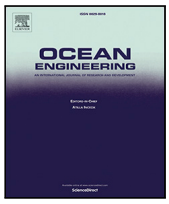
Ocean Engineering 317 (2025) 119963

Contents lists available at ScienceDirect

Ocean Engineering

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



Research paper

Data-driven, non-linear ship response prediction based on time series of irregular, long-crested sea states amidships

Johanna Serr [a,*], Mathies Wedler [d], Merten Stender [c], Nuno Fonseca [f], C. Guedes Soares [g], Norbert Hoffmann [b,e], Sören Ehlers [d], Marco Klein [d]

[a] Institute for Fluid Dynamics and Ship Theory, Hamburg University of Technology (TUHH), Am Schwarzenberg-Campus 4, 21073, Hamburg, Germany
[b] Dynamics Group, Hamburg University of Technology (TUHH), Schlossmühlendamm 30, 21073, Hamburg, Germany
[c] Cyber–Physical Systems in Mechanical Engineering, Technical University of Berlin (TU Berlin), Straße des 17. Juni 135, 10623, Berlin, Germany
[d] Institute for Maritime Energy Systems, German Aerospace Center (DLR), Düneberger Str. 108, 21502, Geesthacht, Germany
[e] Department of Mechanical Engineering, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom
[f] SINTEF Ocean, Postboks 4762 Torgarden, 7465 Trondheim, Norway
[g] Centre for Marine Technology and Ocean Engineering (CENTEC), Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal

ARTICLE INFO

Keywords:
Data-driven prediction
Neural networks
Non-linear ship response prediction
Response prediction zone
New year wave

ABSTRACT

The accurate prediction of vessel responses in waves is crucial for decision-making and contribute to the operational safety and risk minimization. Short-term predictions can be carried out by estimating the vessel's motions and loads based on incident waves. Existing model-based approaches either require computationally intensive simulations that compromise real-time capability or use simplified models affecting the accuracy of the prediction. Therefore, this study explores the feasibility of using neural networks for mapping time signals of surface elevation data and a set of corresponding ship responses, i.e. the heave and pitch motions as well as the vertical bending moment. The approach followed here is built on the assumption that the wave profile amidships is known. A synthetic dataset was generated using a time-domain strip theory solver with considerations of non-linear effects on motions and loads due to large amplitude waves in a variety of irregular, long-crested sea state conditions. We propose two different neural network models, a multi-layer perceptron (MLP) and a fully convolutional neural network (FCNN), and compare their performances on measurement data obtained from model tests in a seakeeping basin. The evaluations also include the freak wave reproduction of the 'new year wave'. The proposed networks are able to estimate the motions and bending moment accurately for a wide range of sea state conditions, surpassing current state-of-the-art models on the given data sets.

1. Introduction

The safety and operational efficiency of vessels advancing in waves are highly dependent on the prevailing sea state. Extreme wave conditions lead to large amplitude motions, cargo loss and critical bending moments. In cases of such events, the real-time prediction of wave-induced responses can provide early-warning information to the mariners and support the crew in decision-making. Therefore, the prediction of vessel responses in critical wave sequences is an on-going topic in the field of naval architecture.

In the past, various numerical methods for estimating the motions and loads of a vessel advancing in waves have been established and have become proven methodology. Hereby, the sea state is commonly described through the wave energy spectrum under the assumption that the wave process is stationary and ergodic within observation time (Ochi, 1990). Response amplitude operators (RAOs) can be utilized to estimate the vessel's responses in frequency domain when assuming linearity between wave excitation and vessel response. However, linear models are restricted to seaway conditions with small to moderate wave steepness with respect to wave prediction (Klein et al., 2020) as well as ship response estimation (Adegeest, 1996; Dannenberg et al., 2010), motivating the usage of non-linear simulation methods. Non-linear effects can be modeled using strip theory to a certain degree of accuracy. Most strip theory solvers are based on the method proposed by Salvesen et al. (1970). By adding modifications to the system of equations, non-linearities associated with large amplitude waves can be considered to some extent (Fonseca and Guedes Soares, 1998, 1970). Computational fluid dynamics (CFD) such as simulations based on the Reynolds-averaged Navier–Stokes equations (RANS) allows an accurate

