the water level of an event, so-called dry points. All dry points were excluded from further performance evaluation since they were learned correctly by the networks. Due to the need of simulating the ETD sandbanks, time series based on buoy measurements could not be obtained. Thus, all predictions were made for the respective time step of the input data.
A fixed number of 30 epochs for training was set for all neural networks and every model was trained and tested in an ensemble of three independent runs to validate the robustness of the prediction performance by its standard deviation with respect to random variations by data shuffling or the optimization algorithm during training. All calculations were performed with Python 3.8 and TensorFlow 2 on an Intel Xeon Gold 6146 CPU 12 Core 3.2Ghz, 64GB RAM, and a Quadro P1000 4GB GPU.

4.2.8 Loss function, optimizer, and error metrics

A mean squared error (MSE) loss function L for all models in this study was computed as follows:

$$L(I) = \frac{1}{b} \cdot \frac{1}{113 \cdot 161} \cdot \sum_{i=1}^{b} \sum_{j=1}^{113 \cdot 161} \mid \hat{y}_{i,j} - y_{i,j} \mid^2 \tag{4.7}$$

where $b \leq n$ is the tunable size of a mini batch of a power of 2 applied for weight updates. n is the total number of two-dimensional input data, $\hat{y}_{i,j}$ and $y_{i,j}$ the jth pixel of the ith predicted and original reshaped output matrix of size 113 $\times$ 161, respectively. Adaptive moment estimation (Adam) (Kingma and Ba, 2015) was used as an optimization algorithm for the backpropagation computation of the gradient of the loss function. Both, the learning rate α and decay of α for a robust convergence of the optimization were tunable network parameters underlying Bayesian hyperparameter optimization.

Error metrics root mean squared error (RMSE), mean absolute error (MAE), mean error or bias, mean arctangent absolute percentage error (MAAPE) (Kim and Kim, 2016), Pearson correlation coefficient r, and the coefficient of determination R^2 were computed to assess the accuracy of the models as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{4.8}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \mid \hat{y}_i - y_i \mid \tag{4.9}$$

$$bias = \frac{1}{n}\sum_{i=1}^{n} (\hat{y}_i - y_i) \tag{4.10}$$

$$MAAPE = \frac{1}{n}\sum_{i=1}^{n} arctan\left(\left|\frac{\hat{y}_i - y_i}{y_i}\right|\right) \tag{4.11}$$

$$r = \frac{cov(\hat{y}, y)}{\sqrt{(var(\hat{y})var(y)}} \tag{4.12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{4.13}$$

where $cov(x, y)$ is the covariance of x and y, $var(x)$ the variance of x, and $\bar{y}$ the mean of the original values. All error metrics presented in this study refer to rescaled (spatial) predictions and do not underlie any normalization.

4.2.9 Hyperparameter tuning

Bayesian hyperparameter optimization was conducted on the training data to find the best network hyperparameters for the proposed model (Bergstra et al., 2011; Snoek et al., 2012). The hyperparameters number of filters in each convolution layer, kernel size of convolution layer, number of nodes in the final fully-connected layer, mini batch size b, learning rate α, decay of learning rate α, and the dropout rate were considered during tuning. 50 tuning iterations were computed for the CNN, ED-CNN, and DFFNN neural networks, respectively, by maximizing the scale-independent R^2 objective function. Since it is not possible to adjust all parameters in the benchmark models, the optimized parameters have been adopted here.
Bayesian hyperparameter optimization evaluates the CNN model as a random black-box function of the hyperparameters and estimates the best parameters by performing only a small number of the expensive objective function evaluations (Brochu et al., 2010). By prior laying a Gaussian process over the objective function, parameters can be selected by likelihood, and after each function evaluation, the next hyperparameters can be chosen by building a posterior (Cabrera et al., 2020; Cornejo-Bueno et al., 2018; Snoek et al., 2012).
Tab. 4.2 gives the respective search domains and results of the Bayesian hyperparameter optimization for CNN, ED-CNN, and DFFNN.

Table 4.2: *Hyperparameter tuning values and results.*

Parameter	Search interval	CNN	ED-CNN	DFFNN
# CNN filter	[8, 128]	32	64	–
kernel size	[2, 10]	3	4	–
# nodes final layer	[5, 10 000]	100	100	50
mini batch size b	[1, 128]	32	64	32
learning rate α	[1e-4, 1e-2]	0.0025	0.025	0.035
decay of learning rate α	[1e-6, 1e-2]	0.0005	0.005	0.005
dropout rate	[0, 0.5]	0.25	0.25	0.35

4.3 Results and discussion

We first present the comparison of the performance results for the spatial prediction of H_{m_0} of the different model architectures. Subsequently, the proposed models have been validated with in-situ buoy measurements as well as the SWAN numerical model. Results of the wave height prediction uncertainty are then shown and discussed in detail for the Norderney coastline. In a final step, the proposed prediction model is applied to an exemplary scenario of sea level rise driven by climate change.

4.3.1 Model performance of spatial prediction

The proposed CNN and ED-CNN architectures of this study were compared to the Resnet-50v2, VGG-19, Inception-v3, and Densenet-121 models. As the performance on the test data of any of these alternative models was consistently weak with an RMSE $>$ 1 m, they were not considered for further analysis. These models were initially designed for classification problems with single value outputs and showed a highly overfitting during the training process. Therefore, we cannot recommend these models for image-to-image predictions with mixed-data input.
Tab. 4.3 provides details of the performance of the CNN, ED-CNN, and DFFNN models. The overall performance of the tuned CNN-based models was very good and in range with the performance of the studies mentioned in the introduction, while the tuned DFFNN had a worse performance. The best prediction results were achieved by the CNN model with an RMSE of 0.097 m for the entire wave field area on average for the three independent model runs. The variance of the different ensemble runs was sufficiently small to be negligible, proving the robustness of the predictive model. The proposed CNN model reached the highest R^2 of 0.977, a MAAPE of 0.067, and a MAE of 0.068 m on the spatial test data. On average proper time

consumption of about 91 min for training and 4 s for predicting the spatial test data was obtained. Similar good results were achieved with the ED-CNN model. While the spatial mean RMSE, R^2, and MAAPE of the test data were 0.103 m, 0.974, and 0.069, respectively, the mean computational time was about 376 min for training and 35 s for predicting. The CNN has a slimmer architecture, thus fewer weights need to be trained. The very similar performance of the ED-CNN and CNN models can be explained by the final fully-connected layer which provides a bottleneck in the CNN-based models. Both models showed relatively high variability in the bias error, probably due to its sign sensitivity. In contrast to the CNN-based models, the DFFNN showed a much higher spatial mean error with an RMSE of 0.265 m, an R^2 of 0.807, and a MAAPE of 0.191. However, the computational time for the training was only about 2 min and for the prediction only about 1 s. Not only the errors were higher for the DFFNN, but also the variance of the three computational runs was greater than for the CNN-based models. Thus, it can be stated that the CNN model can be recommended due to its good computational time and superior performance, and both CNN and ED-CNN models showed very robust results.

Fig. 4.8 shows the prediction results of the proposed CNN model for the four exemplary cases described in Fig. 4.6. By comparing the SWAN-modeled true values of Fig. 4.6 with the CNN-predicted significant wave heights of Fig. 4.8 an overall good agreement was obtained. Interactions of the waves with the simulated bathymetries were quite apparent and the CNN was able to surrogate the numerical SWAN method (Fig. 4.8). Both severe (A and B) and gentle (C and D) sea state events could be spatially predicted using the machine learning method (Fig. 4.8).

To provide a proper comparison of the SWAN-modeled and CNN-predicted wave fields, the bias between them was calculated, where a positive bias indicates an overestimation and a negative bias indicates an underestimation of the H_{m_0} wave field by the CNN. It can be seen from Fig. 4.9 that the prediction bias of the CNN model had a spatial pattern depending on the simulated bathymetry. While exemplary cases A, B, and D showed underestimations at the northern beach, event C showed overestimations at the northern beach area. Significant biases were mostly found around the area of the simulated ETD sandbanks. The areas of the sandbanks were mostly underestimated, while the rest of the nearshore area tended to be overestimated.

Table 4.3: *Spatial prediction of SWH performance results. The mean value and standard deviation of the three independent runs are given, respectively.*

Model	Training Time [min]	Prediction Time [s]	RMSE [m]	R^2	r	MAAPE	MAE [m]	bias [m]
CNN	91.5 ± 0.5	4.4 ± 0.0	**0.097** ± 0.001	**0.977** ± 0.001	**0.989** ± 0.001	**0.067** ± 0.003	**0.068** ± 0.002	**-0.013** ± 0.005
ED-CNN	376.2 ± 2.3	35.4 ± 0.5	0.103 ± 0.003	0.974 ± 0.002	0.987 ± 0.001	0.069 ± 0.003	0.071 ± 0.002	0.016 ± 0.014
DFFNN	2.1 ± 0.0	0.7 ± 0.0	0.265 ± 0.086	0.807 ± 0.126	0.930 ± 0.037	0.191 ± 0.053	0.187 ± 0.063	-0.059 ± 0.071

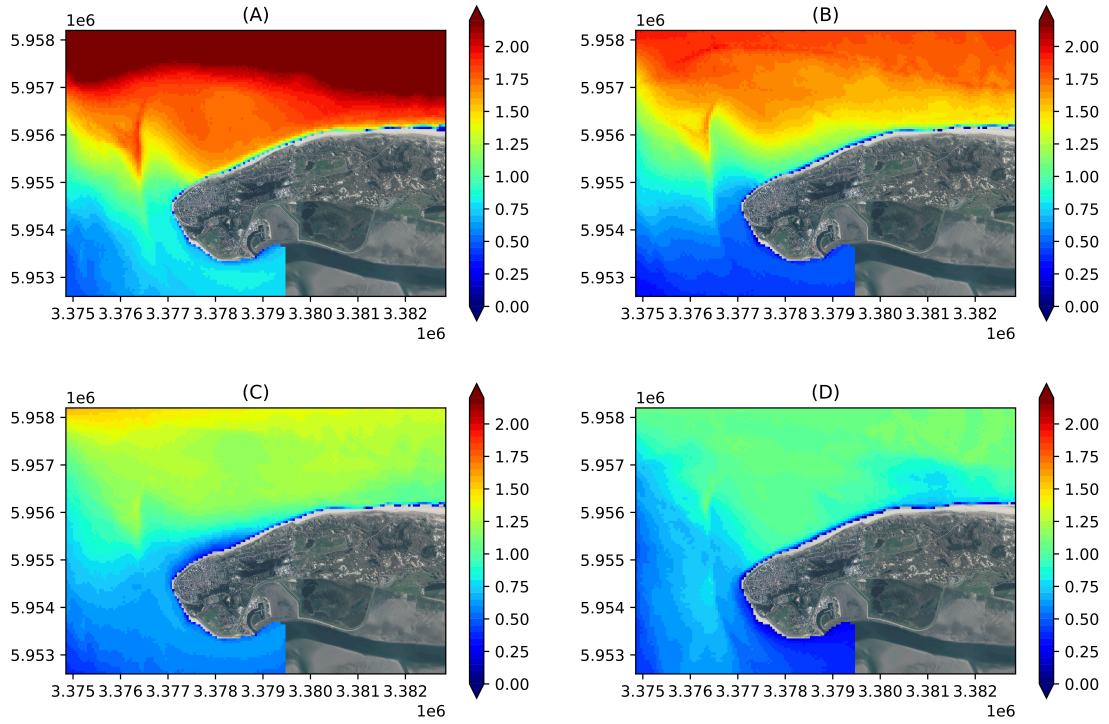

Figure 4.8: *Example of the four CNN-predicted spatial wave fields described in Fig. 4.6. All H_{m_0} values are given in m.*

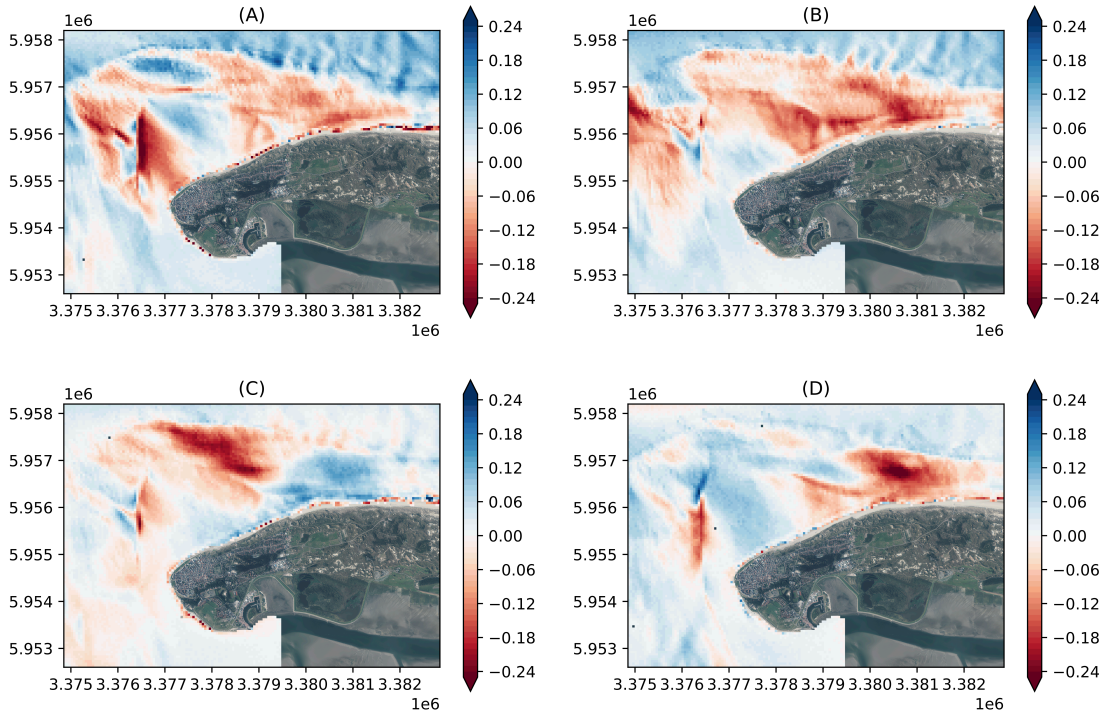

Figure 4.9: *Spatial bias for the four exemplary events presented in Fig. 4.6 and Fig. 4.11. Positive values (blue) indicate an overestimation and negative values (reds) indicate an underestimation at the respective position. All bias values are given in m.*

Fig. 4.10 shows the spatial RMSE of the entire model area of the test data. It indicates strong interactions with the shape of the simulated bathymetries. The largest RMSE of about 0.25 m was found in the area of the ETD sandbanks and the northern coastline. On average, the highest RMSE was obtained at the transition area between the ETD sandbanks and the deeper shipping channel (Fig. 4.1, Fig. 4.10). The RMSE of the ETD sandbank area was mostly less than 0.15 m, which is satisfactorily low compared to the mean significant wave height. North of the ETD sandbanks the RMSE was mainly about 0.06 m. The lowest spatial mean RMSE of about 0.02 m was in the southeast region of the model area. Due to the higher significant wave heights in the northern model area, also higher errors were expected. The RMSE at the Norderney coastline has a heterogeneous structure with small errors in the south, moderate errors at the western beach, and the highest errors at the northern beach. In general, regions of higher errors of the CNN model can be explained by the greater morphodynamic variations of the simulated ETD sandbanks and resulting complex wave interactions in the respective area. Boundary regions that were held constant over all bathymetries therefore tend to have smaller prediction errors. Wave propagation over the ETD sandbanks was accurately predicted by the CNN.

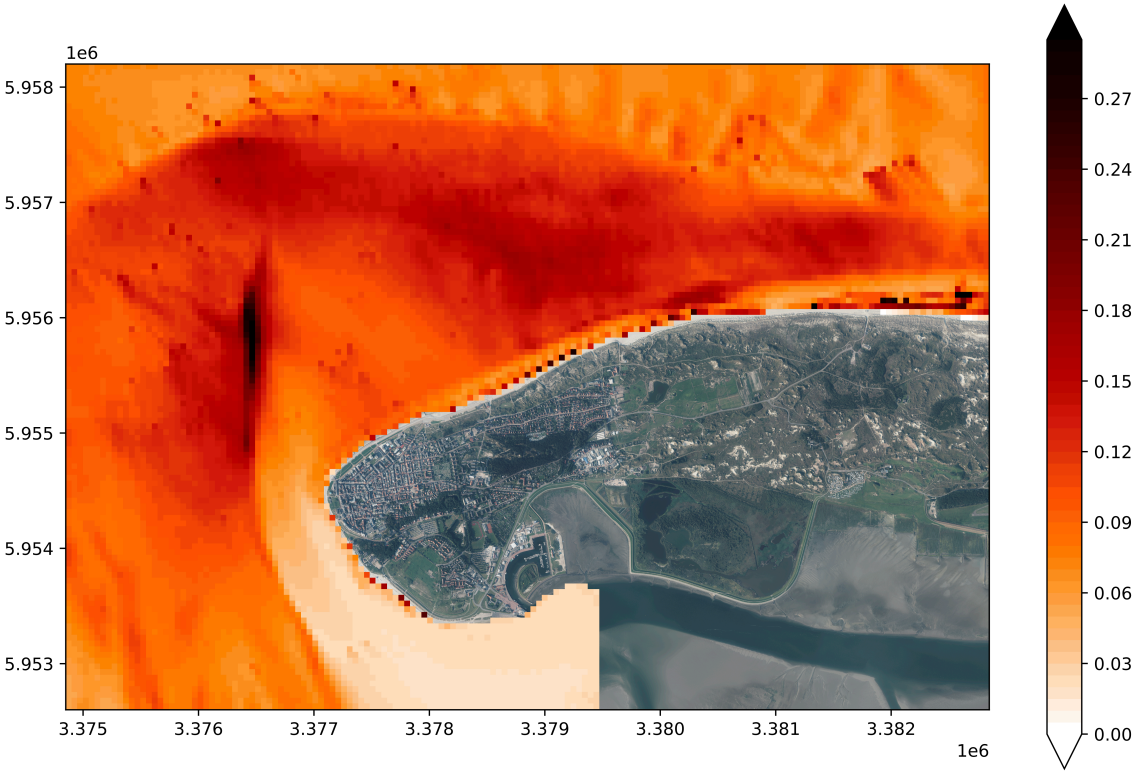

Figure 4.10: *Spatial RMSE of the CNN model for the entire test data. All values are given in m.*

4.3.2 CNN neural networks vs. SWAN numerical model

Compared to SWAN modeling, the computational efficiency of the mixed-data CNN neural networks proposed in this study was greatly higher. Both methods were computed on the same hardware, thus a comparison of the computation time was reasonable. SWAN required about 15.8 days or 22 755 min to model the entire test dataset, while the trained CNN model required only about 4.4 s for the spatial prediction of the same dataset. Training of the mixed-data CNN neural network needed about 91.5 min but had to be performed only once for the study area. Therefore, the proposed machine learning method was able to reduce the computational cost by a factor of more than 300 000 compared to the numerical SWAN model. Feng et al. (2020) obtained an acceleration of computational time by a factor of up to 20 000 with a fully-connected neural network compared to SWAN modeling for the area of Lake Michigan, while James et al. (2018) reached a factor of up to 1 000 for the area of Monterey Bay with a similar setting. Also, Wei and Davison (2022) reached a 22.5 times faster computation with a CNN model compared to SWASH numerical model. It must be noted that the spatial SWAN model was computed on 29 347 nodes, while the CNN model reached a resolution of 113 $\times$ 161 pixels resulting in 18 193 nodes. A higher resolution of the CNN model caused memory issues on the hardware during training, resulting in the need for high-performance computers for huge datasets as well. Furthermore, the training dataset had to be modeled first using the computationally expensive numerical method of SWAN, but nevertheless, the predictions on new data were performed significantly faster by machine learning, providing an accurate real-time operational surrogate model.

The SWAN-modeled spatial wave fields served as ground truth for the machine learning methods. However, the performance of the CNN models evaluated on the in-situ buoy measurements is crucial in the end. Tab. 4.4 compares the prediction performance of the proposed CNN and ED-CNN models to the performance of the SWAN numerical model validated on 59 buoy-measured events at position Coast II (Fig. 4.1). While SWAN achieved an RMSE of 0.218 m and a MAAPE of 14.3%, the CNN and ED-CNN models performed very well with an RMSE of 0.23 m and 0.243 m, and a MAAPE of 15.3% and 16.3%, respectively. Thus, the errors of the CNN models were roughly equal to the error of SWAN, highlighting the power of the CNN models to efficiently provide a substitute to SWAN modeling for the research area.

Fig. 4.11 demonstrates the performance of the SWAN and CNN model by comparing their values to the buoy-measured data at position Coast II for the 59 reference events. The data points of SWAN and CNN were close around the bisector. Both, small and high H_{m_0} values, were modeled or predicted with an unbiased small error.

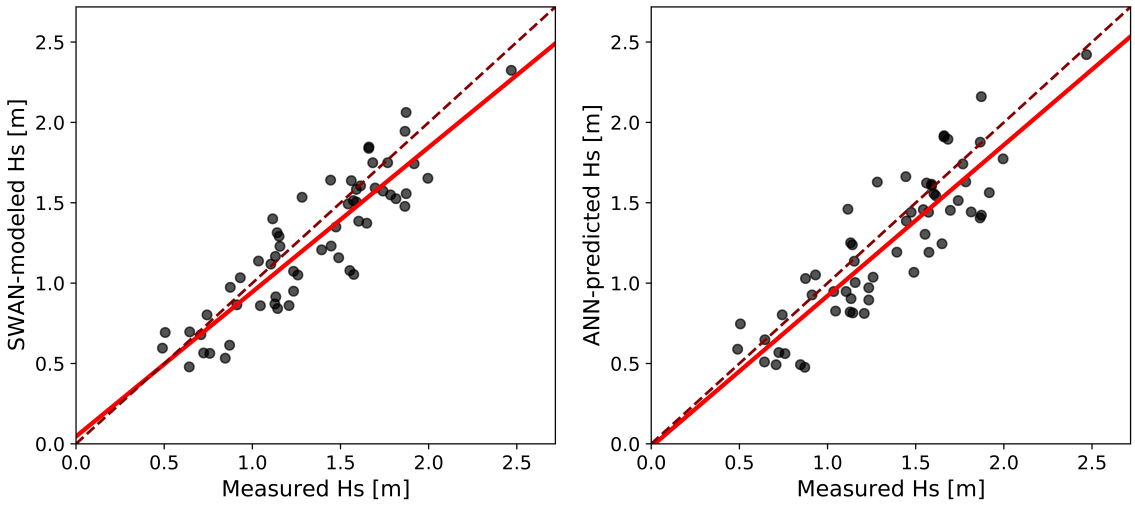

Figure 4.11: *Comparison between buoy-measured and SWAN-modeled (left) and buoy-measured and CNN-predicted (right) significant wave height H_{m_0} at position Coast II for 59 reference events.*

Both techniques showed a minor tendency to underestimate higher H_{m_0} values, as indicated by the solid red line below the bisector. The validation with buoy-measured data was only performed for position Coast II. Since this position is located south of the ETD sandbanks with a medium RMSE, it provided a proper representation, though. For areas having higher prediction errors (Fig. 4.10), no validation could be performed due to a missing data base.

Table 4.4: *Validation of CNN model and SWAN numerical model on 59 Coast II buoy data.*

Model	Error Metric					
	RMSE [m]	R^2	r	MAAPE	MAE [m]	bias [m]
SWAN	0.218	0.725	0.891	0.143	0.180	-0.090
CNN	0.230	0.704	0.885	0.153	0.191	0.081
ED-CNN	0.243	0.669	0.885	0.163	0.204	0.127

4.3.3 High-resolution uncertainty estimation on the coastline

Small-scale predictions along the coastline are of particular interest due to a high variance in sea state conditions over small areas depending on the boundary conditions and the prevailing bathymetry. In order to provide precise and rapid coastal protection (e.g. sandbags or warning of the affected population), a special focus was set on the bias prediction error along the coastline. The mean bias error (Eq. 4.10) and its standard deviation (SDB) were calculated for 80 points on the test

data along the Norderney coastline below 0 m NHN (Fig. 4.1) and are shown in Fig. 4.12. By using 0 m NHN as an upper limit, water-covered areas for THW ± 1.5 h can be ensured. Again, a positive bias indicates an overestimation and a negative bias indicates an underestimation of H_{m_0}. To obtain more robust results, an average of the surrounding points was calculated for each coastline point. In general, the bias along the coastline was quite small with huge variability. While the wave heights tended to be overestimated in the far east, the model was more likely to underestimate H_{m_0} between points 7 and 10 and points 19 and 29. Between point 30 on the north coast and point 80 on the west coast, no clear pattern of overestimation or underestimation could be observed. On the southern coastline between points 60 and 80, the bias was much lower, which can be explained by the overall lower H_{m_0} there.

The SDB was significantly greater than the bias in most cases, indicating very high variability of the bias along the coastline. This can be explained by the high variability in the various sea state events, simulated bathymetries, and other metocean boundary conditions in the test data. Although a spatial pattern could be identified in the coastline points, whether the proposed model under- or overestimates the SWH at the coastline seems to depend more on the specific bathymetric and metocean boundary conditions than on the particular coastline position. Thus, a generalized estimate of model uncertainty is not sufficient for use in coastal protection and any boundary conditions, as well as bathymetry, must be considered for uncertainty estimation.

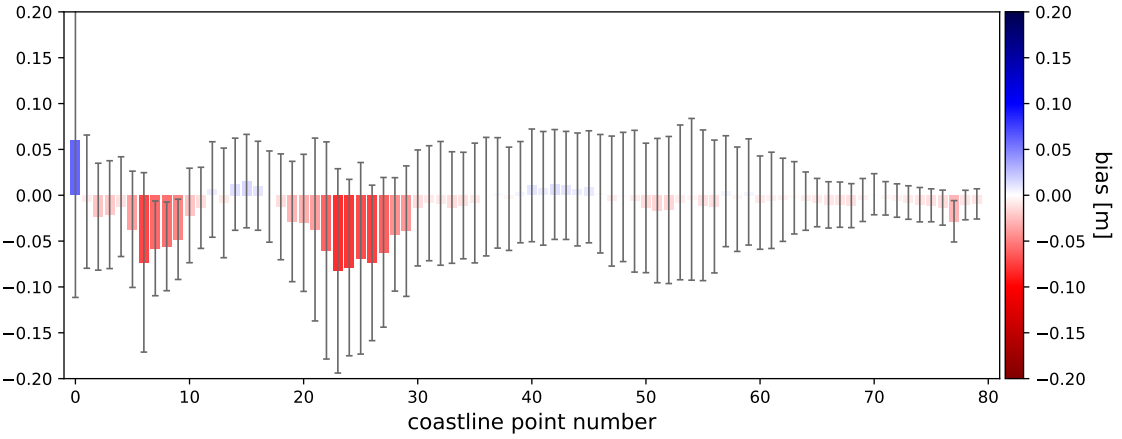

Figure 4.12: *Mean error (bias) and bias standard deviation (SDB) between the CNN-predicted and SWAN-modeled SWH at 80 points along the Norderney coastline. The spatial reference of the points can be seen from the map in Fig. 4.1. Positive values (blues) indicate an overestimation and negative values (reds) indicate an underestimation at the respective points. The grey line indicates the SDB. All bias values are given in m.*

4.3.4 Application of the proposed method for a climate change scenario

The proposed method and CNN prediction model developed in this study can be used as a rapid tool to explore various sea level rise scenarios or morphodynamic changes of the ETD sandbanks in the study area. Furthermore, possible increases in wind speed can be conducted, or coupling to climate change scenarios can be enforced. In the following, the proposed model is applied exemplarily for an event to be studied. The event was created assuming a sea level rise of 1 m in addition to the 5-year-MHW from 2016–2020 of 1.26 m, a wind surge of +3.01 m for a reference event of 1962, and a spring increase of +0.46 m. The resulting water level of 5.73 m represents the so-called design-basis water level according to the single value method, which is taken into account for the construction of coastal protection structures in the study area (Berkenbrink and Wurpts, 2019). A corresponding wind speed of 30 ms^{-1} with a wind direction of 315° was assumed for the maximal wind surge (Witting et al., 2014). According to these boundary conditions, a SWH H_{m_0} of 10 m with a $T_{m_{0,1}}$ of 12 s, a $T_{m_{-1,0}}$ of 13.5 s, and a T_P of 15 s was chosen (Niemeyer, 2001; Niemeyer and Kaiser, 2001). To account for sandbank variability, the event was predicted for all simulated bathymetries and averaged afterwards, resulting in a mean ETD sandbank shape. Due to the simulation method, the synthetic sandbanks do not represent predictions of future morphodynamic developments and therefore do not replace the need to conduct accurate bathymetric surveys. However, they allow for exploration and simulation of various scenarios.

Results are shown in Fig. 4.13 and indicate significant differences in the SWH of the designed reference event along the Norderney coastline. It was found that the high waves coming from the north were uniformly dampened at the northern boundary of the ETD sandbanks, as expected for an averaged ETD bathymetry. Most wave energy affected the island's beach in the northwest and northeast of the study area, with a SWH of up to about 4 m. In between, an area with a SWH up to only about 3.5 m was identified. On the west coast, SWH was measured only up to about 2 m, and the shipping channel showed significantly smaller SWH compared to the surrounding area. The further east, the smaller the SWH became, with a relatively small SWH of about 3 m in the far east. Thus, differences in wave energy were evident in the study area and especially along the coast, even without the availability of bathymetric surveys (e.g., due to high costs) and the assumption of a mean ETD sandbank morphology.

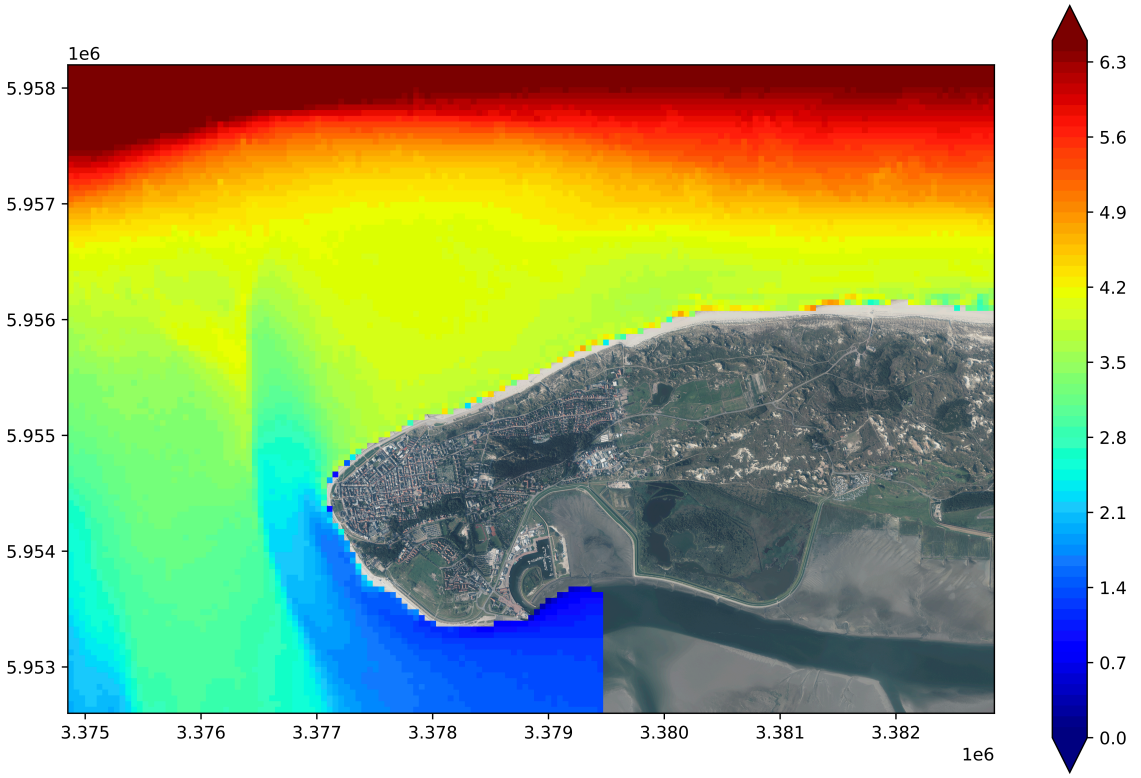

Figure 4.13: *Spatial SWH prediction of a severe event averaged over all simulated bathymetries. All H_{m_0} values are given in m.*

4.4 Conclusions

This study investigated the two-dimensional regional prediction of SWH considering continuously morphodynamically changing ETD sandbanks in the nearshore area of Norderney using a novel mixed-data CNN deep neural network approach. The spatial machine learning model was trained and tested on simulated bathymetries and in-situ measured metocean data from 2004–2011 and 2015–2017, respectively.

For the first time, variogram analysis and random field simulation were used to generate various synthetic spatial ETD sandbank morphologies. Due to the random generation technique, further verification of the applicability of the simulated bathymetries for specific events should be considered in future studies covering the same study area. Accurate bathymetric data are still required for accurate predictions, nevertheless exploration of various morphodynamics has been made available. SWAN-modeled spatial SWH data were used as ground truth for the different mixed-data CNN models tested in the experiments, with hyperparameters tuned by a Bayesian optimization approach. The proposed CNN model architecture achieved an RMSE of 0.097 m, an R^2 of 0.977, and a MAAPE of 6.7% on the test data for the two-dimensional SWH prediction, and performed best among other benchmark

models. The prediction accuracy was within the range of similar studies. Regional variations in RMSE were obtained in the forecast area, with the highest RMSE above the ETD sandbanks.

Further validation on 59 in-situ buoy measurements of SWH showed that the accuracy of the proposed CNN model was very similar to that of the SWAN model. The CNN performed with an RMSE of 0.23 m and a MAAPE of 15.3%, while the numerical SWAN computation achieved only a slightly better performance of 0.218 m RMSE and 14.3% MAAPE on the in-situ buoy data. Since SWAN was used as ground truth for the machine learning model, it was not expected that the CNN would outperform the SWAN model on the in-situ measurements. This suggests that the proposed CNN predictive model is a valid alternative to SWAN in the spatial investigation of SWH.

Compared to the commonly used numerical SWAN model, the computational cost of prediction was reduced by a factor of more than 300 000 by the trained CNN model, with prediction times for the entire test data of 4.4 s and 22 755 min for the CNN and SWAN models, respectively. Thus, despite the need to compute SWAN-modeled ground truth data for model training at a specific region of interest, the proposed CNN model based on simulated bathymetries is an efficient two-dimensional surrogate prediction model for SWH. By applying our proposed model on forecasts of boundary conditions and using the current bathymetry, it is now possible to spatially optimize targeted coastal protection operations in the Norderney area for upcoming storm surges in the short term. Although it is valid only for the specified study area, a transfer of the methodological approach to other regions as well is feasible. However, much computational time would be required to model the ground truth data with SWAN, and the assumption of random morphodynamic sandbank changes further limits rapid transferability.

Based on the overall promising results, more research is needed to investigate spatio-temporal forecasts with longer lead times considering time series data with simulated bathymetries, e.g., using a CNN-LSTM model. Considerations of including additional wave parameters like wave periods and wave direction or the utilization of the entire wave spectrum, including processes like sediment transport and dune erosion, would also be promising. Furthermore, more advanced machine learning simulation techniques like conditional generative adversarial networks (cGANs) should be investigated.

5 Synopsis

This chapter investigates the impact and value of the conducted research with respect to their contributions to the improvement of spatio-temporal SWH predictions under particular consideration of coastal morphodynamics. To accomplish this, a general discussion linking all three studies and the state of the art from Section 1 is performed. Overall conclusions are drawn and an outlook for future research directions is given. A summary of the objectives and major findings of this dissertation is provided in Chapter 6.

5.1 Discussion

The following section discusses the three overarching research questions a–c introduced in Section 1.4. A linkage of all studies is given in each subsection, as well as a general discussion of the central results, contributions, and limitations that go beyond the discussions of the respective papers.

5.1.1 Coastal morphodynamics, ETD variability, and damping effect

ETDs are vital morphological features in many coastal areas around the world. They are a natural coastal protection by sea state damping, provide huge sand reservoirs for nourishment, form shipping channels for navigation, and influence coastal change (Fitzgerald et al., 1984; Pearson et al., 2022). The morphodynamics of ETDs and their sandbanks are mainly determined by the power of tides and waves, while the tidal prism, which is linked to the total sediment volume, as well as local geology, anthropogenic modifications, and sediment supply also affect the morphodynamics (Ford and Dickson, 2018). Since these drivers lead to significant variability in the coastal morphodynamics, a quantification (c.f. Chapter 2) and simulation (c.f. Chapter 4) of the ETD sandbank variability in the nearshore area of Norderney was conducted in this dissertation. A mean velocity of 435–491 my^{-1}, depending on the plates position, was obtained in Chapter 2, which is in the upper range of previous studies (c.f. Section 1.3.3). For further reference, Bezzi et al. (2021) reconstructed a 200-year period of morphodynamic evolution of the Mula di Muggia Bank in the northern Adriatic, Italy. Based on historical cartography, topographic maps, aerial photos, and topo-bathymetric surveys, a migration rate of 12.6 my^{-1} was derived, indicating a much faster migration of the ETD sandbanks in the southern North Sea.

Due to the absence of observations at a high temporal resolution over long periods (c.f. Section 5.1.2), long- and short-term evolution of ETD morphodynamics have

yet been studied by only a few researchers on a global scale (Pearson et al., 2022; Zhang et al., 2020). Elias and van der Spek (2006) analyzed a unique 400-years morphodynamic time series of the Texel Inlet and its ETD. They found a significant influence of human activity (e.g. coastal defense structures) on the ETD evolution. Dallas and Barnard (2011) figured out that the San Francisco Bar, the ETD at the opening of the San Francisco Bay estuary, changed significantly over time and suffered from sand loss in recent years, caused by anthropogenic activities. Due to SLR and other hazards (c.f. Section 1.2), the East Frisian Islands also suffer from sand loss, which is why sand nourishments are carried out (c.f. Section 1.2.3). Zhang et al. (2020) used a half-century time series of remote sensing data to investigate the evolution of the Xincun Lagoon ETD in Hainan Island, China. Considering these worldwide case studies, it becomes clear that potential risks will increase in the future not only for the East Frisian North Sea, but for most coastal regions in the world. Since only few examples of local studies exist, it is important to perform such analyses on a broader variety of coastal areas that take local variations and the complex interactions of the different influencing factors as described above into account. Also, the coastal response and adaption to SLR is assumed to variate on a local scale (Bezzi et al., 2021; Ford and Dickson, 2018; Zhang et al., 2020).
Predictions of future morphodynamics and ETD evolution in the nearshore area of Norderney were not yet taken into account due to the complex interactions and the few measurements available. Processes still remain in the ETD development, evolution, and interaction with adjacent beaches that are not yet fully understood (Harrison et al., 2017). For example, even single severe storm surges can lead to rapid morphodynamic changes of ETD sandbanks (Castelle et al., 2007; Dallas and Barnard, 2011; Wang et al., 2012). However, the mean eastward migration rate can be used as an estimate for future evolution (c.f. Chapter 2). The geostatistical ETD simulation in Chapter 4 is therefore the best way to consider possible future ETD sandbank evolution, without the need to fully understand the complex processes behind. Alternatives to the hydrographic survey data from Chapter 2 are presented in the discussion Section 5.1.2.

Besides the morphodynamic variability of the ETD sandbanks and their evolution over time, the impact of these variations on the nearshore sea state is of particular interest and is the focus of Chapter 2. Since the sediment transport of ETDs and, thus, their influence on the nearshore wave climate, is highly complex and non-linear, numerical modeling using SWAN had to be performed to isolate the morphodynamic signal. A mean sea state damping effect, defined as the ratio of

wave heights seaward and coastside of the ETD sandbanks, of about 41 ± 11% for different representative coastal morphodynamics (bathymetries) and metocean boundary conditions was found for the Norderney nearshore area. The variability of the damping effect along the different bathymetries was 14 ± 2% on average, indicating a strong correlation of the damping effect to the shape, size, and volume of the ETD sandbanks (c.f. Chapter 2). This also implies that the damping effect is difficult to predict due to its variation, but considering coastal morphodynamics is essential for accurate predictions and modeling of the sea state. The pessimistic assumption of adverse bathymetries for future developments is therefore often used. Similar to this research, Ambrosio et al. (2020) analyzed the influence of the ETD morphology on incoming waves in the Cananéia Inlet, São Paulo, Brazil. They used the numerical model MIKE 21 Spectral Waves with four different bathymetries from 1939–2004 to isolate decadal morphological impacts on the wave power distribution, analyzed on eight different incident wave events. Ambrosio et al. (2020) also observed wave energy dissipation at the ETD by refraction and breaking waves. Similar to the conclusion of Chapter 2 and Chapter 3, they also proposed to consider ETD sandbanks for future coastal protection decisions regarding sea state erosion patterns.

Furthermore, it was found in Chapter 2 that the damping effect increased with increasing incident waves, which is beneficial for coastal protection, focusing on severe events. Westerly winds and low water levels also increased the damping effect, while no reliable inferences can be made for changes in the wind direction and SLR counteracts this increase.

5.1.2 Bathymetric data and observation methods for improving predictive sea state modeling

The crucial importance of ETD morphodynamics for proper sea state predictions and coastal protection in the East Frisian North Sea due to its high natural variability and sea state damping was discussed in Section 5.1.1. Based on this, for the first time ML models considering bathymetric data for the SWH reconstruction and prediction in the nearshore area of Norderney were designed (c.f. Section 3). Different algorithms such as random forest (RF), support vector regression (SVR), multiple linear regression (MLR), FFNN, and LSTM as well as a parallel LSTM structure (P-LSTM) were compared, while the LSTM and P-LSTM performed best (Tab. 3.4, Tab. 3.5). Including bathymetric data improved the reconstruction performance of the LSTM and P-LSTM by 16.7% and 12.7% to RMSEs of 0.07 m and 0.069 m, respectively. Also the short-term predictions were improved by including bathymetric

data. Hence, the minimization of the prediction error by considering bathymetries is considerable compared to the respective influences of wind and water level (c.f. Section 3.3.4). Even though extreme events were satisfactorily predicted, a tendency to underestimate high SWHs was observed (c.f. Section 3.3.1). Due to their low frequency in the training data, it is difficult for most ML models to accurately fit extreme events (Qi and Majda, 2019).
Since the models considered highly dynamic morphology, their training relies greatly on accurate and representative bathymetric data (Salameh et al., 2019). However, only few and infrequently sampled observations of the ETD morphodynamics in the nearshore area of Norderney and other ETDs in general exist (Harrison et al., 2017). This insufficient availability of bathymetric data can lead to inaccuracies in the predictions.