* Corresponding author.
E-mail address: johanna.serr@tuhh.de (J. Serr).

https://doi.org/10.1016/j.oceaneng.2024.119963
Received 7 June 2024; Received in revised form 11 November 2024; Accepted 27 November 2024

representation of non-linear effects; however, are of high computational costs. Although these methods give adequate results in terms of motion and load prediction, they lack real-time capability.

The resulting trade-off between accuracy and simulation time that are faced by the mentioned models have spurred research into applying alternative techniques for on-board decision support systems. Triantafyllou and Athans (1981) and Triantafyllou et al. (1983) applied a Kalman filtering technique for predicting ship motions up to 5 s ahead. Hereby, the knowledge of the ship hydrodynamics is required. The method is robust against noisy measurements; however, investigations show that the prediction accuracy highly depends on the estimation of the peak frequency of the present wave spectrum, which limits its application in real-world scenarios. Yumori (1981) proposed an auto-regressive moving average (ARMA) model and by applying time series analysis, the model is able to predict the ship motions for up to $2-4$ s with no previous knowledge about the ship hydrodynamics necessary. Another concept of a predictive model is introduced by Naaijen and Huijsmans (2008, 2010) as well as Naaijen et al. (2018, 2009), which involves the following three steps: (i) the recording of the surrounding wave field, (ii) its phase-resolved propagation prediction to the vessel position and (iii) the estimation of the corresponding vessel responses in future conditions. The authors use wave probe data and wave field data from radar measurements as input into their model. The wave propagation is modeled using linear wave theory and the ship responses are determined using RAOs. The concept was tested in long and short-crested seaways, showing accurate prediction results for a time interval inside of the theoretical prediction zone and with an upper limit of 60 s. The advantages of applying linear theory lie in a minimal calculation time as well as robustness and easy maintainability. The described method was tested on full scale tests in the scope of the joint industry project 'On board Wave and Motion Estimator (OWME)' (Dannenberg et al., 2010) in mild sea state conditions, reaching a prediction horizon up to 120 s. Similar results are achieved by commercially available prediction systems like FutureWaves^{TM} (Kusters et al., 2016) using linear theory.

Besides model-driven approaches, the potential of implementing data-driven techniques in the field of vessel motion prediction has increasingly been investigated in recent years. The following studies use historical ship motion and/or wave data in order to perform a short-term forecast of the wave propagation or vessel responses. Sclavounos and Ma (2018) proposed a support vector machine (SVM) framework for predicting the wave elevation up to 5 s into the future. Furthermore, the authors demonstrated that SVM can also be applied for accurately determine ship roll hydrodynamics under different loading conditions using experimental data. The feasibility of using dynamic mode decomposition (DMD) for extracting spatio-temporal modes from time-series data for predicting maneuvering motions has been studied by Chen et al. (2023a) and the method is compared against higher order dynamic mode decomposition (HODMD) by the same authors (Chen et al., 2023b). Khan et al. (2007) showed that by using a fully-connected neural network (FCNN) trained on historical motion data, roll motion prediction is feasible up to 7 s. Recurrent neural networks (RNNs) are conveniently utilized for processing sequential data such as time series due to their ability to memorize long-term dependencies in the input data. Long short-term memory (LSTM) networks as a variant of the RNN overcome the vanishing or exploding gradients during training and have been deployed to predict heave and surge motions of a semi-submersible by Guo et al. (2021). The authors trained and tested a LSTM-based model on experimental data to forecast the motions both with and without the support of wave data. Lee et al. (2023) have applied a LSTM model in combination with a convolutional neural network (CNN) showing that the heave, roll and pitch responses of the KVLCC2 tanker can be predicted accurately, taking both incident wave excitations as well as motion memory effects into account. Further research expands the field of application to hybrid models combining LSTM and gaussian processes (GP) (Sun et al., 2022), bi-directional LSTM (BiLSTM) frameworks (Jiang et al., 2024; Xue et al., 2024), radial basis function (RBF) networks (Yin et al., 2018) or attention-based LSTM (Zhang et al., 2021). The concept of a digital twin by combining data-driven methods for predicting the wave field evolution and ship responses is introduced by Lee et al. (2022). Applying the backpropagation algorithm to train model weights entails high computational costs. A more cost-effective neural network framework is reservoir computing (RC). In RC, a low-dimensional input vector is fed into the reservoir, a high-dimensional phase space. The approach of only training the output layer enables the dynamic properties of the reservoir and decreases the training costs compared to conventional RNNs. The potential of using RC for forecasting the surge, heave, and pitch motions of a moored barge in head waves to up to 50 s has been demonstrated by Yang et al. (2023a). The training data was obtained from model tests and include one sea state condition for irregular waves. A comparison of the RC framework to LSTM, BiLSTM and gated recurrent unit (GRU) networks for a limited number of sea states was done by the same authors (Yang et al., 2024, 2023b).