Traditionally, single- or multibeam echo-sounder hydrographic surveys are performed by ship to obtain bathymetric data with sonar. For example, the Norderney nearshore area was hydrographically surveyed perpendicular to the depth contours with a sampling rate of 3–10 kHz and different spatial resolutions (up to 5 m) in 1950, 1995, 2005–2008, and 2015 (Bremermann and Meyer, 2012; Meyer, 2014). The entire dataset was used in this dissertation. Although ship-based hydrographic surveys provide very accurate bathymetry retrieval at a high spatial resolution of up to 0.05 m, these missions suffer from limited spatial coverage and low temporal resolution sampling (Salameh et al., 2019). Due to logistical constraints such as the need of low water and calm wind conditions with small sea states, hydrographic surveying is time- and cost-intensive (Bremermann and Meyer, 2012; Spicer et al., 2019). To track rapid morphodynamic changes of the ETD (e.g. impacts of storm surges), high temporal resolution observations are fundamental. Furthermore, there is a great need for accurate and up-to-date bathymetric data for morphodynamic modeling, flood and storm surge forecasting, and coastal management (Pearson et al., 2022; Salameh et al., 2019).

Remote sensing offers alternative methods for surveying bathymetries with ground-based, airborne, and spaceborne approaches and counteracts the problem of insufficient temporal resolution and spatial coverage of in-situ hydrographic surveyed data (Harrison et al., 2017; Lubac et al., 2022; Rogowski et al., 2018). One ground-based approach are monitoring methods like video observations, which are adjustable, easy to install, and provide continuous data automatically (Aarninkhof et al., 2003; Gallop et al., 2009; Holman and Stanley, 2007). Harrison et al. (2017) used video

observations of depth-limited wave breaking at the Raglan ETD at the west coast of New Zealand to identify morphodynamic features over a period of five years. Two video cameras with overlapping views capturing half-hourly averaged RGB-images were installed. The pixel intensity of depth-limited wave breaking allowed for the calculation of wave energy dissipation, which in turn was used to infer the position of shallow sandbanks. No significant correlation of inter-annual morphodynamics and environmental parameters of waves and tides was obtained by Harrison et al. (2017), because some influences or complex interactions were probably not considered. This complexity has also been observed in the analysis of the Norderney ETD (c.f. Sections 2.3 & 5.1.1).
Another high-spatial and high-temporal remote sensing observation technique for ETD mapping is the ground-based coastal radar (Bell et al., 2016; Bird et al., 2017). Rogowski et al. (2018) investigated X-band radar (frequency of 8–12 GHz) for rapid and high-resolution mapping of a shallow ETD and morphodynamics from ground-based and ship-based platforms in the New River Inlet, North Carolina, USA. By averaging the signal response over a specific period (e.g. 5–20 min.), a proxy for underlying bathymetry and shoals was given. Again, best results were obtained for tidal-dominated small sea states (Rogowski et al., 2018). Bird et al. (2017) also used a radar for a waterline elevation mapping for bathymetry retrieval and morphodynamic evolution monitoring. The NLWKN successfully applied this approach with a ground-based radar platform set up for test purposes on the western beach of Norderney as well (NLWKN, 2019). However, the temporal coverage is currently low, so the data were not included in this work.

Besides ground-based remote sensing, air- and spaceborne earth observation (EO) methods are used for bathymetry mapping. Laser scan (LiDAR) flights are periodically conducted to measure the heights of dry areas (topography), because this method is not able to detect data under the sea surface. Airborne LiDAR provides high spatial resolution data, but, similar to echo-sounder, lacks from rather small spatial coverage and low temporal frequency sampling due to costly data acquisition for the investigators by flight missions (Lubac et al., 2022; Salameh et al., 2019). In 2010, the ETD of the Norderney research area was surveyed by airborne missions, while hydrographic surveys of the deeper channels had to be performed timely after the flights due to the laser scan limitations (Meyer, 2014). Data of these cost-intensive mission were used in all studies of this dissertation.
Unmanned aerial vehicles (UAV) introduce a highly automated low-cost airborne EO method with high-resolution acquisition. Nevertheless, the disadvantage of UAVs is their very small spatial coverage due to the limitations of batteries, flight speed,

and weather conditions, as well as regulations and flight restrictions (Salameh et al., 2019).

Spaceborne EO for bathymetry observation is a partial alternative to airborne laser scan flights. Ford and Dickson (2018) used 17 years of records of cloud-masked multispectral Landsat 7 and Landsat 8 satellite images from the Manukau ETD near Auckland, New Zealand, to observe persistent depth-limited wave breaking. Similar to the ground-based video observations, time-averaged scenes were used to classify into broken and unbroken water classes by thresholding and to calculate morphodynamic features from the frequency of the detected wave breaking. Lubac et al. (2022) compared different spaceborne remote bathymetry mapping methods and their uncertainties by using multispectral satellite data of Landsat 7, Landsat 8, and Sentinel-2 A/B at the Arcachon Bay inlet, France from 2013–2020. Compared to echo-sounder ground truth data, they obtained the best performance and highest robustness with a cluster-based approach using a green band log-linear regression model. However, uncertainties of satellite-based bathymetry observations have been poorly studied so far (Lubac et al., 2022).

In general, satellite imaging has the ability to acquire large areas for morphodynamic evolution monitoring with high temporal frequency at low-cost for the investigators due to an unprecedented free data access (Salameh et al., 2019; Stewart et al., 2016). Spatial resolution, on the other hand, is limited e.g. to 30 m for Landsat 8 and 10 m for Sentinel-2 A/B. High-resolution bathymetry mapping with the required vertical accuracy, such as with ship-based echo-sounder, is a major limitation of multispectral satellite sensors due to their inability to observe below the sea surface (Ford and Dickson, 2018; Salameh et al., 2019).

In addition to passive multispectral satellite sensors, active radar satellites, such as Sentinel-1, also observe the Earth (Salameh et al., 2019). Stewart et al. (2016) used high-resolution Sentinel-1 synthetic aperture radar (SAR) data and a new inversion algorithm developed by Renga et al. (2013, 2014) to derive the bathymetry of a southern North Sea region.

Future new multispectral, hyperspectral and radar satellite missions will significantly increase the potential of spaceborne bathymetry observation and substantially reduce the current problem of low spatial resolution (Lubac et al., 2022; Salameh et al., 2019).

5.1.3 ML vs. numerical modeling – Spatial wave predictions using synthetic and real bathymetry data

Physics-based numerical wave models for the nearshore area can cope with sparse and cost-intensive spatio-temporal in-situ wave measurements due to their physi-

cal representation (c.f. Section 1.3.1). Besides WAM, WAVEWATCH, and SWAN, XBeach is commonly used for nearshore wave simulation (Roelvink et al., 2009). For precise modeling, physics-based models require an accurate consideration of initial states and boundary conditions of the model domain. Often, noisy or even missing boundary conditions are found in real-world physical problems (ill-posed problem), which makes it challenging to achieve good simulations even though precise physical equations exists (Karniadakis et al., 2021; Wang et al., 2022). Generating cumbersome computational grids for complex study areas like the ETD of the East Frisian North Sea are time-consuming and prone to errors. Thus, physics-based numerical models are not suitable for immediate predictions of wave fields, such as now-casting for marine operations or storm surge protection (Chen et al., 2021a).
However, soft-computing methods such as ANNs are not suitable to fully replace those physics-based numerical models, which have been successfully applied for decades and are still under research (c.f. Section 1.3.2). But ANNs can be used as surrogates to reduce computational costs by handling non-linear and high-dimensional data-driven problems (Wang et al., 2022). Feng et al. (2020) obtained an acceleration of up to a factor of 20 000 with a FFNN compared to a SWAN model in the area of Lake Michigan, USA. James et al. (2018) also used a FFNN for spatial wave field predictions in the area of Monterey Bay, USA and achieved a speed-up by a factor of 1 000 compared to SWAN, while Chen et al. (2021a) designed a RF model to reduce the computational time by a factor of 100 compared to SWAN in a study area of Cornwall, UK. Thus, a few studies exist for oceanic coastal areas, as well as lakes, which provide a surrogate for numerical simulation with SWAN for a given study area. As figured out in Section 4, a massive speed-up by a factor of more than 300 000 was achieved in this dissertation by the proposed CNN surrogate compared to SWAN modeling. This factor is very high compared to the other studies, supposedly showing the efficiency of the CNN model for spatial predictions used in this work.
Wei and Davison (2022) also developed a CNN surrogate model but only achieved a speed-up by a factor of 22.5. However, they used SWASH rather than SWAN as ground truth, hence a comparison is more critical.

Albeit the significant speed-up of the proposed ML approach developed in this dissertation, the surrogate model relies on the SWAN modeled ground truth data. Thus, it must be taken into account, that the speed-up only applies after the cost-intensive numerical modeling for new and rapid predictions in the study area. Nevertheless, the predictions on new data in the study area were performed significantly faster by ML, providing an accurate real-time operational surrogate model for exploring

various sea state and morphodynamic scenarios. Validation on 59 in-situ buoy and metocean measurements covering the period 2004–2017 showed similar performance of the proposed CNN and SWAN model with an RMSE of 0.23 m for the CNN and 0.218 m for SWAN, respectively (c.f. Section 4.3.2).
Applying the approach to other areas, new ground truth modeling is required. Thus, a challenging drawback of using ANNs for spatial predictions of SWH is the need for a huge amount of training data for accurate predictions when data availability is limited (c.f. Section 5.1.2), leading also to a lack of generality for unseen data (out-of-sample cases). Additionally, the problem that physics are not explicitly learned during ML training and results may be inconsistent with the laws of physics is even more challenging (Jia et al., 2019; Wang et al., 2022).

To overcome this limitation of traditional ML methods, some recent studies combined physics-based knowledge with data-driven ML methods (Chen and Sun, 2020; Chen et al., 2021c, 2022; Willard et al., 2020; Zhang et al., 2019). A common method are so-called physics-informed neural networks (PINN) (Raissi et al., 2019). While Jin et al. (2021) proposed a PINN which learned the Navier-Stokes equation of fluid flows (e.g. ocean waves), Wang et al. (2022) published a study about integrating physics of nearshore waves coming from the numerical model XBeach to a ML algorithm for wave field prediction. The advantage of PINNs is their ability to solve ill-posed problems by the consideration of physics-based knowledge during the training, which compensates for any insufficient training data and leads to robust and physics-consistent prediction results. The performance of PINNs for wave height prediction is determined by the user's decision of how many and which physical laws (e.g. diffraction, refraction) should be considered during the training.

Even for PINNs, their still remain some general concerns regarding the optimization of the solution of ANNs. As for traditional ANNs, non-convex optimization algorithms are used and local-minimal solutions instead of global-minimal solutions will be computed (Sun et al., 2020). Also the amount of training data still plays an essential role for the ML model performance (Wang et al., 2022). Depending on the spatial extent of the study area, the temporal period of prediction, and the physical parameters, sampling and generation (e.g. simulated synthetic bathymetries) of ground truth data can be very time-consuming and cost-intensive.

Another approach to address the problem of limited generality of the ANNs is transfer learning, which may be a promising method to reuse trained ANNs for the initialization of other ANNs and save computational time for new domains. Transfer

learning requires a certain similarity of the boundary conditions (e.g. bathymetry) between the pre-trained and the new model (Wang et al., 2022). Since bathymetric data were trained within the model presented in this dissertation for the first time, there is no need to train additional networks for the study area of Norderney for changes in the bathymetry. This provides a great cost and time saving for this domain. However, transferability to other areas of the East Frisian North Sea is still not straightforward, although the proposed models of this dissertation could be used for transfer learning. The methodological approach, on the other hand, is generalizable to many regions.

Section 4 focused on the two-dimensional prediction of regional SWH considering morphodynamics with a mixed-data deep CNN (c.f. Section 1.3.2). Therefore, for the first time, synthetic bathymetries were simulated by geostatistical variogram analysis and random fields simulation. The quality and accuracy of the resulting generated spatial ETD sandbanks morphologies relies on the methodological approach and was qualitatively validated by experts of the NLWKN. Except for a general eastward migration of the ETD sandbanks (c.f. Section 2), the complex interactions and storm surge impacts prevent accurate morphodynamic predictions, as discussed in Section 5.1.1. Therefore, the random-based generation approach discussed in Section 4.2.3 was suitable and has made possible the exploration of various morphodynamic scenarios without the need of cost-intensive bathymetric surveys. However, accurate bathymetries are still required for accurate predictions of specific events (Chen et al., 2021a).

A closer look at the accuracy of the single-point prediction (c.f. Section 3) and the spatial model (c.f. Section 4) revealed that the performance was similarly high for both approaches, indicating that both methods extract the maximum information from the data: While the proposed single-point model had an RMSE of 0.069 m, R^2 of 0.965, and MAAPE of 9.3% at position Coast II, the proposed spatial CNN model reached an RMSE of 0.230 m, R^2 of 0.704, and MAAPE of 15.3% at the same position. The spatial averaged performance of the CNN was an RMSE of 0.097 m, R^2 of 0.977, and MAAPE of 6.7%. Since a similar dataset was used for both models, the scale-dependent RMSE can be compared here. It was expected that the specialized single-point model fits the prediction at position Coast II better than the spatial model, which is strongly influenced by the ETD sandbanks. The highest RMSE of the spatial prediction was observed over the ETD sandbanks, likely caused by the complex wave interactions and refraction (c.f. Section 4.3.1). Additionally, a huge variability of the bias error was found along the coastline, mainly due to the high

variability of the predicted sea state events, simulated bathymetries, and metocean boundary conditions (c.f. Section 4.3.1). Therefore, using these three sources of information will be very important for precise predictions.

5.2 Conclusions

The aim of this dissertation was to contribute to the improvement of spatial and temporal predictions of nearshore wave heights with neural networks considering coastal morphodynamics. For the first time, simulated and surveyed coastal bathymetry data were included in ML predictions, providing an efficient surrogate to common numerical modeling with SWAN. The focus was set on ML model development and optimization for temporal predictions with LSTM neural networks and spatial predictions with deep mixed-data CNNs, both considering coastal morphodynamics. Therefore, the highly variable coastal morphodynamics of the East Frisian North Sea research area of Norderney were studied in detail for a better understanding of the offshore ETD sandbanks and their influence on nearshore sea states. Based on this, it was possible to simulate synthetic bathymetries with respect to the almost random morphological dynamics of the ETD sandbanks.
The research questions a–c from Section 1.4 are answered by the research conducted in Sections 2, 3, and 4.

Regarding the quantification of the influence of coastal morphodynamics on the nearshore sea state (RQ-a), great variability in the morphological conditions of the ETD sandbanks and their influence on the sea state damping was found for the study area. By proposing validated numerical SWAN modeling, a systematic quantification of the isolated influence of the ETD sandbanks was conducted for the first time. With a mean sea state damping effect of 41 $\pm$ 11% for different modeled metocean boundary conditions, a critical impact was found. Analysis of in-situ buoy-measured data revealed a mean damping effect of 37 $\pm$ 18%, which was within the range of the numerical model. Variations of the damping effect up to 14 $\pm$ 2% due to five different isolated ETD sandbank conditions were modeled for storm events at tidal high water. Therefore, it can be concluded that the evolution of the ETD sandbanks plays a crucial role in nearshore sea state conditions as a natural coastal protection by sea state damping. Besides eastward migration with an observed mean velocity of 435–491 my^{-1}, the highly complex and almost random variability of ETD morphology is not yet fully understood. Thus, it is proposed for coastal management to account for pessimistic ETD sandbank evolution scenarios with changing natural damping effects.

Since variability in coastal morphodynamics was found to exert a huge impact on nearshore wave climate, RQ-b focused on using this finding to improve single-point predictions of SWH with ANNs. For the first time, LSTM neural networks were applied to shallow waters with great morphodynamic variability. By including a transect of bathymetric features of the ETD sandbanks, the single-point reconstruction and short- and long-term prediction of SWH with the proposed LSTM and an ensemble of parallel LSTM (P-LSTM) models was significantly improved by 16.7% and 12.7%, respectively, in terms of RMSE. All models used were improved by including bathymetric data, with substantial effects compared to wind and water level. SWH reconstruction achieved robust performance with RMSEs of 0.07 m and 0.069 m for the LSTM and P-LSTM, respectively. As the accuracy was improved by including bathymetric data, consideration of natural coastal morphodynamics is once again crucial for coastal protection and further sea state predictions in the nearshore area of the East Frisian North Sea. The good reconstructions and predictions provide the potential to optimize the cost-intensive operation of the wave buoys or to completely terminate single positions.

Based on the major finding that consideration of coastal morphodynamics improves nearshore sea state predictions, RQ-c aimed to transfer the approach to spatial predictions of SWH. A two-dimensional mixed-data CNN wave height prediction model was developed for the first time using continuously morphodynamically changing ETD sandbanks. Since only few bathymetric data were available, synthetic sandbanks were simulated using the geostatistical methods of variogram analysis and random fields for the first time. Trained on SWAN-modeled ground truth data, the proposed CNN surrogate model performed with a spatial averaged RMSE of 0.097 m and thus was within the range of other studies. Highest RMSEs were found above the ETD sandbanks, probably due to the high morphodynamic changes and complex wave interactions. Further validation with 59 in-situ buoy measurements revealed similar accuracy of the CNN (RMSE = 0.23 m) to the SWAN-modeled ground truth data (RMSE = 0.218 m). Once trained, the CNN model reduced the computation time by a factor of more than 300 000 compared to the numerical SWAN model. However, the CNN training required SWAN-modeled ground truth data. Since synthetic coastal morphodynamics were used, accurate bathymetries are still required for accurate predictions. Nevertheless, the proposed surrogate model enabled real-time, spatial high-resolution forecasting of SWH for e.g. storm surge events and the exploration of various synthetic morphodynamic and metocean scenarios for coastal protection, such as climate-related SLR, wind speed increase, or a pessimistic ETD morphology.

5.3 Outlook

Future research may benefit from the ML and geostatistical methods developed in this dissertation as well as the main finding of prediction improvement by including coastal morphodynamics in spatio-temporal modeling of SWH with ANNs. While the morphodynamic influence of the ETD sandbanks is determinant for Norderney, other influences may be crucial in other regions, which is why locally adapted models are more reliable and accurate than global models. Albeit this dissertation advanced the scientific understanding of the influence of coastal morphodynamics on the nearshore sea state in the East Frisian North Sea, still knowledge gaps exist. Further investigations of the complex dependencies and processes of coastal morphodynamcis and their interactions with waves and currents are required for an advanced understanding of the system, which allows for a better parameterization of numerical wave models as well as generalization of data-driven ML methods (Pearson et al., 2022; Zhang et al., 2020). Sediment transport was not studied in detail in this work, but is also essential for understanding and modeling coastal morphodynamics. Since traditional bathymetric surveys are time- and cost-intensive, advanced acquisition methods for sufficient high-temporal and high-spatial resolution data are needed. Harrison et al. (2017) introduced remote sensing video observations for bathymetry mapping, which might be also implemented on Norderney, where appropriate tall buildings for mounting the cameras are located near the beach. Expanded research and case-studies on coastal radar for bathymetry observations as conducted by Rogowski et al. (2018) and Bird et al. (2017) seem very promising. Finally, the spaceborne EO approaches discussed in Lubac et al. (2022) also provide great potential. For a better representation of the physical processes behind, physics-informed neural networks (PINN) may play a major role, as well as transfer learning in light of an efficient model transfer to other regions (Wang et al., 2022). Additionally, it would be of interest to consider further wave parameters like wave period and direction, or to use the entire wave spectrum as an input for ML, as in numerical models. Optimizing network architectures to better handle spatial mixed-data time series may lead to longer lead time of the predictions. For example, Chen et al. (2021b) investigated combined CNN-LSTM neural networks for spatial wind speed forecasting, which could be transferred to spatial SWH forecasting including the use of synthetic bathymetries. In order to improve ETD sandbank simulations for situations where no proper bathymetric observations are available for model training, sophisticated methods such as conditional generative adversarial networks (cGANs) can be applied (Willard et al., 2020).

New data and methods will enhance future research, while remaining mandatory due to climate-related challenges in coastal protection.

6 Summary

Coastal regions are particularly affected by climate change impacts such as SLR and changing storms. Accurate predictions of SWH are therefore essential for coastal protection and marine disaster prevention. Especially for the nearshore areas of the East Frisian North Sea, proper understanding of the complex coastal morphodynamics and ETD sandbank evolution is crucial for the consideration of bathymetric data in ML models. Moreover, by providing ANN surrogate models to the common numerical wave model SWAN, computational time can be drastically reduced for real-time spatial wave height forecasts, as required for e.g. storm surge events.

The major objective of this dissertation was therefore to improve spatial and temporal ML predictions of nearshore wave heights in the research area of Norderney, Germany by considering coastal morphodynamics, as well as a better understanding of their influence on the sea state.
The following three overarching open research questions were identified and addressed within this work:

RQ-a. How can the influence of coastal morphodynamics and its variability on the nearshore sea state be quantified?