In contrast to the mentioned studies, the authors del Águila Ferrandis et al. (2021) follow a different approach and suggest to obtain the vessel responses from given wave sequence data in the sense of a mapping task. CFD simulations are used to create the training data, which includes two wave spectra that represent mild and extreme conditions, for various vessel types operating in irregular, long-crested seas. Since the proposed LSTM network is limited to the chosen sea states, the authors conclude that a generalization of the model can be achieved with more extensive training using a wider range of wave spectra. A summary of the listed studies is given in Appendix A.

Similar to latter publication, this study explores the feasibility of predicting vessel responses in the framework of a non-linear mapping task. The basic prerequisite is the knowledge of the wave profile at the midship section of the ship, which will serve as input into the predictive model. Most notably, the authors specify that the models must be able to function in a large range of sea states and, in particular, to cover extreme-events. We propose two neural network models, a multi-layer perceptron and a fully convolutional neural network, to perform a sequence-to-sequence mapping of the non-linear relations between surface elevation data amidships and a set of corresponding responses, i.e. the vertical motions pitch and heave as well as the vertical bending moment. This concept aims at expanding the available models for ship response prediction by the following characteristics: (i) a non-linear mapping method is presented which has the ability of close-to-real time execution, (ii) in addition to ship motions, the vertical bending moment is predicted and (iii) the model is generalizable to many sea states, covering a variety of mild to severe conditions. As current state-of-the-art, a LSTM network similar to the one proposed by del Águila Ferrandis et al. (2021) was recreated and a comparison to the established models was drawn. The investigated networks were trained on simulation data. The waves were generated using linear wave theory for a variety of sea state conditions including steep waves of high amplitude. The vessel responses were simulated using a strip theory code with considerations of non-linearities in the calculation of the hydrostatic and Froude–Krylov forces. In order to assess the reliability of the model's predictions, the neural networks were tested on experimental data obtained from model tests of the vessel in a seakeeping basin. The testing involves mild sea state as well as the reproduction of the 'new year wave' as an example of a freak-wave.

The paper is organized as follows: Section 2 describes the numerical method used for data generation. Furthermore, physical constraints regarding the prediction horizon are discussed. The data-driven approach is addressed in Section 3, comprising the architectures of the established neural networks, the generation of the simulation data as well as the description of the model test setup. Following in Section 4, the models are investigated regarding the physical constraints and the resulting impact on the mapping capability is highlighted. Based on these findings, the final results are presented in Section 5 followed by the conclusion elaborated in Section 6.

2. Physics-based model for data generation

A sufficient amount of data is required for the training of neural networks. In order to keep the costs of data generation moderate it is reasonable to find a trade-off between computational time and accuracy of the data. Therefore, a strip theory solver by Fonseca and Guedes Soares (1998), see also Clauss et al. (2012), was selected, as it requires comparatively low computational effort while giving results that are sufficiently accurate with regard to the problem statement. The solver applies potential flow theory and non-linear effects due to large amplitude motions are considered in the hydrostatic and Froude–Krylov forces. The method is presented in more detail in the following. In addition to that, this section also covers the physical constraints of the mapping problem resulting from the point-wise input data.

2.1. Non-linear strip theory solver

A fixed coordinate system is defined as right-handed Cartesian coordinate system with z-axis pointing positive in the upward direction passing the center of gravity, x-axis pointing positive toward the bow and the y-axis pointing positive in port-direction perpendicular to the others. The origin is located in the plane of the undisturbed free surface. The flow is assumed to be frictionless and irrotational and thus the potential flow theory can be applied. The vector of velocities $\mathbf{U}$ can be written as gradient of the velocity potential $\mathbf{U} = \nabla\Phi$ with $\mathbf{x}$ being the vector of the Cartesian coordinates. Using a linear approach, the total velocity potential Φ^{T} can be simplified to the summation of steady and unsteady components

$$\Phi^{\mathrm{T}} = -Ux + \Phi^{\mathrm{S}} + \Phi^{\mathrm{J}} + \Phi^{\mathrm{D}} + \sum_{j=1}^{6} \Phi_j^{\mathrm{R}} \tag{1}$$

with U being the forward speed of the vessel and Φ^{S} the perturbation potential due to steady translation. Φ^{J} represents the incident wave potential, Φ^{D} the diffracted wave potential and Φ_j^{R} the radiation potential due to motion in the jth mode. Based on Eq. (1) and the linearized Bernoulli Equation, the hydrodynamic pressure can be approximated, resulting in the hydrodynamic pressure forces associated with oscillatory motions after integrating over the wetted surface area. These can be categorized into the radiation forces, hydrostatic forces and wave exciting forces; whereby the latter can be split into Froude–Krylov and diffraction forces. Large non-linearities due to the severe change in the submerged volume of the hull when advancing in high-amplitude waves are assumed to be dominated by buoyancy effects. Therefore, the hydrostatic pressures as well as the Froude–Krylov forces are integrated over the whole wetted surface area, including the deck in case of green water.

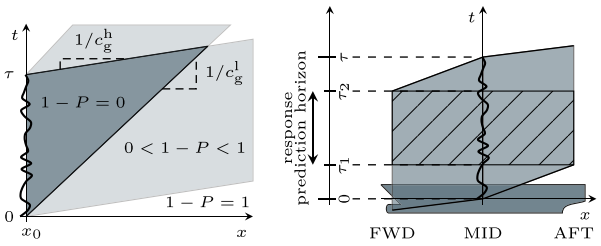
The rigid body motions of a ship consist of three translations of the origin and three rotations around the axes; however only head sea is modeled in this specific case, which reduces the ship motions to the vertical translatory motion heave X_3 and rotational motion pitch X_5. The ship motions are approximated by solving the coupled equations of motion in time domain. The vertical bending moment VBM is calculated at each cross-section based on the moment of inertia I_5 and the moments due to hydrodynamic forces

$$VBM(t) = I_5(t) - R_5(t) - D_5(t) - K_5(t) - H_5(t) \tag{2}$$

with contributions due to radiation R_5, diffraction D_5, Froude–Krylov force K_5 and hydrostatic force H_5. The moment contributions are integrated over the part of the hull that is forward of the respective cross section.