RQ-b. How can data of coastal morphodynamics improve the single-point reconstruction and prediction of significant wave height by neural networks?

RQ-c. How can simulated bathymetries be used to predict two-dimensional significant wave height with convolutional neural networks?

Three studies were conducted that addressed these research questions respectively, focusing on spatial and temporal predictions and the role of coastal morphodynamics. The main results of this work are summarized below:

RQ-a. How can the influence of coastal morphodynamics and its variability on the nearshore sea state be quantified?

- Validated SWAN numerical models were used to quantify the influence of ETD sandbanks on the nearshore sea state. A significant impact was found in the sense of a mean modeled sea state damping effect of 41 $\pm$ 11 % for different metocean boundary conditions.
- A mean buoy-measured damping of the sea state through the ETD sandbanks was measured of 37 $\pm$ 18 % and was thus within the range of the numerical model.

- By isolating the morphodynamic influence of five different ETD sandbank states, variations of the damping effect up to 14 ± 2 % were found for storm events at tidal high water. The evolution of the ETD sandbanks as natural coastal protection therefore plays a critical role in wave heights along the coastline.
- With increasing incoming wave height, the damping effect increased. Consequently, especially during storm surges, the ETD sandbanks provide a great protection for the island.
- Besides an eastward migration of the ETD sandbanks with a mean velocity between 435–491 my^{-1}, depending on the plates position, a high variability of their morphology was found, which leads to the conclusion that changes can be considered as almost random. Even single storm surges can lead to strong changes in morphology.

RQ-b. How can data of coastal morphodynamics improve the single-point reconstruction and prediction of significant wave height by neural networks?

- LSTM neural networks were applied to shallow waters with great morphodynamic influences for the first time. A transect of bathymetric features was included to account for coastal morphodynamics and therefore improve single-point predictions of SWH.
- The proposed LSTM and an ensemble of parallel LSTM (P-LSTM) model performed best among other ML benchmarks. They showed robust results with RMSEs of 0.07 m and 0.069 m for the reconstruction of SWH, respectively.
- Including bathymetric data to the LSTM and P-LSTM models improved their performance by 16.7% and 12.7% in terms of RMSE, respectively. All model predictions were improved by including bathymetric data, with substantial levels of improvement compared to the effects of wind and water level.
- The reconstruction and short- and long-term predictions were also performed on deep ocean NOAA buoys for further validation – with satisfactory results. Extreme events were also predicted well.
- As the accuracy was improved by including bathymetric data, the consideration of the natural sea state damping and its dynamic behavior due to changing ETD sandbanks is essential for coastal protection.

RQ-c. How can simulated bathymetries be used to predict two-dimensional significant wave height with convolutional neural networks?

- A two-dimensional mixed-data CNN wave height prediction model considering the continuously morphodynamically changing ETD sandbanks was developed for the first time.
- For the first time, synthetic sandbanks were generated using the geostatistical methods of variogram analysis and random fields.
- The proposed CNN surrogate model had a spatially averaged RMSE of 0.097 m. Highest RMSEs were found above the ETD sandbanks probably due to the high morphodynamic changes and complex wave interactions.
- Compared to the ground truth SWAN data (RMSE = 0.218 m), the CNN achieved a similar performance (RMSE = 0.23 m) on in-situ buoy measurements.
- The trained CNN was able to reduce the computational time by a factor of more than 300 000 compared to the numerical SWAN model. However, the training of the CNN model required SWAN modeled ground truth data. Applications include real-time forecasting of wave height and the exploration of various metocean and morphodynamic scenarios for coastal protection, e.g. SLR. For accurate predictions still accurate bathymetries are required.

The above findings of this dissertation advanced the scientific understanding of coastal morphodynamics regarding the influence of ETD sandbanks on the nearshore sea state and their indispensable consideration in predictive modeling. Future research may also benefit from the methodological approaches of ANNs for wave height prediction using bathymetric data and the geostatistical simulation of synthetic bathymetries. Operational coastal protection applications may profit from the trained models and described results, as exploration of various climate and morphodynamic scenarios as well as real-time predictions of high-resolution coastal wave heights has now been made possible.

Based on the synopsis, a promising path for future advances in ML methods with consideration of coastal morphodynamics for sea state predictions emerges. For further improved representation of physical processes, PINNs may play a major role, while CNN-LSTM architectures and transfer learning can help to achieve longer lead times and an efficient transfer to other regions or other wave parameters like wave period and direction. Since bathymetric data are time- and cost-intensive, sophisticated methods such as spaceborne EO, remote sensing video observations, or coastal radar are required to better describe, quantify, and understand the complex coastal morphodynamics.

Bibliography

Aarninkhof, S. G., Turner, I. L., Dronkers, T. D., Caljouw, M., and Nipius, L.: A video-based technique for mapping intertidal beach bathymetry, Coastal Engineering, 49, 275–289, doi:doi.org/10.1016/s0378-3839(03)00064-4, 2003.

Abed-Elmdoust, A. and Kerachian, R.: Wave height prediction using the rough set theory, Ocean Engineering, 54, 244–250, doi:doi.org/10.1016/j.oceaneng.2012.07.020, 2012.

Abramowitz, M. and Stegun, I. E.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, U.S. Department of Commerce, National Bureau of Standards, Washington DC, 10 edn., 1972.

Aggarwal, C. C.: Neural Networks and Deep Learning, Springer, Cham, Switzerland, doi:doi.org/10.1007/978-3-319-94463-0, 2018.

Akpinar, A., Bingolbali, B., and Van Vledder, G. P.: Wind and wave characteristics in the Black Sea based on the SWAN wave model forced with the CFSR winds, Ocean Engineering, 126, 276–298, doi:doi.org/10.1016/j.oceaneng.2016.09.026, 2016.

Ali, M. and Prasad, R.: Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition, Renewable and Sustainable Energy Reviews, 104, 281–295, doi:doi.org/10.1016/j.rser.2019.01.014, 2019.

Altunkaynak, A.: Prediction of significant wave height using geno-multilayer perceptron, Ocean Engineering, 58, 144–153, doi:doi.org/10.1016/j.oceaneng.2012.08.005, 2013.

Ambrosio, B. G., Sousa, P. H., Gagliardi, M. H., and Siegle, E.: Wave energy distribution at inlet channel margins as a function of ebb tidal delta morphology: Cananéia Inlet, São Paulo, Brazil, Anais da Academia Brasileira de Ciências, 92, doi:doi.org/10.1590/0001-3765202020180677, 2020.

Anthony, E. J.: Storms, shoreface morphodynamics, sand supply, and the accretion and erosion of coastal dune barriers in the southern North Sea, Geomorphology, 199, 8–21, doi:doi.org/10.1016/j.geomorph.2012.06.007, 2013.

Awad, M. and Khanna, R.: Efficient Learning Machines, Apress, doi:doi.org/10.1007/978-1-4302-5990-9, 2015.

Axer, T., Bistry, T., Fietze, S., Müller, M., and Prechtl, M.: Sturmdokumentation Deutschland 2005-2008, Report, Deutsche Rück, Deutsche Rückversicherung AG, 2005-2008.

Bacon, S. and Carter, D. J. T.: Wave Climate Changes in the North-Atlantic and North-Sea, International Journal of Climatology, 11, 545–558, URL WOS:A1991FY33400006, 1991.

Bai, G., Wang, Z., Zhu, X., and Feng, Y.: Development of a 2-D deep learning regional wave field forecast model based on convolutional neural network and the application in South China Sea, Applied Ocean Research, 118, doi:doi.org/10.1016/j.apor.2021.103012, 2022.

Bakker, J. P., Baas, A. C., Bartholdy, J., Jones, L., Ruessink, G., Temmerman, S., and van de Pol, M.: Environmental Impacts—Coastal Ecosystems, in: North Sea Region Climate Change Assessment, pp. 275–314, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_9, 2016.

Balas, C. E., Koc, L., and Balas, L.: Predictions of Missing Wave Data by Recurrent Neuronets, Journal of Waterway, Port, Coastal and Ocean Engineering, 130, 256–265, 2004.

Bell, P. S., Bird, C. O., and Plater, A. J.: A temporal waterline approach to mapping intertidal areas using X-band marine radar, Coastal Engineering, 107, 84–101, doi:doi.org/10.1016/j.coastaleng.2015.09.009, 2016.

Bengtsson, L., Hodges, K. I., and Keenlyside, N.: Will Extratropical Storms Intensify in a Warmer Climate?, Journal of Climate, 22, 2276–2301, doi:doi.org/10.1175/2008jcli2678.1, 2009.

Bennett, W. G., Karunarathna, H., Reeve, E. D., and Mori, N.: Computational modelling of morphodynamic response of a macro-tidal beach to future climate variabilities, Marine Geology, 415, 2019.

Berbić, J., Ocvirk, E., Carević, D., and Lončar, G.: Application of neural networks and support vector machine for significant wave height prediction, Oceanologia, 59, 331–349, doi:doi.org/10.1016/j.oceano.2017.03.007, 2017.

Bergillos, R. J., Lopez-Ruiz, A., Ortega-Sanchez, M., Masselink, G., and Losada, M. A.: Implications of delta retreat on wave propagation and longshore sediment transport - Guadalfeo case study (southern Spain), Marine Geology, 382, 1–16, doi:doi.org/10.1016/j.margeo.2016.09.011, 2016.

Bergstra, J., Bardenet, R., Bengio, Y., and Kégl, B.: Algorithms for Hyper-Parameter Optimization, in: 24th International Conference on Neural Information Processing Systems, pp. 2546–2554, 2011.

Berkenbrink, C. and Niemeyer, H. D.: Changing Salinity Intrusion Due To Estuarine Waterway Deeping - Analysis Based On Artificial Neural Networks, in: IAHR World Congress, 2017.

Berkenbrink, C. and Wurpts, A.: Aktuelle Bemessungspraxis für See- und Ästuardeiche in Niedersachsen, Die Küste, 87, 1–18, doi:doi.org/10.18171/1.087107, 2019.

Berkenbrink, C., Bremm, G., and van der Wel, R.: Messkette Norderney – Jahresbericht 2019 (Datengrundlage 2017/2018), Report, Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz, Forschungsstelle Küste, 2019.

Bezzi, A., Casagrande, G., Fracaros, S., Martinucci, D., Pillon, S., Sponza, S., Bratus, A., Fattor, F., and Fontolan, G.: Geomorphological Changes of a Migrating Sandbank: Multidecadal Analysis as a Tool for Managing Conflicts in Coastal Use, Water, 13, 3416, doi:doi.org/10.3390/w13233416, 2021.

Bird, C. O., Bell, P. S., and Plater, A. J.: Application of marine radar to monitoring seasonal and event-based changes in intertidal morphology, Geomorphology, 285, 1–15, doi:doi.org/10.1016/j.geomorph.2017.02.002, 2017.

Bitner-Gregersen, E. M., Vanem, E., Gramstad, O., Horte, T., Aarnes, O. J., Reistad, M., Breivik, O., Magnusson, A. K., and Natvig, B.: Climate change and safe design of ship structures, Ocean Engineering, 149, 226–237, doi:doi.org/10.1016/j.oceaneng.2017.12.023, 2018.

Bloomfield, P.: Fourier Analysis of Time Series: An Introduction: Second Edition, John Wiley & Sons, inc, New York, Chichester, Weinheim, Brisbane, Singapore, Toronto, 2000.

Booij, N., Ris, R. C., and Holthuijsen, L. H.: A third-generation wave model for coastal regions - 1. Model description and validation, Journal of Geophysical Research: Oceans, 104, 7649–7666, doi:doi.org/10.1029/98jc02622, 1999.

Bottou, L.: Online Learning and Stochastic Approximations, 1998.

Brander, K. M., Ottersen, G., Bakker, J. P., Beaugrand, G., Herr, H., Garthe, S., Gilles, A., Kenny, A., Siebert, U., Skjoldal, H. R., and Tulp, I.: Environmental Impacts—Marine Ecosystems, in: North Sea Region Climate Change Assessment, pp. 241–274, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_8, 2016.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, 2001.

Bremermann, A. and Meyer, C.: Verlagerungsgeschwindigkeiten von Platen im Norderneyer Riffbogen. Technical Report 01/2012, doi:Unpublished\results, 2012.

Brochu, E., Cora, V. M., and de Freitas, N.: A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning, doi:arXiv:1012.2599, 2010.

Cabrera, D., Guamán, A., Zhang, S., Cerrada, M., Sánchez, R.-V., Cevallos, J., Long, J., and Li, C.: Bayesian approach and time series dimensionality reduction to LSTM-based model-building for fault diagnosis of a reciprocating compressor, Neurocomputing, 380, 51–66, doi:doi.org/10.1016/j.neucom.2019.11.006, 2020.

Castelle, B., Bourget, J., Molnar, N., Strauss, D., Deschamps, S., and Tomlinson, R.: Dynamics of a wave-dominated tidal inlet and influence on adjacent beaches, Currumbin Creek, Gold Coast, Australia, Coastal Engineering, 54, 77–90, doi:doi.org/10.1016/j.coastaleng.2006.08.007, 2007.

Cavaleri, L.: Wave Modeling—Missing the Peaks, Journal of Physical Oceanography, 39, 2757–2778, doi:doi.org/10.1175/2009jpo4067.1, 2009.

Chang, C.-C. and Lin, C.-J.: LIBSVM: A Library for Support Vector Machines, Report, National Taiwan University, URL `https://t1p.de/5puc,https://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf`, https://web.archive.org/web/20210224172125/, https://www.csie.ntu.edu.tw/ cjlin/papers/libsvm.pdf, 2021.

Chen, J., Pillai, A. C., Johanning, L., and Ashton, I.: Using machine learning to derive spatial wave data: A case study for a marine energy site, Environmental Modelling and Software, 142, 105 066, doi:doi.org/10.1016/j.envsoft.2021.105066, 2021a.

Chen, Y., Wang, Y., Dong, Z., Su, J., Han, Z., Zhou, D., Zhao, Y., and Bao, Y.: 2-D regional short-term wind speed forecast based on CNN-LSTM deep

learning model, Energy Conversion and Management, 244, doi:doi.org/10.1016/j.enconman.2021.114451, 2021b.

Chen, Z. and Sun, H.: Sparse representation for damage identification of structural systems, Structural Health Monitoring, 20, 1644–1656, doi:doi.org/10.1177/1475921720926970, 2020.

Chen, Z., Chen, B., Dhaene, J., and Yang, T.: Fair dynamic valuation of insurance liabilities via convex hedging, Insurance: Mathematics and Economics, 98, 1–13, doi:doi.org/10.1016/j.insmatheco.2021.01.001, 2021c.

Chen, Z., Liu, Y., and Sun, H.: Forecasting of nonlinear dynamics based on symbolic invariance, Computer Physics Communications, 277, 108 382, doi:doi.org/10.1016/j.cpc.2022.108382, 2022.

Choi, H., Park, M., Son, G., Jeong, J., Park, J., Mo, K., and Kang, P.: Real-time significant wave height estimation from raw ocean images based on 2D and 3D deep neural networks, Ocean Engineering, 201, doi:doi.org/10.1016/j.oceaneng.2020.107129, 2020.

Chung, M. K.: Gaussian kernel smoothing, arXiv, p. 18, doi:doi.org/10.48550/arXiv.2007.09539, 2021.

Church, J. A. and White, N. J.: Sea-level rise from the late 19th to the early 21st century, Surveys in Geophysics, 32, 585–602, doi:doi.org/10.1007/s10712-011-9119-1, 2011.

Cooper, J. and Navas, F.: Natural bathymetric change as a control on century-scale shoreline behavior, Geology, 32, 513, doi:doi.org/10.1130/g20377.1, 2004.

Cornejo-Bueno, L., Garrido-Merchán, E. C., Hernández-Lobato, D., and Salcedo-Sanz, S.: Bayesian optimization of a hybrid system for robust ocean wave features prediction, Neurocomputing, 275, 818–828, doi:doi.org/10.1016/j.neucom.2017.09.025, 2018.

Cowpertwait, P. S. and Metcalfe, A. V.: Introductory Time Series with R, Springer, Dordrecht, Heidelberg, London, New York, doi:doi.org/10.1007/978-0-387-88698-5, 2009.

Dallas, K. L. and Barnard, P. L.: Anthropogenic influences on shoreline and nearshore evolution in the San Francisco Bay coastal system, Estuarine, Coastal and Shelf Science, 92, 195–204, doi:doi.org/10.1016/j.ecss.2010.12.031, 2011.

Datawell: Operating Manual W@ves21 Software for Datawell Waverider Buoys. From VERSION 2.1.12 (RfBuoy 2.1.18), Report, 2014.

Deka, P. C. and Prahlada, R.: Discrete wavelet neural network approach in significant wave height forecasting for multistep lead time, Ocean Engineering, 43, 32–42, doi:doi.org/10.1016/j.oceaneng.2012.01.017, 2012.

Deo, M. and Naidu, C. S.: Real time wave forecasting using neural networks, Ocean Engineering, 26, 191–203, 1999.

Deo, M. C., Jha, A., Chaphekar, A. S., and Ravikant, K.: Neural networks for wave forecasting, Ocean Engineering, 28, 889–898, doi:doi.org/10.1016/S0029-8018(00)00027-5, 2001.

Dhanachandra, N., Manglem, K., and Chanu, Y. J.: Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm, Procedia Computer Science, 54, 764–771, doi:doi.org/10.1016/j.procs.2015.06.090, 2015.

Dietrich, C. R. and Newsam, G. N.: Fast And Exact Simulation Of Stationary Gaussian Processes Throught Circulant Embedding Of The Covariance Matrix, Society for Industrial and Applied Mathematics, 18, 1088–1107, 1997.

Dixit, P. and Londhe, S.: Prediction of extreme wave heights using neuro wavelet technique, Applied Ocean Research, 58, 241–252, doi:doi.org/10.1016/j.apor.2016.04.011, 2016.

Dronkers, J. and Stojanovic, T.: Socio-economic Impacts—Coastal Management and Governance, in: North Sea Region Climate Change Assessment, pp. 475–488, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_19, 2016.

Durán-Rosal, A. M., Hervás-Martínez, C., Tallón-Ballesteros, A. J., Martínez-Estudillo, A. C., and Salcedo-Sanz, S.: Massive missing data reconstruction in ocean buoys with evolutionary product unit neural networks, Ocean Engineering, 117, 292–301, doi:doi.org/10.1016/j.oceaneng.2016.03.053, 2016.

Eldeberky, Y.: Nonlinear Transformation of Wave Spectra in the Nearshore Zone, Ph.D. thesis, 1996.

Elias, E. P. and van der Spek, A. J.: Long-term morphodynamic evolution of Texel Inlet and its ebb-tidal delta (The Netherlands), Marine Geology, 225, 5–21, doi:doi.org/10.1016/j.margeo.2005.09.008, 2006.

Eshleman, J. L., Barnard, P. L., Erikson, L. H., and Hanes, D. M.: Coupling alongshore variations in wave energy to beach morphologic change using SWAN wave model at Ocean Beach, San Francisco, CA, 10th International Workshop on Wind Hindcasting and Forecasting and Coastal Hazard Symposium: North Shore, Oahu, November 11-16, 2007, 2007.

Fan, S., Xiao, N., and Dong, S.: A novel model to predict significant wave height based on long short-term memory network, Ocean Engineering, 205, doi:doi.org/10.1016/j.oceaneng.2020.107298, 2020.

Feng, X., Ma, G., Su, S.-F., Huang, C., Boswell, M. K., and Xue, P.: A multi-layer perceptron approach for accelerated wave forecasting in Lake Michigan, Ocean Engineering, 211, doi:doi.org/10.1016/j.oceaneng.2020.107526, 2020.

Fernández, J. C., Salcedo-Sanz, S., Gutiérrez, P. A., Alexandre, E., and Hervás-Martínez, C.: Significant wave height and energy flux range forecast with machine learning classifiers, Engineering Applications of Artificial Intelligence, 43, 44–53, doi:doi.org/10.1016/j.engappai.2015.03.012, 2015.

Fitzgerald, D. M., Penland, S., and Nummedal, D.: Control Of Barrier Island Shape By Inlet Sediment Bypassing: East Frisean Islands, West Germany, Marine Geology, 60, 355–376, 1984.

Ford, M. R. and Dickson, M. E.: Detecting ebb-tidal delta migration using Landsat imagery, Marine Geology, 405, 38–46, doi:doi.org/10.1016/j.margeo.2018.08.002, 2018.

Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V. W., Dangendorf, S., Hogarth, P., Zanna, L., Cheng, L., and Wu, Y. H.: The causes of sea-level rise since 1900, Nature, 584, 393–397, doi:doi.org/10.1038/s41586-020-2591-3, URL `https://www.ncbi.nlm.nih.gov/pubmed/32814886`, 2020.

Gallop, S. L., Bryan, K. R., and Coco, G.: Video Observations of Rip Currents on an Embayed Beach, Journal of Coastal Research, pp. 49–53, URL `http://www.jstor.org/stable/25737535`, 2009.

Gaye, J. and Walther, F.: Die Wanderung der Sandriffe vor den ostfriesischen Inseln, Die Bautechnik, 41, 555–567, 1935.