The solver models the irregular seaway by superposing linear wave components of varying heights and frequencies. A linear wave is characterized by the wave length L, wave period T, angular frequency ω, wave number k and phase speed c. The parameters are linked as follows

$$T = \frac{2\pi}{\omega},\quad L = T^2\,\frac{g}{2\pi},\quad k = \frac{\omega}{c}\,. \tag{3}$$



(a) Construction of the prediction zone (b) Definition of the response prediction horizon

Fig. 1. Prediction zone following linear wave theory: On the left-hand side, the theoretical prediction zone is outlined in dark gray for an observation of the surface elevation $\zeta(x_0, t \in [0, \dots \tau])$. Values of predictability $1 - P$ are given. The slope of the limits are determined by the group velocities c_g. The resulting prediction zone of the present case is shown on the right-hand side. The area in which the wave profile can be reconstructed along the entire hull is marked in dashed lines, resulting in non-predictability over parts of the hull outside of the response prediction horizon $\mathbf{T}_{\mathrm{Pred}} = [\,\tau_1, \dots,\ \tau_2\,]$. The x-axis denotes the direction of the flow.

In case of an uni-directional seaway, the superposition is simplified to

$$\zeta(x,\ t) = \sum_n \zeta_n \cos\left(\, k_n x - \omega_n t + \Phi_n \,\right) \tag{4}$$

with ζ_n, k_n, ω_n and Φ_n being the amplitude, wave number, angular frequency and phase shift of the nth wave component. The amplitude ζ_n is calculated based on the spectral density $S(\omega_n)$ in order to conserve the energy of a natural seaway.

2.2. Physical constraints of the prediction horizon

In case of a finite spatio-temporal observation of the wave field, the forecast of its development is only feasible for a specific time period (Morris et al., 1998). The spatio-temporal zone in which the seaway can be fully reconstructed based on an observation of the surface elevation $\zeta(x_0\ , t \in \mathbf{T} = [0, \dots, \tau])$ for a duration τ is called prediction zone and commonly visualized in a time-distance diagram, see Fig. 1(a). Its deterministic construction is addressed by Naaijen and Huijsmans (2008), Naaijen et al. (2014) and Wu (2004), among others. The predictability of linear waves can be measured deterministically by summing up the relative amount of energy arriving at position x at time t, leading to the theoretical prediction indicator

$$P(x,t) = \frac{\int_{\omega^{\mathrm{l}}}^{\omega^{\mathrm{h}}} S(\omega)\, d\omega}{\int_0^{\infty} S(\omega)\, d\omega} \tag{5}$$

with limits of integration ω^{l} and ω^{h} as lowest and highest angular frequencies of the wave field (Naaijen et al., 2014). Following Wu (2004), the prediction error indicator $1 - P \in [0, \dots, 1]$ instead of P is used in this study, whereby a value of $1 - P = 0$ indicates predictability. More specifically, the prediction zone will be defined as domain where $1 - P < 0.15$ holds. The value of 0.15 as limit of predictability has been chosen following Law et al. (2020), who investigated a data-driven method for wave propagation prediction.

The resulting prediction zone for the case presented in this paper is depicted in Fig. 1(b). At the temporal boundaries, the surface elevation is not predictable along the entire hull and therefore the wave profile is partly unknown by the neural networks, which affects the capability of the models to estimate the corresponding vessel responses. This leads to physical constraints of the mapping problem and restricts the time interval in which the models can be evaluated. This time interval will be referred to as *response prediction horizon* $\mathbf{T}_{\mathrm{Pred}} := [\tau_1, \dots,\ \tau_2]$.

3. Data-driven approach

Neural networks are a subset of Machine Learning and based on the idea of mathematically mimicking the behavior of biological neurons.

Supervised learning in a regression problem is characterized by the ability to learn a functional dependence $g : \mathbb{R}^{d_x} \rightarrow \mathbb{R}^{d_y}$ between labeled input $\mathbf{x}$ (features) and related output $\mathbf{y}$ (targets) of a training data set $\boldsymbol{\mathcal{D}} = \left\{\mathcal{D}_i = (\mathbf{x}, \mathbf{y})_i \,|\, \mathbf{x} \in \mathbb{R}^{d_x},\ \mathbf{y} \in \mathbb{R}^{d_y},\ i = 1, 2, \ldots, N\right\}$ consisting of N data samples. The dimensions of the feature and target space are depicted as d_x and d_y. During the training process, g is approximated such that the discrepancy between approximation and ground-truth is minimal according to a chosen metric. Neural networks are tuned to the distribution of the training data and do not extrapolate well to samples that lie beyond the training set. The latter should therefore cover the data space that is relevant for the application to ensure that the model can make reliable estimations. The capability of generalization is given if the model performs equally well on new, unseen data that has a similar distribution as the training set, also referred to as test set. The theory of neural networks is briefly summarized in the following section and the established neural network architectures are presented. More elaborated details on the theory of neural networks is given by Goodfellow et al. (2016).

3.1. Neural network theory

The most basic unit of a conventional neural network is the artificial neuron, which connects the sum of input $\mathbf{x}$ multiplied by weights $\mathbf{w}$ with the output $\mathbf{y}$ by an activation function f

$$\mathbf{y} = f(\mathbf{w},\ \mathbf{x}) = f\left(\text{bias} + \sum_{i=1}^{N} w_i x_i\right). \tag{6}$$

The activation function enables the neural network to establish non-linear relations between input and output. Arranging multiple neurons as a network, a multivariate function can be formulated as

$$\mathbf{y} = f_L\left(\mathbf{w}_L, f_{L-1}\left(\mathbf{w}_{L-1}, \ldots, f_1(\mathbf{w}_1, \mathbf{x})\right)\right), \tag{7}$$

where L corresponds to the number of layers. The most widespread neural network type is the multi-layer perceptron (MLP) consisting of the input and output layers as well as at least one hidden layer in between. MLPs are feed-forward, fully-connected neural networks, meaning the connections between the neurons are strictly directed toward the output layer and each neuron is connected to each neuron in the next layer. Another common architecture type is the convolutional neural network (CNN), which is a feed-forward model that is able to detect and extract features of interest from the input data by using the principle of convolution. During the process of a convolution, the input is filtered by a kernel and the extracted features are stored in feature maps. The kernel size determines the number of input data elements that gets further processed to compute one entry in the feature map. The region in the input data space that is seen by a certain feature map layer, called receptive field, can be expanded by using dilation, which inflates the increments of the kernel compared to the increment of the data. In case of only performing convolutional operations, the model is called fully convolutional neural network (FCNN).

During the training process, the model parameters are first initialized randomly and further optimized in order to represent the distribution of the data as best as possible within the capacity of the model. The discrepancy between prediction and ground-truth is measured by the loss function. The loss can be minimized by using gradient-descent and backpropagation. The choice of a suitable loss function depends on the underlying problem. In this study, the Surface Similarity Parameter

$$\text{SSP} = \frac{\sqrt{\int |F_{\mathbf{y}_1}(\mathbf{k}) - F_{\mathbf{y}_2}(\mathbf{k})|^2 d\mathbf{k}}}{\sqrt{\int |F_{\mathbf{y}_1}(\mathbf{k})|^2 d\mathbf{k}} + \sqrt{\int |F_{\mathbf{y}_2}(\boldsymbol{k})|^2 d\mathbf{k}}} \in [0,\ 1] \tag{8}$$

is applied, which compares two surfaces $\mathbf{y}_1$ and $\mathbf{y}_2$ in Fourier space, denoted as $F_{\mathbf{y}}$, and measures the similarity of the frequency components with respect to amplitude and phase. A value of SSP = 0 indicates perfect agreement and with increasing value of SSP, the similarity decreases. In a study of Wedler et al. (2022) it has been demonstrated that the SSP benefits machine learning models in the training of oscillatory sequences.

Table 1
Main particulars of the chemical tanker.