Gers, F. A., Schmidhuber, J., and Cummins, F.: Learning to forget: continual prediction with LSTM, Neural Comput, 12, 2451–71, doi:doi.org/10.1162/

089976600300015015, URL https://www.ncbi.nlm.nih.gov/pubmed/11032042, 2000.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, MIT Press, URL http://www.deeplearningbook.org, 2016.

Grabemann, I. and Weisse, R.: Climate change impact on extreme wave conditions in the North Sea: an ensemble study, Ocean Dynamics, 58, 199–212, doi:doi.org/10.1007/s10236-008-0141-x, 2008.

Graves, A. and Jaitly, N.: Towards End-To-End Speech Recognition with Recurrent Neural Networks, in: International Conference on Machine Learning, edited by Xing, E. P. and Jebara, T., vol. 32, pp. 1764–1772, Proceedings of Machine Learning Research, URL http://proceedings.mlr.press, 2014.

Graves, A., Liwicki, M., Fernandez, S., Bertolami, R., Bunke, H., and Schmidhuber, J.: A novel connectionist system for unconstrained handwriting recognition, IEEE Trans Pattern Anal Mach Intell, 31, 855–68, doi:doi.org/10.1109/TPAMI.2008.137, URL https://www.ncbi.nlm.nih.gov/pubmed/19299860, 2009.

Gulli, A. and Pal, S.: Deep Learning with Keras, Packt Publishing, 2017.

Hagen, R., Freund, J., Plüß, A., and Ihde, R.: Validierungsdokument EasyGSH-DB Nordseemodell. Teil. UnTRIM2 – SediMorph – UnK, Bundesanstalt für Wasserbau, doi:doi.org/10.18451/k2_easygsh_1, 2020.

Halsnæs, K., Drews, M., and Clausen, N.-E.: Socio-economic Impacts—Offshore Activities/Energy, in: North Sea Region Climate Change Assessment, pp. 409–415, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_14, 2016.

Handcock, M. S. and Stein, M. L.: A Bayesian Analysis of Kriging, Technometrics, 35, doi:doi.org/10.2307/1270273, 1993.

Hansen, J. E., Elias, E., and Barnard, P. L.: Changes in surfzone morphodynamics driven by multi-decadal contraction of a large ebb-tidal delta, Marine Geology, 345, 221–234, doi:doi.org/10.1016/j.margeo.2013.07.005, 2013.

Hanson, J. L. and Phillips, O. M.: Automated analysis of ocean surface directional wave spectra, Journal of Atmospheric and Oceanic Technology, 18, 277–293, doi:doi.org/10.1175/1520-0426(2001)018<0277:Aaoosd>2.0.Co;2, 2001.

Hanson, J. L., Tracy, B. A., Tolman, H. L., and Scott, R. D.: Pacific Hindcast Performance of Three Numerical Wave Models, Journal of Atmospheric and Oceanic Technology, 26, 1614–1633, doi:doi.org/10.1175/2009jtecho650.1, 2009.

Harrison, S. R., Bryan, K. R., and Mullarney, J. C.: Observations of morphological change at an ebb-tidal delta, Marine Geology, 385, 131–145, doi:doi.org/10.1016/j.margeo.2016.12.010, 2017.

Hartigan, J. A. and Wong, M. A.: Algorithm AS 136: A K-Means Clustering Algorithm, Applied Statistics, 28, doi:doi.org/10.2307/2346830, 1979.

Harvey, N. and Nicholls, R.: Global sea-level rise and coastal vulnerability, Sustainability Science, 3, 5–7, doi:doi.org/10.1007/s11625-008-0049-x, 2008.

Hasselmann, K., Barnett, T., Bouws, E., Carlson, H., Cartwright, D., Enke, K., Ewing, J., Gienapp, H., Hasselmann, D., Kruseman, P., Meerburg, A., Müller, P., Olbers, D., Richter, K., Sell, W., and Walden, H.: Measurement of Wind-Wave Growth and Swell Decay during the Joint North Sea Wave Project (JONSWAP), Deutsche Hydrographische Zeitschrift, A,8, 95, 1973.

Hasselmann, S., Hasselmann, K., Allender, J., and Barnett, T.: Computations and Parameterizations of the Nonlinear Energy Transfer in a Gravity-Wave Spectrum. Part II: Parameterizations of the Nonlinear Energy Transfer for Application in Wave Models, Journal Of Physical Oceanography, 15, 1378–1391, 1985.

Hasselmann, S., Bruning, C., Hasselmann, K., and Heimbach, P.: An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra, Journal of Geophysical Research: Oceans, 101, 16 615–16 629, doi:doi.org/10.1029/96jc00798, 1996.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition, Springer Series in Statistics, Springer, New York, doi:doi.org/10.1007/b94608, 2009.

Hay, C. C., Morrow, E., Kopp, R. E., and Mitrovica, J. X.: Probabilistic reanalysis of twentieth-century sea-level rise, Nature, 517, 481–484, doi:doi.org/10.1038/nature14093, 2015.

Hayes, M. O.: General morphology and sediment patterns in tidal inlets, Sedimentary Geology, 26, 139–156, 1980.

He, K., Zhang, X., Ren, S., and Sun, J.: Identity Mappings in Deep Residual Networks, arXiv, doi:doi.org/10.48550/arXiv.1603.05027, 2016.

Helmholtz-Zentrum hereon: SeaLevel-Monitor.eu, URL https://meeresspiegel-monitor.de/, https://t1p.de/imqnw;https://web.archive.org/web/20221009095603/https://meeresspiegel-monitor.de/norderney/sla/index.php.en, Accessed 20.08.2022.

Herman, A., Kaiser, R., and Niemeyer, H. D.: MOSES A (03KIS040): KFKI-Projekt MOSES Schlussbericht, Report, NLWKN - Coastal Research Station, URL https://t1p.de/6keu,https://edocs.tib.eu/files/e01fb08/559767226.pdf, http://web.archive.org/web/20201201150527/, https://edocs.tib.eu/files/e01fb08/559767226.pdf, https://t1p.de/6keu, 2007.

Herman, A., Kaiser, R., and Niemeyer, H. D.: Wind-wave variability in a shallow tidal sea—Spectral modelling combined with neural network methods, Coastal Engineering, 56, 759–772, doi:doi.org/10.1016/j.coastaleng.2009.02.007, 2009.

Herrero, A., Sedano, J., Baruque, B., Quintian, H., and Corchado, E.: 10th International Conference on Soft Computing Models in Industrial and Environmental Applications, Advances in Intelligent Systems and Computing, 368, 471, doi:doi.org/10.1007/978-3-319-19719-7, 2015.

Hochreiter, S. and Schmidhuber, J.: Long short-term memory, Neural Comput, 9, 1735–80, doi:doi.org/10.1162/neco.1997.9.8.1735, URL https://www.ncbi.nlm.nih.gov/pubmed/9377276, 1997.

Holman, R. and Stanley, J.: The history and technical capabilities of Argus, Coastal Engineering, 54, 477–491, doi:doi.org/10.1016/j.coastaleng.2007.01.003, 2007.

Holthuijsen, L. H.: Waves in Oceanic and Coastal Waters, Cambridge University Press, New York, doi:doi.org/10.1016/B978-044450556-9/50070-8, 2007.

Hoque, M. A., Perrie, W., and Solomon, S. M.: Application of SWAN model for storm generated wave simulation in the Canadian Beaufort Sea, Journal of Ocean Engineering and Science, doi:doi.org/10.1016/j.joes.2019.07.003, 2019.

Hu, H., van der Westhuysen, A. J., Chu, P., and Fujisaki-Manome, A.: Predicting Lake Erie wave heights and periods using XGBoost and LSTM, Ocean Modelling, 164, doi:doi.org/10.1016/j.ocemod.2021.101832, 2021.

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q.: Densely Connected Convolutional Networks, arXiv, doi:doi.org/10.48550/arXiv.1608.06993, 2018.

Huthnance, J., Weisse, R., Wahl, T., Thomas, H., Pietrzak, J., Souza, A. J., van Heteren, S., Schmelzer, N., van Beusekom, J., Colijn, F., Haigh, I., Hjøllo, S., Holfort, J., Kent, E. C., Kühn, W., Loewe, P., Lorkowski, I., Mork, K. A., Pätsch, J., Quante, M., Salt, L., Siddorn, J., Smyth, T., Sterl, A., and Woodworth, P.: Recent Change—North Sea, in: North Sea Region Climate Change Assessment, pp. 85–136, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_3, 2016.

IPCC: Summary for Policymakers, book section Spm, pp. 3–36, Cambridge University Press, Cambridge, UK and New York, USA, doi:doi.org/10.1017/9781009157964.001, 2019.

Jain, A. K.: Data clustering: 50 years beyond K-means, Pattern Recognition Letters, 31, 651–666, doi:doi.org/10.1016/j.patrec.2009.09.011, 2010.

James, S. C., Zhang, Y. S., and O'Donncha, F.: A machine learning framework to forecast wave conditions, Coastal Engineering, 137, 1–10, doi:doi.org/10.1016/j.coastaleng.2018.03.004, 2018.

Jia, X., Willard, J., Karpatne, A., Read, J., Zwart, J., Steinbach, M., and Kumar, V.: Physics Guided RNNs for Modeling Dynamical Systems: A Case Study in Simulating Lake Temperature Profiles, pp. 558–566, doi:doi.org/10.1137/1.9781611975673.63, 2019.

Jin, X., Cai, S., Li, H., and Karniadakis, G. E.: NSFnets (Navier-Stokes flow nets): Physics-informed neural networks for the incompressible Navier-Stokes equations, Journal of Computational Physics, 426, 109 951, doi:doi.org/10.1016/j.jcp.2020.109951, 2021.

Jörges, C.: Rekonstruktion von Seegangsdaten mit Hilfe Künstlicher Neuronaler Netze im Küstenvorfeld von Norderney, Thesis, [Master-Thesis], 2018.

Jörges, C., Berkenbrink, C., and Stumpe, B.: Wave data prediction and reconstruction by recurrent neural networks at the nearshore area of Norderney, in: EGU General Assembly, doi:doi.org/10.5194/egusphere-egu2020-19772, 2020.

Jörges, C., Berkenbrink, C., Heil, J., and Stumpe, B.: Quantification of morphodynamic variability and sea state damping of plates at the nearshore area

in the East Frisian North Sea, Coastal Engineering, 164, doi:doi.org/10.1016/j.coastaleng.2020.103779, 2021a.

Jörges, C., Berkenbrink, C., and Stumpe, B.: Prediction and reconstruction of ocean wave heights based on bathymetric data using LSTM neural networks, Ocean Engineering, 232, doi:doi.org/10.1016/j.oceaneng.2021.109046, 2021b.

Kagemoto, H.: Forecasting a water-surface wave train with artificial intelligence- A case study, Ocean Engineering, 207, doi:doi.org/10.1016/j.oceaneng.2020.107380, 2020.

Kaiser, R. and Niemeyer, H. D.: Strömungen, p. 200, Eugen Ulmer GmbH & Co., Stuttgart, 1999.

Kaiser, R. and Niemeyer, H. D.: Effect Of Barrier Island Morphology On Tidal Basin Wave Climate, in: Coastal Engineering 2006 - 30th International Conference, edited by Jane, McKee, and Smith, vol. 5, doi:doi.org/10.1142/6439, 2007.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L.: Physics-informed machine learning, Nature Reviews Physics, 3, 422–440, doi:doi.org/10.1038/s42254-021-00314-5, 2021.

Khan, A., Sohail, A., Zahoora, U., and Qureshi, A. S.: A survey of the recent architectures of deep convolutional neural networks, Artificial Intelligence Review, 53, 5455–5516, doi:doi.org/10.1007/s10462-020-09825-6, 2020.

Kim, H. Y. and Won, C. H.: Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models, Expert Systems with Applications, 103, 25–37, doi:doi.org/10.1016/j.eswa.2018.03.002, 2018.

Kim, S. and Kim, H.: A new metric of absolute percentage error for intermittent demand forecasts, International Journal of Forecasting, 32, 669–679, doi:doi.org/10.1016/j.ijforecast.2015.12.003, 2016.

Kingma, D. P. and Ba, J. L.: ADAM: A Method For Stochastic Optimization, 2015.

Koch, M. and Niemeyer, H. D.: Sturmtiden-Strommessungen im Bereich des Norderneyer Seegats, Jber. Forschungsstelle Norderney, 29, 91–108, 1978.

Kreilkamp, E., von Bergner, N. M., and Mauser, C.: Socio-economic Impacts—Recreation, in: North Sea Region Climate Change Assessment, pp. 447–455, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_17, 2016.

Kuhn, M. and Johnson, K.: Applied Predictive Modeling, Springer, New York, doi:doi.org/10.1007/978-1-4614-6849-3, 2013.

Kumar, N. K., Savitha, R., and Mamun, A. A.: Regional ocean wave height prediction using sequential learning neural networks, Ocean Engineering, 129, 605–612, doi:doi.org/10.1016/j.oceaneng.2016.10.033, 2017.

Kumar, N. K., Savitha, R., and Al Mamun, A.: Ocean wave height prediction using ensemble of Extreme Learning Machine, Neurocomputing, 277, 12–20, doi:doi.org/10.1016/j.neucom.2017.03.092, 2018.

Kutupoğlu, V., Çakmak, R. E., Akpınar, A., and van Vledder, G. P.: Setup and evaluation of a SWAN wind wave model for the Sea of Marmara, Ocean Engineering, 165, 450–464, doi:doi.org/10.1016/j.oceaneng.2018.07.053, 2018.

Lantuéjoul, C.: Geostatistical Simulation. Models and Algorithms, Springer, Berlin, Heidelberg, 1 edn., doi:doi.org/10.1007/978-3-662-04808-5, 2002.

Law, Y. Z., Santo, H., Lim, K. Y., and Chan, E. S.: Deterministic wave prediction for unidirectional sea-states in real-time using Artificial Neural Network, Ocean Engineering, 195, doi:doi.org/10.1016/j.oceaneng.2019.106722, 2020.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D.: Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation, 1, 541–551, 1989.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep learning, Nature, 521, 436–44, doi:doi.org/10.1038/nature14539, URL `https://www.ncbi.nlm.nih.gov/pubmed/26017442`, 2015.

Lewis, M. J., Palmer, T., Hashemi, R., Robins, P., Saulter, A., Brown, J., Lewis, H., and Neill, S.: Wave-tide interaction modulates nearshore wave height, Ocean Dynamics, 69, 367–384, doi:doi.org/10.1007/s10236-018-01245-z, 2019.

Li, X., Zhang, G., Huang, H. H., Wang, Z., and Zheng, W.: Performance Analysis of GPU-Based Convolutional Neural Networks, doi:doi.org/10.1109/icpp.2016.15, 2016.

Li, Y., Wu, H., and Liu, H.: Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction, Energy Conversion and Management, 167, 203–219, doi:doi.org/10.1016/j.enconman.2018.04.082, 2018.

Lippe, C.: Küstenschutz in Niedersachsen: Ein erster Überblick. Nds. Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz, URL https://www.nlwkn.niedersachsen.de/startseite/hochwasser_kustenschutz/kustenschutz/kuestenschutz-in-niedersachsen-sicherheit-fuer-die-menschen-45612.html, https://t1p.de/ib3wc;https://web.archive.org/web/20221009094738/https://www.nlwkn.niedersachsen.de/startseite/hochwasser_kustenschutz/kustenschutz/kuestenschutz-in-niedersachsen-sicherheit-fuer-die-menschen-45612.html, Accessed 20.09.2022.

Londhe, S. N.: Soft computing approach for real-time estimation of missing wave heights, Ocean Engineering, 35, 1080–1089, doi:doi.org/10.1016/j.oceaneng.2008.05.003, 2008.

Londhe, S. N. and Panchang, V.: Correlation of wave data from buoy networks, Estuarine, Coastal and Shelf Science, 74, 481–492, doi:doi.org/10.1016/j.ecss.2007.05.003, 2007.

Lubac, B., Burvingt, O., Lerma, A. N., and Sénéchal, N.: Performance and Uncertainty of Satellite-Derived Bathymetry Empirical Approaches in an Energetic Coastal Environment, Remote Sensing, 14, 2350, doi:doi.org/10.3390/rs14102350, 2022.

MacQueen, J.: Some Methods For Classification And Analysis Of Multivariate Observations, University Of California, p. 17, 1967.

Mahjoobi, J., Etemad-Shahidi, A., and Kazeminezhad, M. H.: Hindcasting of wave parameters using different soft computing methods, Applied Ocean Research, 30, 28–36, doi:doi.org/10.1016/j.apor.2008.03.002, 2008.

Mai, S., Paesler, C., and Zimmermann, C.: Wellen und Seegang an Küsten und Küstenbauwerken, Report, Lehrstuhl für Wasserbau und Küsteningenieurwesen, Franzius-Institut, Universität Hannover, 2004.

Malcherek, A.: Gezeiten und Wellen. Die Hydromechanik der Küstengewässer, vol. 1, Vieweg+Teubner, Wiesbaden, 2010.

Malekmohamadi, I., Ghiassi, R., and Yazdanpanah, M. J.: Wave hindcasting by coupling numerical model and artificial neural networks, Ocean Engineering, 35, 417–425, doi:doi.org/10.1016/j.oceaneng.2007.09.003, 2008.

Malekmohamadi, I., Bazargan-Lari, M. R., Kerachian, R., Nikoo, M. R., and Fallahnia, M.: Evaluating the efficacy of SVMs, BNs, ANNs and ANFIS in wave height prediction, Ocean Engineering, 38, 487–497, doi:doi.org/10.1016/j.oceaneng.2010.11.020, 2011.

Mandal, S. and Prabaharan, N.: Ocean Wave Prediction Using Numerical and Neural Network Models, The Open Ocean Engineering Journal, 3, 12–17, 2010.

Matheron, G.: Les variables régionalisées et leur estimation: Une application de la théorie des fonctions aléatoires aux sciences de la nature, Thesis, 1965.

Matérn, B.: Spatial Variation, Meddelanden från Statens Skogsforskningsinstitut, 49, doi:doi.org/10.1007/978-1-4615-7892-5, [2nd Edition (1986): Lecture Notes in Statistics. Springer, New York.], 1960.

McCulloch, W. S. and Pitts, W.: A logical calculus of the ideas immanent in nervous activity, The Bulletin of Mathematical Biophysics, 5, 115–133, doi:doi.org/10.1007/bf02478259, 1943.

Meyer, C.: Morphodynamische Analysen im Bereich des Norderneyer Seegats und seines Einzugsgebietes. Technical Report 01/2014, doi:Unpublished\results, 2014.

Minasny, B. and McBratney, A. B.: The Matérn function as a general model for soil variograms, Geoderma, 128, 192–207, doi:doi.org/10.1016/j.geoderma.2005.04.003, 2005.

Mori, N., Shimura, T., Yasuda, T., and Mase, H.: Multi-model climate projections of ocean surface variables under different climate scenarios-Future change of waves, sea level and wind, Ocean Engineering, 71, 122–129, doi:doi.org/10.1016/j.oceaneng.2013.02.016, 2013.

Ng, A. Y.: Feature selection, L1 vs. L2 regularization, and rotational invariance, in: International Conference on Machine Learning, 2004.

Ni, C. and Ma, X.: An integrated long-short term memory algorithm for predicting polar westerlies wave height, Ocean Engineering, 215, doi:doi.org/10.1016/j.oceaneng.2020.107715, 2020.

Nicholls, R., Wong, P., Burkett, V., Codignotto, J., Hay, J., McLean, R., Ragoonaden, S., and Woodroffe, C.: Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, chap. Coastal systems

and low-lying areas, pp. 315–356, Cambridge University Press, Cambridge, UK, 2007.

Nie, F., Huang, H., Cai, X., and Ding, C.: Efficient and Robust Feature Selection via Joint l2,1-Norms Minimization, in: Advances in Neural Information Processing Systems, vol. 23, pp. 1813–1821, 2010.

Niemeyer, H. D.: Untersuchungen zum Seegangsklima im Bereich der Ostfriesischen Inseln und Küste, Die Küste, 34, 53–70, 1979.

Niemeyer, H. D.: Bemessung von See- und Ästuardeichen in Niedersachsen, Die Küste, 64, doi:https://hdl.handle.net/20.500.11970/101452, 2001.

Niemeyer, H. D. and Kaiser, R.: Effects Of Ebb Delta Migration On Local Wave Climate, 1999.

Niemeyer, H. D. and Kaiser, R.: Evaluation of Design Water Levels and Design Wave Run-Up for an Estuarine Coastal Protection Master Plan, doi:doi.org/10.1061/40549(276)118, 2001.

Niemeyer, H. D., Beaufort, G., Mayerle, R., Monbaliu, J., Townend, I., Madsen, H. T., de Vriend, H., and Wurpts, A.: Socio-economic Impacts—Coastal Protection, in: North Sea Region Climate Change Assessment, pp. 457–474, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_18, 2016.

Niemeyer, H. D., Berkenbrink, C., Ritzmann, A., Knaack, H., Wurpts, A., and Kaiser, R.: Evaluierung von Küstenschutz-Strategien im Hinblick auf Klimaänderungsfolgen, Die Küste. Kuratorium für Forschung im Küsteningenieurwesen (KFKI), 86, doi:https://hdl.handle.net/20.500.11970/106389, 2018.

Nitsure, S. P., Londhe, S. N., and Khare, K. C.: Wave forecasts using wind information and genetic programming, Ocean Engineering, 54, 61–69, doi:doi.org/10.1016/j.oceaneng.2012.07.017, 2012.

NLWKN: Jahresbericht 2018/2019, 2019.

NLWKN: Küstenschutz: Unruhige Sturmflutsaison sorgt für kurzfristigen Handlungsbedarf auf den Inseln. Nds. Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz, URL `https://t1p.de/rhdxd`, accessed 20.09.2022, 2022.

NLWKN Pressestelle, a.: Zahlen aus Wasserwirtschaft und Naturschutz 2019/2020. Nds. Landesbetrieb für Wasserwirtschaft, Küsten- und

Naturschutz, URL https://www.nlwkn.niedersachsen.de/jb2020/zahlen-aus-wasserwirtschaft-und-naturschutz-2019-2020-191018.html, https://t1p.de/7xi2s;https://web.archive.org/web/20221009100538/https://www.nlwkn.niedersachsen.de/jb2020/zahlen-aus-wasserwirtschaft-und-naturschutz-2019-2020-191018.html, Accessed 20.09.2022, 2020.

Olesen, J. E.: Socio-economic Impacts—Agricultural Systems, in: North Sea Region Climate Change Assessment, pp. 397–407, Springer International Publishing, doi: doi.org/10.1007/978-3-319-39745-0_13, 2016.

Oost, A. P., Hofstede, J., Weisse, R., Baart, F., Janssen, G., and Zijlstra, R.: Climate change, Wadden Sea Quality Status Report 2017, URL qsr.waddensea-worldheritage.org/reports/climate-change, 2017.

O'Donncha, F., Zhang, Y., Chen, B., and James, S. C.: An integrated framework that combines machine learning and numerical models to improve wave-condition forecasts, Journal of Marine Systems, 186, 29–36, doi:doi.org/10.1016/j.jmarsys.2018.05.006, 2018.

Pascanu, R., Mikolov, T., and Bengio, Y.: On the difficulty of training recurrent neural networks, in: 30th International Conference on Machine Learning, vol. 28, 2013.

Pearson, S. G., Elias, E. P., van Prooijen, B. C., van der Vegt, H., van der Spek, A. J., and Wang, Z. B.: A novel approach to mapping ebb-tidal delta morphodynamics and stratigraphy, Geomorphology, 405, 108 185, doi:doi.org/10.1016/j.geomorph.2022.108185, 2022.

Pebesma, E. J.: gstat user's manual, Report, Dept. of Physical Geography, Utrecht University, URL http://www.gstat.org/, 2014.

Peres, D. J., Iuppa, C., Cavallaro, L., Cancelliere, A., and Foti, E.: Significant wave height record extension by neural networks and reanalysis wind data, Ocean Modelling, 94, 128–140, doi:doi.org/10.1016/j.ocemod.2015.08.002, 2015.

Petersen, J.: Ostfriesische Inseln - Vegetation im Wandel der Zeit. Mit Vegetationskarten, Schlütersche, 2005.

Phillips, M. S., Harley, M. D., Turner, I. L., Splinter, K. D., and Cox, R. J.: Shoreline recovery on wave-dominated sandy coastlines: the role of sandbar morpho-

dynamics and nearshore wave parameters, Marine Geology, 385, 146–159, doi: doi.org/10.1016/j.margeo.2017.01.005, 2017.

Pierson, W. J. and Moskowitz, L.: A Proposed Spectral Form for Fully Developed Wind Seas Based on the Similarity Theory of S. A. Kitaigorodskii, Journal of Geophysical Research, 69, 5181–5190, 1964.

Pinnegar, J. K., Engelhard, G. H., Jones, M. C., Cheung, W. W., Peck, M. A., Rijnsdorp, A. D., and Brander, K. M.: Socio-economic Impacts—Fisheries, in: North Sea Region Climate Change Assessment, pp. 375–395, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_12, 2016.

Pirhooshyaran, M. and Snyder, L. V.: Forecasting, hindcasting and feature selection of ocean waves via recurrent and sequence-to-sequence networks, Ocean Engineering, 207, doi:doi.org/10.1016/j.oceaneng.2020.107424, 2020.

Pirhooshyaran, M., Scheinberg, K., and Snyder, L. V.: Feature engineering and forecasting via derivative-free optimization and ensemble of sequence-to-sequence networks with applications in renewable energy, Energy, 196, doi:doi.org/10.1016/j.energy.2020.117136, 2020.

Portilla, J., Ocampo-Torres, F. J., and Monbaliu, J.: Spectral Partitioning and Identification of Wind Sea and Swell, Journal of Atmospheric and Oceanic Technology, 26, 107–122, doi:doi.org/10.1175/2008jtecho609.1, 2009.

Puscasu, R. M.: Integration of Artificial Neural Networks into Operational Ocean Wave Prediction Models for Fast and Accurate Emulation of Exact Nonlinear Interactions, Procedia Computer Science, 29, 1156–1170, doi:doi.org/10.1016/j.procs.2014.05.104, 2014.

QGIS Development Team: QGIS Geographic Information System, Report, Open Source Geospatial Foundation Project, URL `http://qgis.osgeo.org`, 2019.

Qi, D. and Majda, A. J.: Using machine learning to predict extreme events in complex systems, Proceedings of the National Academy of Sciences, 117, 52–59, doi:doi.org/10.1073/pnas.1917285117, 2019.

Qing, X. and Niu, Y.: Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM, Energy, 148, 461–468, doi:doi.org/10.1016/j.energy.2018.01.177, 2018.

Quante, M. and Colijn, F.: North Sea Region Climate Change Assessment, Regional Climate Studies, Springer, doi:doi.org/10.1007/978-3-319-39745-0, 2016.

Quedens, G.: Nordsee und Wattenmeer., Breklumer Vlg., Bred., 2000.

Raissi, M., Perdikaris, P., and Karniadakis, G.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational Physics, 378, 686–707, doi:doi.org/10.1016/j.jcp.2018.10.045, 2019.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection, doi:doi.org/10.1109/cvpr.2016.91, 2016.

Reikard, G., Pinson, P., and Bidlot, J.-R.: Forecasting ocean wave energy: The ECMWF wave model and time series methods, Ocean Engineering, 38, 1089–1099, doi:doi.org/10.1016/j.oceaneng.2011.04.009, 2011.

Renga, A., Rufino, G., D'Errico, M., Moccia, A., Boccia, V., Graziano, M. D., Aragno, C., and Zoffoli, S.: Analysis of spaceborne SAR monitoring capabilities for coastal areas bathymetry with COSMO-SkyMed and ALOS data, doi:doi.org/10.1117/12.2028974, 2013.

Renga, A., Rufino, G., D'Errico, M., Moccia, A., Boccia, V., Graziano, M. D., Aragno, C., and Zoffoli, S.: SAR Bathymetry in the Tyrrhenian Sea by COSMO-SkyMed Data: A Novel Approach, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 2834–2847, doi:doi.org/10.1109/jstars.2014.2327150, 2014.

Ritzmann, A. and Kaiser, R.: Messkette Norderney – Jahresbericht 2009 (Datensatz 2007/2008), NLWKN-Forschungsstelle Küste, doi:Unpublished\results, 2010.

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

Rogowski, P., Paolo, T., Terrill, E., and McNinch, J.: X-Band Radar Mapping of Morphological Changes at a Dynamic Coastal Inlet, Journal of Geophysical Research: Earth Surface, 123, 3034–3054, doi:doi.org/10.1029/2018jf004676, 2018.

Ruggiero, P., Komar, P. D., and Allan, J. C.: Increasing wave heights and extreme value projections: The wave climate of the U.S. Pacific Northwest, Coastal Engineering, 57, 539–552, doi:doi.org/10.1016/j.coastaleng.2009.12.005, 2010.

Salameh, E., Frappart, F., Almar, R., Baptista, P., Heygster, G., Lubac, B., Raucoules, D., Almeida, L., Bergsma, E., Capo, S., Michele, M. D., Idier, D., Li, Z., Marieu, V., Poupardin, A., Silva, P., Turki, I., and Laignel, B.: Monitoring Beach Topography and Nearshore Bathymetry Using Spaceborne Remote Sensing: A Review, Remote Sensing, 11, 2212, doi:doi.org/10.3390/rs11192212, 2019.

Salman, A. G., Heryadi, Y., Abdurahman, E., and Suparta, W.: Single Layer & Multi-layer Long Short-Term Memory (LSTM) Model with Intermediate Variables for Weather Forecasting, Procedia Computer Science, 135, 89–98, doi:doi.org/10.1016/j.procs.2018.08.153, 2018.

Salmon, J. E. and Holthuijsen, L. H.: Modeling depth-induced wave breaking over complex coastal bathymetries, Coastal Engineering, 105, 21–35, doi:doi.org/10.1016/j.coastaleng.2015.08.002, 2015.

Salmon, J. E., Holthuijsen, L. H., Zijlema, M., van Vledder, G. P., and Pietrzak, J. D.: Scaling depth-induced wave-breaking in two-dimensional spectral wave models, Ocean Modelling, 87, 30–47, doi:doi.org/10.1016/j.ocemod.2014.12.011, 2015.

Schlather, M.: Simulation and Analysis of Random Fields, Report, 2001.

Schlather, M.: Construction of Covariance Functions and Unconditional Simulation of Random Fields, Lecture Notes in Statistics, Springer, Berlin, Heidelberg, 1 edn., doi:doi.org/10.1007/978-3-642-17086-7, 2012.

Schrum, C., Lowe, J., Meier, H. E. M., Grabemann, I., Holt, J., Mathis, M., Pohlmann, T., Skogen, M. D., Sterl, A., and Wakelin, S.: Projected Change—North Sea, in: North Sea Region Climate Change Assessment, pp. 175–217, Springer International Publishing, doi:doi.org/10.1007/978-3-319-39745-0_6, 2016.

Sebastian, A., Proft, J., Dietrich, J. C., Du, W., Bedient, P. B., and Dawson, C. N.: Characterizing hurricane storm surge behavior in Galveston Bay using the SWAN+ADCIRC model, Coastal Engineering, 88, 171–181, doi:doi.org/10.1016/j.coastaleng.2014.03.002, 2014.

Sheng, W. A. and Li, H.: A Method for Energy and Resource Assessment of Waves in Finite Water Depths, Energies, 10, doi:doi.org/10.3390/en10040460, 2017.

Shumway, R. H. and Stoffer, D. S.: Time Series Analysis and Its Applications. With R Examples. Second Edition, Springer Texts In Statistics, Springer, New York, 2006.

Sierra, J. P., Genius, A., Lionello, P., Mestres, M., Mösso, C., and Marzo, L.: Modelling the impact of climate change on harbour operability: The Barcelona port case study, Ocean Engineering, 141, 64–78, doi:doi.org/10.1016/j.oceaneng.2017.06.002, 2017.

Simonyan, K. and Zissermann, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv, doi:doi.org/10.48550/arXiv.1409.1556, 2015.

Snoek, J., Larochelle, H., and Adams, R. P.: Practical Bayesian Optimization of Machine Learning Algorithms, Advances in Neural Information Processing Systems, 25, 2012.

Song, X., Liu, Y., Xue, L., Wang, J., Zhang, J., Wang, J., Jiang, L., and Cheng, Z.: Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model, Journal of Petroleum Science and Engineering, 186, doi:doi.org/10.1016/j.petrol.2019.106682, 2020.

Spicer, B. A., Brodie, K. L., Hesser, T. J., and Smith, J. M.: Applying dynamically updated nearshore bathymetry estimates to operational nearshore wave modeling, Coastal Engineering, 145, 53–64, doi:doi.org/10.1016/j.coastaleng.2018.12.005, 2019.

Stefanakos, C.: Fuzzy time series forecasting of nonstationary wind and wave data, Ocean Engineering, 121, 1–12, doi:doi.org/10.1016/j.oceaneng.2016.05.018, 2016.

Stein, M. L.: Interpolation of Spatial Data. Some Theory for Kriging, Springer Series in Statistics, Springer, New York, doi:doi.org/10.1007/978-1-4612-1494-6, 1999.

Stewart, C., Renga, A., Gaffney, V., and Schiavon, G.: Sentinel-1 bathymetry for North Sea palaeolandscape analysis, International Journal of Remote Sensing, 37, 471–491, doi:doi.org/10.1080/01431161.2015.1129563, 2016.

Stolz, A.: Küstenschutz und Deichbau in Niedersachsen: Antworten auf häufig gestellte Fragen. Nds. Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz, URL `https://www.nlwkn.niedersachsen.de/startseite/hochwasser_kustenschutz/kustenschutz/antworten_auf_haufig_gestellte_fragen/`

kuestenschutz-und-deichbau-in-niedersachsen-45182.html, https://t1p.de/5zd49;https://web.archive.org/web/20221009095451/https://www.nlwkn.niedersachsen.de/startseite/hochwasser_kustenschutz/kustenschutz/antworten_auf_haufig_gestellte_fragen/kuestenschutz-und-deichbau-in-niedersachsen-45182.html, (Accessed 20.09.2022), 2021.

Sultana, F., Sufian, A., and Dutta, P.: Advancements in Image Classification using Convolutional Neural Network, doi:doi.org/10.1109/ICRCICN.2018.8718718, 2018.

Sun, L., Gao, H., Pan, S., and Wang, J.-X.: Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data, Computer Methods in Applied Mechanics and Engineering, 361, 112 732, doi:doi.org/10.1016/j.cma.2019.112732, 2020.

Surface-Water Modeling System: SMS User Manual v12.1, Report, Aquaveo, 2012.

SWAN: SWAN User Manual. SWAN Cycle III version 41.20AB, Report, Delft University of Technology. Faculty of Civil Engineering and Geosciences. Environmental Fluid Mechanics Section, 2019a.

SWAN: SWAN Scientific And Technical Documentation. SWAN Cycle III version 41.20AB, Report, Delft University of Technology. Faculty of Civil Engineering and Geosciences. Environmental Fluid Mechanics Section, 2019b.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z.: Rethinking the Inception Architecture for Computer Vision, arXiv, doi:doi.org/10.48550/arXiv.1512.00567, 2015.

Teich, T., Groll, N., and Weisse, R.: Long-term statistics of potentially hazardous sea states in the North Sea 1958-2014, Ocean Dynamics, 68, 1559–1570, doi:doi.org/10.1007/s10236-018-1210-4, 2018.

Thai, L. H., Hai, T. S., and Thuy, N. T.: Image Classification using Support Vector Machine and Artificial Neural Network, International Journal of Information Technology and Computer Science, 4, 32–38, doi:doi.org/10.5815/ijitcs.2012.05.05, 2012.

The Wamdi Group: The WAM Model—A Third Generation Ocean Wave Prediction Model, Journal of Physical Oceanography, 18, 1775–1810, doi:doi.org/10.1175/1520-0485(1988)018<1775:Twmtgo>2.0.Co;2, 1988.

Ti, Z., Zhang, M., Wu, L., Qin, S., Wei, K., and Li, Y.: Estimation of the significant wave height in the nearshore using prediction equations based on the Response Surface Method, Ocean Engineering, 153, 143–153, doi:doi.org/10.1016/j.oceaneng.2018.01.081, 2018.

Toimil, A., Losada, I. J., Nicholls, R. J., Dalrymple, R. A., and Stive, M. J.: Addressing the challenges of climate change risks and adaptation in coastal areas: A review, Coastal Engineering, 156, doi:doi.org/10.1016/j.coastaleng.2019.103611, 2020.

Tolman, H. L., Balasubramaniyan, B., Burroughs, L. D., Chalikov, D. V., Chao, Y. Y., Chen, H. S., and Gerald, V. M.: Development and Implementation of Wind-Generated Ocean Surface Wave Modelsat NCEP*, Weather and Forecasting, 17, 311–333, doi:doi.org/10.1175/1520-0434(2002)017<0311:Daiowg>2.0.Co;2, 2002.

van der Westhuysen, A. J.: Advances in the spectral modelling of wind waves in the nearshore, Ph.D. thesis, 2007.

van der Westhuysen, A. J.: Modelling of depth-induced wave breaking over sloping and horizontal beds, 2009.

van der Westhuysen, A. J.: Modeling of depth-induced wave breaking under finite depth wave growth conditions, Journal of Geophysical Research, 115, doi:doi.org/10.1029/2009jc005433, 2010.

van der Westhuysen, A. J.: Spectral modeling of wave dissipation on negative current gradients, Coastal Engineering, 68, 17–30, doi:doi.org/10.1016/j.coastaleng.2012.05.001, 2012.

Vanem, E. and Walker, S.-E.: Identifying trends in the ocean wave climate by time series analyses of significant wave height data, Ocean Engineering, 61, 148–160, doi:doi.org/10.1016/j.oceaneng.2012.12.042, 2013.

Vapnik, V. N.: Statistical Learning Theory, Wiley, New York, 1998.

Vapnik, V. N.: The Nature of Statistical Learning Theory. Second Edition, Springer, New York, doi:doi.org/10.1007/978-1-4757-3264-1, 2000.

Venables, W. and Ripley, B.: Modern Applied Statistics with S. Fourth Edition, Springer, New York, doi:doi.org/10.1007/978-0-387-21706-2, 2002.

Verhagen, H. J., van Vledder, G., and Eslami Arab, S.: A Practical Method For Design Of Coastal Structures In Shallow Water, 2008.

Vieira, F., Cavalcante, G., and Campos, E.: Analysis of wave climate and trends in a semi-enclosed basin (Persian Gulf) using a validated SWAN model, Ocean Engineering, 196, doi:doi.org/10.1016/j.oceaneng.2019.106821, 2020.

WAFO: WAFO - a Matlab Toolbox for Analysis of Random Waves and Loads: Tutorial for WAFO version 2.5, Report, Mathematical Statistics, Lund University, 2011.

Wahl, T., Haigh, I., Woodworth, P., Albrecht, F., Dillingh, D., Jensen, J., Nicholls, R., Weisse, R., and Wöppelmann, G.: Observed mean sea level changes around the North Sea coastline from 1800 to present, Earth Science Reviews, 124, 51–67, doi:doi.org/10.1016/j.earscirev.2013.05.003, 2013.

Wang, D. W. and Hwang, P. A.: An operational method for separating wind sea and swell from ocean wave spectra, Journal of Atmospheric and Oceanic Technology, 18, 2052–2062, doi:doi.org/10.1175/1520-0426(2001)018<2052:Aomfsw>2.0.Co;2, 2001.

Wang, N., Chen, Q., and Chen, Z.: Reconstruction of nearshore wave fields based on physics-informed neural networks, Coastal Engineering, 176, 104 167, doi:doi.org/10.1016/j.coastaleng.2022.104167, 2022.

Wang, Z. B., Hoekstra, P., Burchard, H., Ridderinkhof, H., De Swart, H. E., and Stive, M. J. F.: Morphodynamics of the Wadden Sea and its barrier island system, Ocean & Coastal Management, 68, 39–57, doi:doi.org/10.1016/j.ocecoaman.2011.12.022, 2012.

Wei, Z. and Davison, A.: A convolutional neural network based model to predict nearshore waves and hydrodynamics, Coastal Engineering, 171, doi:doi.org/10.1016/j.coastaleng.2021.104044, 2022.

Weisse, R. and Gunther, H.: Wave climate and long-term changes for the Southern North Sea obtained from a high-resolution hindcast 1958-2002, Ocean Dynamics, 57, 161–172, doi:doi.org/10.1007/s10236-006-0094-x, 2007.

Weisse, R., von Storch, H., Niemeyer, H. D., and Knaack, H.: Changing North Sea storm surge climate: An increasing hazard?, Ocean & Coastal Management, 68, 58–68, 2012.

Willard, J., Jia, X., Xu, S., Steinbach, M., and Kumar, V.: Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems, doi:doi.org/10.48550/ARXIV.2003.04919, 2020.

Witting, M., Berkenbrink, C., and Wurpts, A.: Bemessungsseegang für die Deckwerke am Nord- und Weststrand von Norderney. Technical Report 01/2014, doi: Unpublished\results, 2014.

Wu, Y., Schuster, M., Chen, Z., V. Le, Q., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, L., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M., and Dean, J.: Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation, doi:https://arxiv.org/abs/1609.08144v2, 2016.

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., and Chen, Z.: Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach, Remote Sensing of Environment, 233, doi:doi.org/10.1016/j.rse.2019.111358, 2019.

Young, I.: Wind Generated Ocean Waves, vol. 2 of *Elsevier Ocean Engineering Book Series*, Elsevier, Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, 1999.

Yuan, Y., Chen, X., and Wang, J.: Object-Contextual Representations for Semantic Segmentation, doi:doi.org/10.1007/978-3-030-58539-6_11, 2020.

Zamani, A., Solomatine, D., Azimian, A., and Heemink, A.: Learning from data for wind–wave forecasting, Ocean Engineering, 35, 953–962, doi:doi.org/10.1016/j.oceaneng.2008.03.007, 2008.

Zhang, H., Li, D., Wang, J., Zhou, H., Guan, W., Lou, X., Cao, W., Shi, A., Chen, P., Fan, K., Ren, L., Zheng, G., and Li, Y.: Long time-series remote sensing analysis of the periodic cycle evolution of the inlets and ebb-tidal delta of Xincun Lagoon, Hainan Island, China, ISPRS Journal of Photogrammetry and Remote Sensing, 165, 67–85, doi:doi.org/10.1016/j.isprsjprs.2020.05.006, 2020.

Zhang, R., Chen, Z., Chen, S., Zheng, J., Büyüköztürk, O., and Sun, H.: Deep long short-term memory networks for nonlinear structural seismic response prediction, Computers and Structures, 220, 55–68, doi:doi.org/10.1016/j.compstruc.2019.05.006, 2019.

Zhou, S., Xie, W., Lu, Y., Wang, Y., Zhou, Y., Hui, N., and Dong, C.: ConvLSTM-Based Wave Forecasts in the South and East China Seas, Frontiers in Marine Science, 8, doi:doi.org/10.3389/fmars.2021.680079, 2021.

Zijlema, M. and van der Westhuysen, A. J.: On convergence behaviour and numerical accuracy in stationary SWAN simulations of nearshore wind wave spectra, Coastal Engineering, 52, 237–256, doi:doi.org/10.1016/j.coastaleng.2004.12.006, 2005.

Appendix

Appendix I

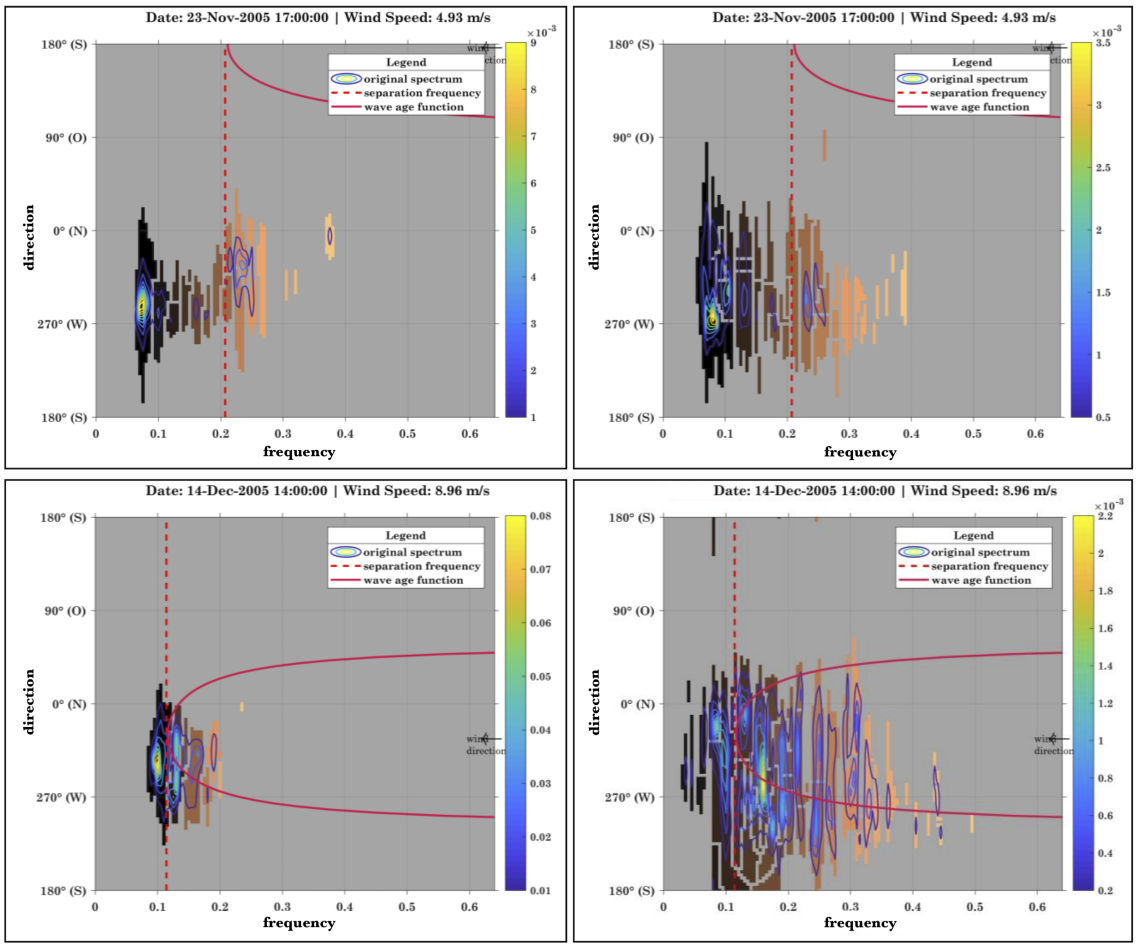

Figure I.1: *Examples of the spectral separation methods performed in this study. The separation frequency (dashed line) and the wave age function (solid line) are plotted over the original spectrum and the several partitions calculated by the watershed algorithm are presented (brown areas). Some spectrums even are not well separated by the wave age criterion (above) while most spectrums are separated satisfactory by both methods (under).*

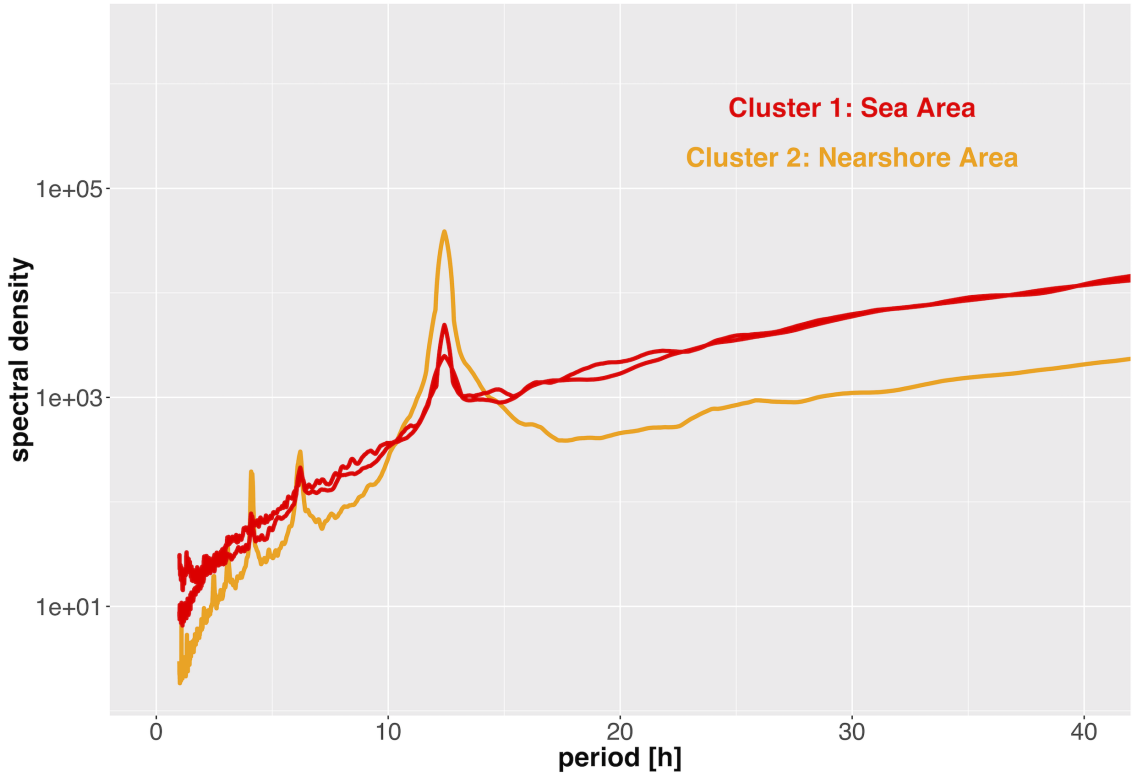

Figure I.2: *Periodogram of H_{m_0} time series from 1992–2017 for the three buoys. The spectral density shows a tidal peak at a period of 12.5 h.*

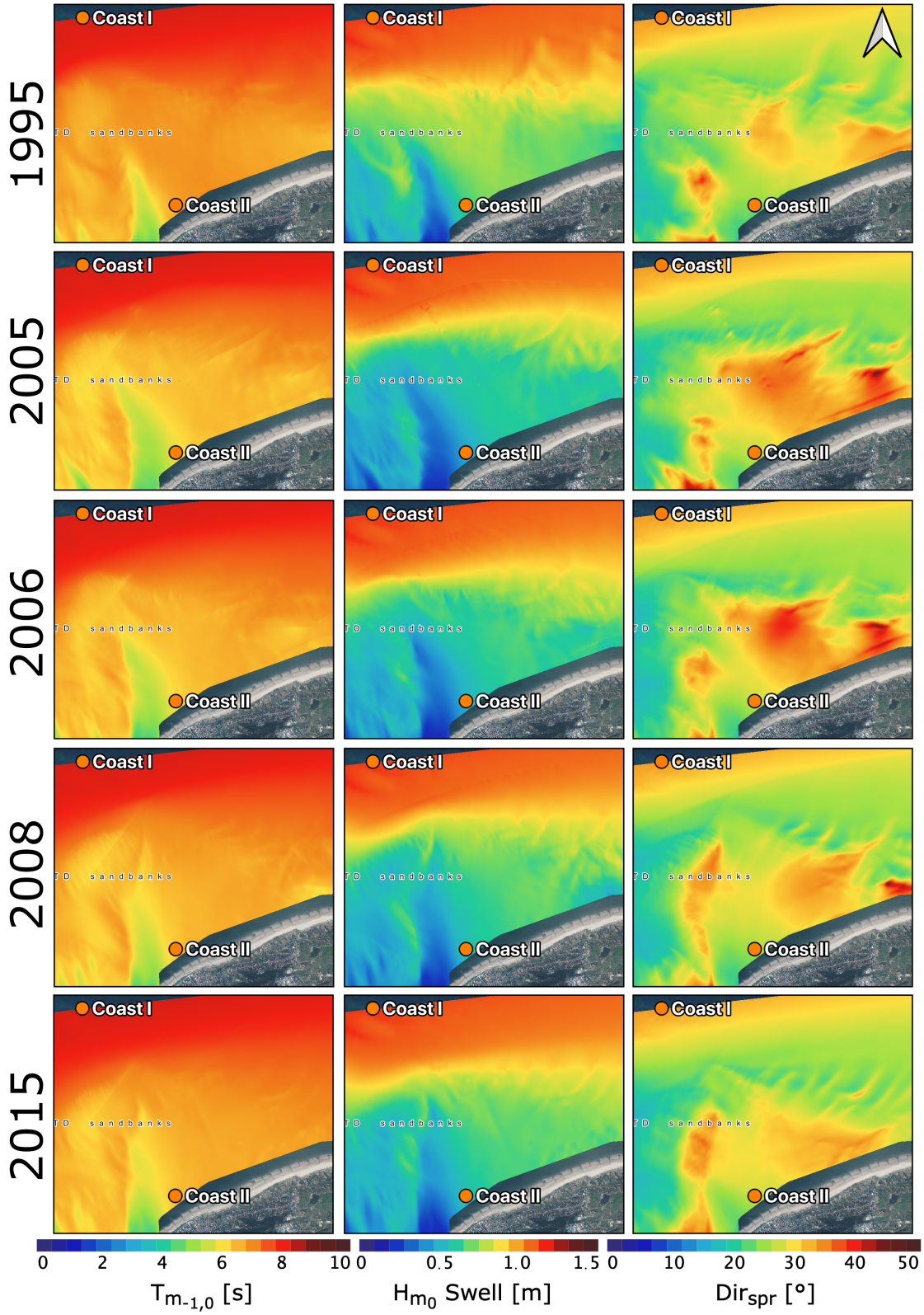

Figure I.3: *Sea state parameter mean wave period $T_{m_{-1,0}}$, significant wave height of the swell H_{m_0} swell and mean directional spreading Dir_{Spr} for the years 1995, 2005, 2006, 2008 and 2015 for the simulation of test case #3 (05-Sep-2015 16:00).*

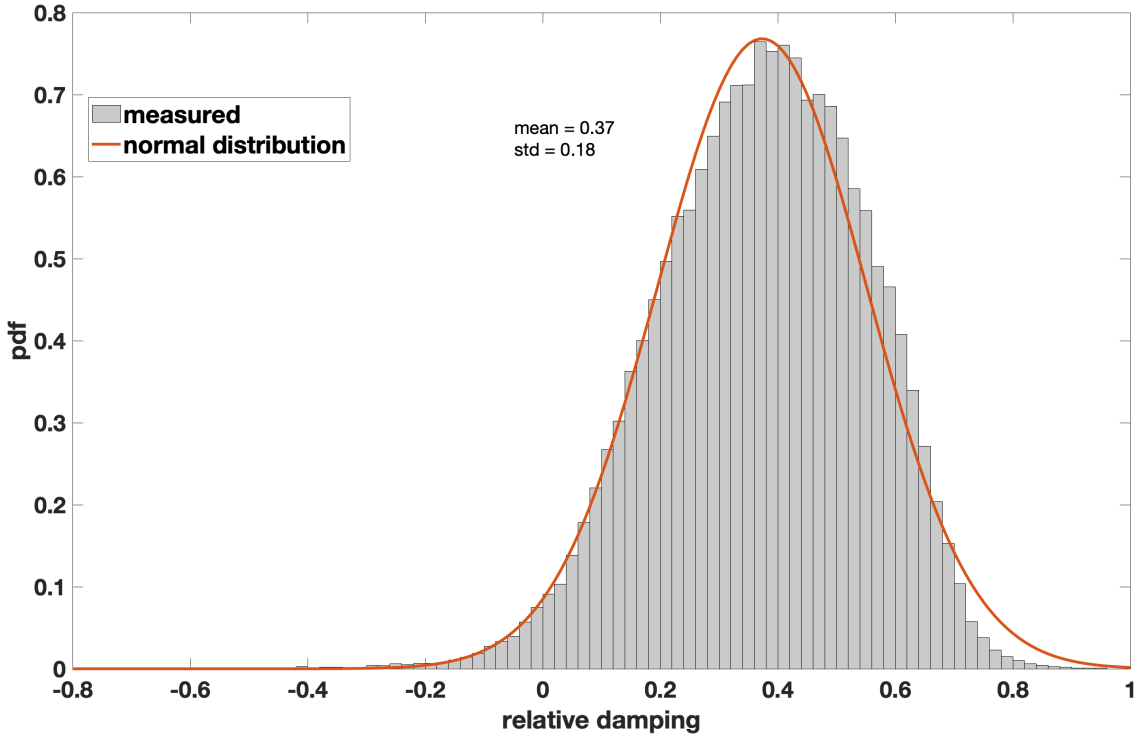

Figure I.4: *Distribution of the buoy-measured relative damping effect for the years 2004–2017. The red line indicates a normal distribution with the respective parameters of the mean relative damping and the standard deviation given in the figure.*

Appendix II

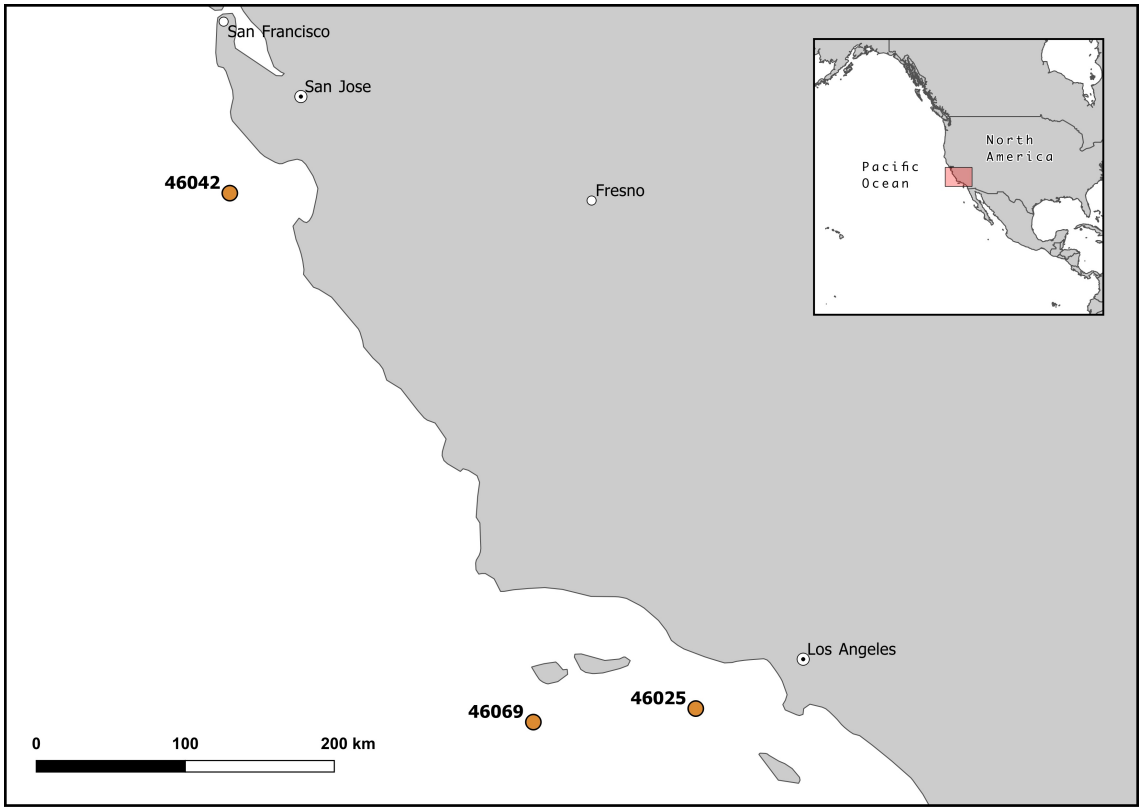

Figure II.1: *Map of the positions of the three NOAA deep ocean wave buoys.*

Table II.1: *Basic information of the three NOAA deep ocean wave buoys.*

NOAA Station ID	SWH Mean [m]	SWH Std. [m]	SWH Range [m]	Mean wave period [s]	# Data	Location	Water depth [m NHN]
46042	2.23	0.93	[0.66; 7.19]	7.57	9 267	N 36.785 W 122.398	1 646
46069	2.07	0.79	[0.66; 6.92]	7.43	9 267	N 33.677 W 120.213	978
46025	1.09	0.44	[0.34; 4.09]	6.55	9 267	N 33.758 W 119.044	890

Table II.2: *Feature Selection excluding wind speed. Short-term predictions of SWH in the nearshore area of Norderney, Germany. The mean value and standard deviation of the three independent runs are given, respectively.*

Problem task	Model	RMSE [m]	R^2	r	MAAPE	MAE [m]
Reconstruction	LSTM	**0.079** ± 0.002	**0.954** ± 0.002	0.979 ± 0.000	0.107 ± 0.001	**0.059** ± 0.001
	P-LSTM	0.081 ± 0.003	0.953 ± 0.002	**0.980** ± 0.000	**0.104** ± 0.001	0.060 ± 0.002
	FFNN	0.099 ± 0.001	0.928 ± 0.003	0.965 ± 0.000	0.134 ± 0.002	0.073 ± 0.001
	SL-FFNN	0.112 ± 0.002	0.907 ± 0.004	0.954 ± 0.002	0.155 ± 0.002	0.083 ± 0.002
	SVR	0.118 ± 0.000	0.897 ± 0.001	0.950 ± 0.001	0.194 ± 0.001	0.095 ± 0.000
	RF	0.107 ± 0.000	0.916 ± 0.000	0.957 ± 0.000	0.146 ± 0.001	0.080 ± 0.000
	MLR	0.152 ± 0.001	0.829 ± 0.002	0.911 ± 0.001	0.213 ± 0.001	0.114 ± 0.000
3h-ahead	LSTM	0.105 ± 0.001	0.919 ± 0.001	0.960 ± 0.000	0.139 ± 0.002	0.077 ± 0.000
	P-LSTM	**0.102** ± 0.001	**0.924** ± 0.001	**0.963** ± 0.000	**0.135** ± 0.002	**0.075** ± 0.001
	FFNN	0.128 ± 0.002	0.880 ± 0.004	0.939 ± 0.002	0.175 ± 0.004	0.095 ± 0.002
	SL-FFNN	0.145 ± 0.002	0.846 ± 0.005	0.924 ± 0.003	0.184 ± 0.003	0.105 ± 0.001
	SVR	0.141 ± 0.001	0.855 ± 0.002	0.928 ± 0.001	0.219 ± 0.002	0.111 ± 0.001
	RF	0.135 ± 0.000	0.867 ± 0.000	0.931 ± 0.000	0.180 ± 0.000	0.099 ± 0.000
	MLR	0.169 ± 0.002	0.791 ± 0.003	0.890 ± 0.002	0.234 ± 0.001	0.127 ± 0.001
6h-ahead	LSTM	0.140 ± 0.002	0.859 ± 0.004	0.930 ± 0.002	0.179 ± 0.003	0.102 ± 0.001
	P-LSTM	**0.133** ± 0.002	**0.871** ± 0.006	**0.935** ± 0.002	**0.173** ± 0.002	**0.097** ± 0.001
	FFNN	0.165 ± 0.003	0.804 ± 0.004	0.898 ± 0.003	0.215 ± 0.007	0.121 ± 0.002
	SL-FFNN	0.187 ± 0.004	0.747 ± 0.013	0.874 ± 0.002	0.239 ± 0.021	0.137 ± 0.006
	SVR	0.170 ± 0.002	0.790 ± 0.005	0.892 ± 0.002	0.256 ± 0.002	0.134 ± 0.001
	RF	0.168 ± 0.001	0.796 ± 0.003	0.892 ± 0.002	0.225 ± 0.002	0.124 ± 0.000
	MLR	0.196 ± 0.002	0.722 ± 0.003	0.849 ± 0.002	0.266 ± 0.001	0.148 ± 0.002
9h-ahead	LSTM	0.160 ± 0.001	0.811 ± 0.004	0.905 ± 0.003	0.205 ± 0.004	0.117 ± 0.001
	P-LSTM	**0.152** ± 0.003	**0.832** ± 0.006	**0.915** ± 0.001	**0.199** ± 0.005	**0.112** ± 0.001
	FFNN	0.191 ± 0.002	0.728 ± 0.004	0.854 ± 0.001	0.244 ± 0.004	0.140 ± 0.001
	SL-FFNN	0.208 ± 0.001	0.679 ± 0.004	0.830 ± 0.003	0.261 ± 0.005	0.152 ± 0.001
	SVR	0.195 ± 0.002	0.716 ± 0.005	0.850 ± 0.003	0.283 ± 0.000	0.151 ± 0.001
	RF	0.192 ± 0.002	0.726 ± 0.007	0.852 ± 0.004	0.255 ± 0.003	0.142 ± 0.002
	MLR	0.219 ± 0.002	0.643 ± 0.005	0.802 ± 0.003	0.295 ± 0.001	0.166 ± 0.001
12h-ahead	LSTM	0.179 ± 0.005	0.767 ± 0.014	0.878 ± 0.007	0.230 ± 0.005	0.132 ± 0.004
	P-LSTM	**0.169** ± 0.002	**0.795** ± 0.004	**0.894** ± 0.002	**0.219** ± 0.001	**0.124** ± 0.001
	FFNN	0.217 ± 0.002	0.654 ± 0.004	0.811 ± 0.002	0.268 ± 0.001	0.158 ± 0.000
	SL-FFNN	0.244 ± 0.005	0.565 ± 0.018	0.772 ± 0.002	0.292 ± 0.015	0.176 ± 0.005
	SVR	0.216 ± 0.001	0.660 ± 0.003	0.814 ± 0.003	0.303 ± 0.002	0.166 ± 0.000
	RF	0.212 ± 0.002	0.670 ± 0.006	0.819 ± 0.004	0.281 ± 0.001	0.158 ± 0.001
	MLR	0.234 ± 0.001	0.599 ± 0.002	0.774 ± 0.002	0.312 ± 0.002	0.178 ± 0.001

Table II.3: *Feature Selection excluding wind direction. Short-term predictions of SWH in the nearshore area of Norderney, Germany. The mean value and standard deviation of the three independent runs are given, respectively.*

Problem task	Model	RMSE [m]	R^2	r	MAAPE	MAE [m]
Reconstruction	LSTM	**0.089** ± 0.003	**0.940** ± 0.003	**0.974** ± 0.001	0.119 ± 0.002	**0.067** ± 0.002
	P-LSTM	**0.088** ± 0.001	**0.944** ± 0.001	**0.974** ± 0.001	**0.115** ± 0.002	**0.065** ± 0.001
	FFNN	**0.088** ± 0.003	0.942 ± 0.004	**0.974** ± 0.000	0.127 ± 0.011	0.066 ± 0.002
	SL-FFNN	0.091 ± 0.006	0.938 ± 0.008	0.972 ± 0.000	0.131 ± 0.007	0.069 ± 0.006
	SVR	0.105 ± 0.001	0.918 ± 0.001	0.958 ± 0.000	0.164 ± 0.002	0.081 ± 0.000
	RF	0.104 ± 0.000	0.919 ± 0.000	0.959 ± 0.000	0.143 ± 0.001	0.077 ± 0.000
	MLR	0.157 ± 0.002	0.815 ± 0.002	0.903 ± 0.001	0.212 ± 0.002	0.116 ± 0.001
3h-ahead	LSTM	0.118 ± 0.002	0.899 ± 0.002	0.950 ± 0.001	0.155 ± 0.004	0.086 ± 0.001
	P-LSTM	**0.116** ± 0.001	**0.902** ± 0.001	**0.951** ± 0.000	**0.152** ± 0.003	**0.085** ± 0.001
	FFNN	0.121 ± 0.002	0.893 ± 0.002	0.947 ± 0.001	0.171 ± 0.007	0.090 ± 0.001
	SL-FFNN	0.124 ± 0.001	0.887 ± 0.004	0.945 ± 0.002	0.161 ± 0.003	0.090 ± 0.001
	SVR	0.132 ± 0.001	0.872 ± 0.001	0.937 ± 0.000	0.206 ± 0.003	0.103 ± 0.001
	RF	0.132 ± 0.000	0.871 ± 0.001	0.934 ± 0.001	0.178 ± 0.001	0.097 ± 0.000
	MLR	0.176 ± 0.001	0.772 ± 0.004	0.879 ± 0.002	0.240 ± 0.001	0.132 ± 0.000
6h-ahead	LSTM	0.155 ± 0.002	0.826 ± 0.003	0.911 ± 0.002	**0.198** ± 0.002	0.113 ± 0.001
	P-LSTM	**0.152** ± 0.001	**0.833** ± 0.004	**0.914** ± 0.002	**0.198** ± 0.003	**0.111** ± 0.000
	FFNN	0.157 ± 0.002	0.822 ± 0.004	0.908 ± 0.002	0.219 ± 0.005	0.117 ± 0.002
	SL-FFNN	0.162 ± 0.002	0.809 ± 0.007	0.904 ± 0.003	0.230 ± 0.015	0.123 ± 0.005
	SVR	0.166 ± 0.002	0.802 ± 0.004	0.898 ± 0.002	0.246 ± 0.002	0.127 ± 0.002
	RF	0.167 ± 0.001	0.798 ± 0.002	0.894 ± 0.002	0.223 ± 0.001	0.123 ± 0.000
	MLR	0.206 ± 0.001	0.693 ± 0.002	0.832 ± 0.001	0.277 ± 0.001	0.155 ± 0.000
9h-ahead	LSTM	0.186 ± 0.004	0.748 ± 0.011	0.868 ± 0.002	0.233 ± 0.006	0.135 ± 0.002
	P-LSTM	**0.179** ± 0.001	**0.766** ± 0.001	**0.877** ± 0.000	**0.230** ± 0.004	**0.130** ± 0.001
	FFNN	0.187 ± 0.002	0.744 ± 0.006	0.865 ± 0.003	0.247 ± 0.010	0.138 ± 0.000
	SL-FFNN	0.190 ± 0.003	0.736 ± 0.004	0.861 ± 0.000	0.240 ± 0.007	0.138 ± 0.000
	SVR	0.193 ± 0.002	0.728 ± 0.002	0.854 ± 0.002	0.271 ± 0.004	0.146 ± 0.001
	RF	0.192 ± 0.002	0.730 ± 0.004	0.855 ± 0.002	0.255 ± 0.001	0.143 ± 0.001
	MLR	0.227 ± 0.002	0.623 ± 0.002	0.790 ± 0.002	0.304 ± 0.002	0.172 ± 0.001
12h-ahead	LSTM	0.208 ± 0.002	0.682 ± 0.004	0.833 ± 0.001	**0.255** ± 0.005	0.151 ± 0.000
	P-LSTM	**0.202** ± 0.001	**0.701** ± 0.002	**0.840** ± 0.001	**0.259** ± 0.005	**0.148** ± 0.002
	FFNN	0.210 ± 0.002	0.678 ± 0.005	0.826 ± 0.001	0.271 ± 0.000	0.154 ± 0.001
	SL-FFNN	0.215 ± 0.003	0.660 ± 0.011	0.820 ± 0.001	0.269 ± 0.010	0.157 ± 0.000
	SVR	0.216 ± 0.002	0.658 ± 0.003	0.813 ± 0.001	0.297 ± 0.004	0.163 ± 0.002
	RF	0.215 ± 0.001	0.663 ± 0.003	0.815 ± 0.002	0.283 ± 0.004	0.160 ± 0.002
	MLR	0.242 ± 0.001	0.571 ± 0.004	0.756 ± 0.002	0.321 ± 0.004	0.183 ± 0.001

Table II.4: *Feature Selection excluding water level. Short-term predictions of SWH in the nearshore area of Norderney, Germany. The mean value and standard deviation of the three independent runs are given, respectively.*

Problem task	Model	RMSE [m]	R^2	r	MAAPE	MAE [m]
Reconstruction	LSTM	**0.134** ± 0.002	**0.866** ± 0.005	**0.932** ± 0.001	0.159 ± 0.002	**0.096** ± 0.001
	P-LSTM	**0.132** ± 0.000	**0.873** ± 0.000	**0.937** ± 0.000	**0.157** ± 0.000	**0.094** ± 0.000
	FFNN	**0.178** ± 0.001	**0.765** ± 0.004	**0.876** ± 0.001	0.195 ± 0.003	**0.122** ± 0.000
	SL-FFNN	**0.201** ± 0.002	**0.700** ± 0.008	**0.846** ± 0.005	0.221 ± 0.003	**0.139** ± 0.001
	SVR	0.169 ± 0.000	0.788 ± 0.003	0.889 ± 0.002	0.225 ± 0.000	0.127 ± 0.000
	RF	0.161 ± 0.000	0.808 ± 0.002	0.899 ± 0.001	0.190 ± 0.001	0.117 ± 0.000
	MLR	0.180 ± 0.000	0.761 ± 0.002	0.872 ± 0.001	0.219 ± 0.002	0.130 ± 0.000
3h-ahead	LSTM	0.168 ± 0.001	0.797 ± 0.003	0.894 ± 0.001	0.206 ± 0.002	0.122 ± 0.001
	P-LSTM	**0.162** ± 0.000	**0.807** ± 0.001	**0.900** ± 0.002	**0.199** ± 0.003	**0.118** ± 0.000
	FFNN	**0.202** ± 0.001	0.708 ± 0.005	**0.842** ± 0.003	0.239 ± 0.004	**0.145** ± 0.000
	SL-FFNN	**0.221** ± 0.001	**0.650** ± 0.005	**0.815** ± 0.003	0.246 ± 0.002	**0.156** ± 0.001
	SVR	0.197 ± 0.000	0.723 ± 0.003	0.851 ± 0.001	0.265 ± 0.002	0.150 ± 0.001
	RF	0.192 ± 0.001	0.736 ± 0.001	0.859 ± 0.001	0.239 ± 0.001	0.142 ± 0.000
	MLR	0.205 ± 0.001	0.698 ± 0.001	0.836 ± 0.001	0.260 ± 0.001	0.152 ± 0.001
6h-ahead	LSTM	0.196 ± 0.004	0.719 ± 0.011	0.852 ± 0.002	**0.239** ± 0.003	0.144 ± 0.002
	P-LSTM	**0.187** ± 0.004	**0.745** ± 0.008	**0.865** ± 0.005	**0.236** ± 0.005	**0.138** ± 0.003
	FFNN	**0.231** ± 0.002	0.610 ± 0.005	0.783 ± 0.004	**0.273** ± 0.006	**0.167** ± 0.002
	SL-FFNN	**0.257** ± 0.009	**0.514** ± 0.029	**0.740** ± 0.013	0.282 ± 0.011	**0.182** ± 0.004
	SVR	0.226 ± 0.002	0.624 ± 0.003	0.792 ± 0.002	0.304 ± 0.003	0.174 ± 0.001
	RF	0.221 ± 0.001	0.641 ± 0.001	0.801 ± 0.001	0.281 ± 0.001	0.166 ± 0.001
	MLR	0.233 ± 0.002	0.602 ± 0.003	0.776 ± 0.002	0.296 ± 0.002	0.174 ± 0.001
9h-ahead	LSTM	0.207 ± 0.003	0.688 ± 0.007	0.829 ± 0.004	**0.263** ± 0.004	0.153 ± 0.002
	P-LSTM	**0.201** ± 0.002	**0.702** ± 0.004	**0.842** ± 0.002	**0.249** ± 0.004	**0.147** ± 0.001
	FFNN	0.238 ± 0.000	0.587 ± 0.005	0.769 ± 0.004	0.286 ± 0.005	0.173 ± 0.001
	SL-FFNN	0.259 ± 0.008	0.510 ± 0.035	0.738 ± 0.007	0.313 ± 0.013	0.188 ± 0.005
	SVR	0.238 ± 0.001	0.586 ± 0.007	0.767 ± 0.004	0.317 ± 0.004	0.181 ± 0.001
	RF	**0.234** ± 0.001	**0.601** ± 0.005	**0.776** ± 0.003	0.299 ± 0.001	**0.175** ± 0.001
	MLR	0.246 ± 0.001	0.560 ± 0.005	0.748 ± 0.003	0.313 ± 0.001	0.183 ± 0.001
12h-ahead	LSTM	0.217 ± 0.001	0.650 ± 0.008	0.809 ± 0.005	**0.271** ± 0.004	0.160 ± 0.001
	P-LSTM	**0.208** ± 0.002	**0.678** ± 0.006	**0.826** ± 0.004	**0.261** ± 0.001	**0.153** ± 0.001
	FFNN	**0.247** ± 0.001	0.548 ± 0.006	0.742 ± 0.003	0.301 ± 0.007	**0.181** ± 0.002
	SL-FFNN	0.270 ± 0.005	0.461 ± 0.017	**0.700** ± 0.009	**0.323** ± 0.012	**0.196** ± 0.005
	SVR	0.246 ± 0.001	0.550 ± 0.009	0.743 ± 0.005	0.327 ± 0.000	0.187 ± 0.000
	RF	**0.245** ± 0.001	**0.557** ± 0.007	**0.747 ± 0.005**	0.313 ± 0.001	**0.183 ± 0.001**
	MLR	0.255 ± 0.001	0.519 ± 0.006	0.721 ± 0.004	0.325 ± 0.003	0.189 ± 0.001

Table II.5: *Feature Selection with SWH only. Short-term predictions of SWH in the nearshore area of Norderney, Germany. The mean value and standard deviation of the three independent runs are given, respectively.*

Problem task	Model	RMSE [m]	R^2	r	MAAPE	MAE [m]
Reconstruction	LSTM	**0.175** ± 0.001	**0.779** ± 0.003	**0.884** ± 0.002	0.207 ± 0.003	**0.126** ± 0.001
	P-LSTM	**0.173** ± 0.000	**0.780** ± 0.001	**0.884** ± 0.001	**0.200** ± 0.001	**0.125** ± 0.000
	FFNN	**0.172** ± 0.001	**0.781** ± 0.003	**0.884** ± 0.002	0.205 ± 0.002	**0.124** ± 0.000
	SL-FFNN	**0.170** ± 0.000	**0.781** ± 0.000	**0.884** ± 0.000	0.207 ± 0.000	**0.124** ± 0.000
	SVR	0.176 ± 0.001	0.772 ± 0.001	0.880 ± 0.000	0.233 ± 0.001	0.131 ± 0.000
	RF	0.173 ± 0.001	0.779 ± 0.003	0.883 ± 0.002	0.204 ± 0.001	0.125 ± 0.000
	MLR	0.194 ± 0.002	0.722 ± 0.002	0.850 ± 0.001	0.246 ± 0.002	0.141 ± 0.001
3h-ahead	LSTM	0.197 ± 0.001	0.720 ± 0.002	0.850 ± 0.001	0.247 ± 0.007	0.146 ± 0.001
	P-LSTM	**0.198** ± 0.000	**0.711** ± 0.001	**0.843** ± 0.000	**0.239** ± 0.004	**0.145** ± 0.001
	FFNN	**0.195** ± 0.001	0.724 ± 0.002	**0.852** ± 0.001	0.248 ± 0.010	**0.145** ± 0.002
	SL-FFNN	**0.195** ± 0.002	**0.725** ± 0.004	**0.852** ± 0.002	0.244 ± 0.006	**0.145** ± 0.002
	SVR	0.200 ± 0.002	0.712 ± 0.004	0.846 ± 0.002	0.272 ± 0.003	0.151 ± 0.001
	RF	0.197 ± 0.001	0.718 ± 0.003	0.848 ± 0.002	0.244 ± 0.001	0.146 ± 0.001
	MLR	0.216 ± 0.001	0.663 ± 0.002	0.814 ± 0.001	0.278 ± 0.001	0.160 ± 0.01
6h-ahead	LSTM	0.232 ± 0.002	0.606 ± 0.002	0.781 ± 0.001	**0.282** ± 0.007	0.171 ± 0.001
	P-LSTM	**0.231** ± 0.001	**0.609** ± 0.003	**0.781** ± 0.001	**0.284** ± 0.001	**0.172** ± 0.000
	FFNN	**0.230** ± 0.002	0.612 ± 0.003	0.784 ± 0.000	**0.282** ± 0.007	**0.170** ± 0.001
	SL-FFNN	**0.230** ± 0.002	**0.613** ± 0.004	**0.785** ± 0.001	0.285 ± 0.002	**0.170** ± 0.001
	SVR	0.233 ± 0.002	0.600 ± 0.001	0.777 ± 0.000	0.313 ± 0.003	0.177 ± 0.001
	RF	0.231 ± 0.002	0.609 ± 0.003	0.780 ± 0.002	0.290 ± 0.001	0.172 ± 0.001
	MLR	0.247 ± 0.001	0.551 ± 0.005	0.743 ± 0.003	0.321 ± 0.002	0.185 ± 0.000
9h-ahead	LSTM	0.249 ± 0.004	0.549 ± 0.008	0.742 ± 0.004	**0.304** ± 0.001	0.183 ± 0.002
	P-LSTM	**0.251** ± 0.000	**0.548** ± 0.000	**0.740** ± 0.000	**0.307** ± 0.000	**0.185** ± 0.000
	FFNN	0.248 ± 0.004	0.554 ± 0.008	0.746 ± 0.004	0.306 ± 0.004	0.183 ± 0.002
	SL-FFNN	0.246 ± 0.003	0.558 ± 0.004	0.748 ± 0.003	0.313 ± 0.004	0.184 ± 0.001
	SVR	0.249 ± 0.004	0.547 ± 0.006	0.742 ± 0.004	0.327 ± 0.003	0.187 ± 0.002
	RF	**0.245** ± 0.003	**0.560** ± 0.007	**0.749** ± 0.004	0.306 ± 0.003	**0.182** ± 0.002
	MLR	0.263 ± 0.002	0.497 ± 0.004	0.705 ± 0.003	0.337 ± 0.002	0.195 ± 0.002
12h-ahead	LSTM	0.254 ± 0.001	0.528 ± 0.003	0.727 ± 0.002	**0.325** ± 0.003	0.189 ± 0.001
	P-LSTM	**0.256** ± 0.003	**0.517** ± 0.002	**0.720** ± 0.001	**0.318** ± 0.005	**0.189** ± 0.000
	FFNN	**0.253** ± 0.001	0.533 ± 0.002	0.731 ± 0.001	0.324 ± 0.005	**0.188** ± 0.000
	SL-FFNN	0.254 ± 0.004	0.529 ± 0.011	**0.732** ± 0.001	**0.318** ± 0.011	**0.188** ± 0.002
	SVR	0.255 ± 0.001	0.527 ± 0.001	0.727 ± 0.001	0.335 ± 0.002	0.191 ± 0.000
	RF	**0.253** ± 0.001	**0.535** ± 0.001	**0.732** ± 0.001	0.323 ± 0.001	**0.188** ± 0.000
	MLR	0.268 ± 0.000	0.475 ± 0.002	0.690 ± 0.002	0.350 ± 0.001	0.200 ± 0.000

Appendix III

List of publications

1. Jörges, C.; Berkenbrink, C.; Heil, J. and Stumpe, B. (2021): Quantification of morphodynamic variability and sea state damping of plates at the nearshore area in the East Frisian North Sea, Coastal Engineering, 164, doi: 10.1016/j.coastaleng.2020.103779.

   Peer-Review: The manuscript was submitted for publication in Coastal Engineering on 17 February 2020 and after peer-review with two anonymous reviewers, it was accepted for publication on 30 August 2020.

   Modifications: References to figures, tables, sections, and appendix were modified. Unification of notation from U_{Dir} to U_θ. Harmonization of variables and units.

2. Jörges, C.; Berkenbrink, C. and Stumpe, B. (2021): Prediction and reconstruction of ocean wave heights based on bathymetric data using LSTM neural networks, Ocean Engineering, 232, doi: 10.1016/j.oceaneng.2021.109046.

   Peer-Review: The manuscript was submitted for publication in Ocean Engineering on 17 December 2020 and after peer-review with two anonymous reviewers, it was accepted for publication on 18 April 2021.

   Modifications: References to figures, tables, sections, and appendix were modified. Harmonization of section names, variables and units.

3. Jörges, C.; Berkenbrink, C.; Gottschalk, H. and Stumpe, B. (2023): Spatial ocean wave height prediction with CNN mixed-data deep neural networks using random field simulated bathymetry, Ocean Engineering, 271, doi: 10.1016/j.oceaneng.2023.113699.

   Peer-Review: The manuscript was submitted for publication in Ocean Engineering on 23 June 2022 and after peer-review with two anonymous reviewers, it was accepted for publication on 10 January 2023.

   Modifications: References to figures, tables, and sections were modified. Unification of notation from H_S to H_{m_0}. Harmonization of variables and units.