Parameter	Unit	Abbreviation	Full scale	Model scale
Length overall	[m]	L_{oa}	170	2.428
Length between perpendiculars	[m]	L_{pp}	161	2.3
Breadth	[m]	B_{WL}	28	0.4
Depth	[m]	D	13	0.186
Draught	[m]	D_{MS}	9	0.129
Displacement	[kg]	Δ	30 666 000	89.405
Block coefficient	[−]	c_{B}	0.75	0.75
Longitudinal center of gravity	[m]	L_{cg}	82.4663	1.17809
Vertical center of gravity	[m]	V_{cg}	6.1516	0.08788
Radius gyration for rolling	[m]	R_{xx}	9.31	0.133
Radius gyration for pitching	[m]	R_{yy}	32.9	0.47
Radius gyration for yawing	[m]	R_{zz}	33.53	0.479

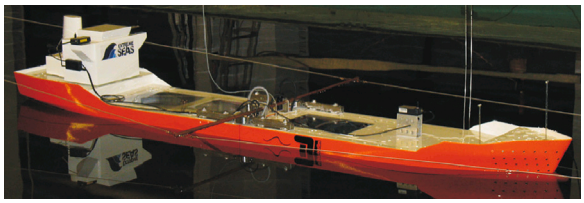


Fig. 2. Model of the chemical tanker in the seakeeping basin at the Technical University of Berlin.

3.2. Presentation of simulated and experimental data

The vessel featured in this study is a chemical tanker with main particulars listed in Table 1. Fig. 2 shows the model during conducted experiments at the Technical University of Berlin.

The vessel was part of a numerical and experimental study on the identification of critical wave sequences by Clauss et al. (2012). The authors applied an iterative optimization program using a strip theory solver in order to identify wave patterns that result in maximal vertical bending moments. The worst case scenario obtained from the optimization run is characterized by an extreme sagging condition with a wave length close to L_{pp} and a wave height at the limit of the wave steepness constraint. These extreme cases are of particular interest with regards of ship safety and should be well represented by the models. Therefore, the goal was to create a data set that is (i) representative of the real world and (ii) rich of information, i.e to include large vertical bending moments and large amplitude motions. Thus, the ranges of the Jonswap parameters were chosen based on the findings of the study mentioned above and a wave steepness constraint. The ratio between peak wave length to ship length $r := L_{\text{p}}/L_{\text{pp}} \in \{0.6, \ldots, 1.5\}$ and the wave steepness $\epsilon := \pi H_{\text{s}}/L_{\text{p}} \in \{0.025, \ldots, 0.15\}$ are introduced. Values of $r \approx 1.1$, meaning a peak wave length L_{p} close to ship length L_{pp}, were included in order to provoke large bending moments. The ranges of the peak enhancement factor were set to $\gamma \in \{1, \ldots, 7\}$. A total number of 5 000 Jonswap spectra were generated by defining vectors $\boldsymbol{\epsilon}$, $\mathbf{r}$ and $\boldsymbol{\gamma}$ with random values uniformly distributed in the corresponding range. Next, the spectral parameters were derived, whose histograms are shown in Fig. 3. The joint distribution of H_{s} and T_{p} is given in Fig. 4, compared to measured sea state statistics of the Northern Atlantic. The distributions of simulated data and measurements differ and the simulated sea state conditions are more evenly distributed inside of the defined range.

The frequency spectrum was discretized in $0\,\text{rad/s} < \omega \leq 8\,\text{rad/s}$ using $\Delta\omega = 0.01\,\text{rad/s}$. The time domain was discretized for a duration of $T = 180\,\text{s}$ in 512 time steps, resulting in a time step size of $\Delta t \approx 0.35\,\text{s}$. The dimensions of one datasample therefore equals $D = \left\{\ (\mathbf{x}, \mathbf{y}) = \left(\boldsymbol{\zeta}(x_{\text{mid}}), \boldsymbol{X}_3, \boldsymbol{X}_5, \boldsymbol{VBM}\right) \in \mathbb{R}^{512\times 4}\ \right\}$. The speed of the ship was set to zero and one loading condition was modeled. Next to $\zeta(x_{\text{mid}},\ t)$, the strip theory solver also prints out the surface elevation at forward perpendicular $\zeta(x_{\text{fwd}},\ t)$ and aft perpendicular $\zeta(x_{\text{aft}},\ t)$. Fig. 5 shows

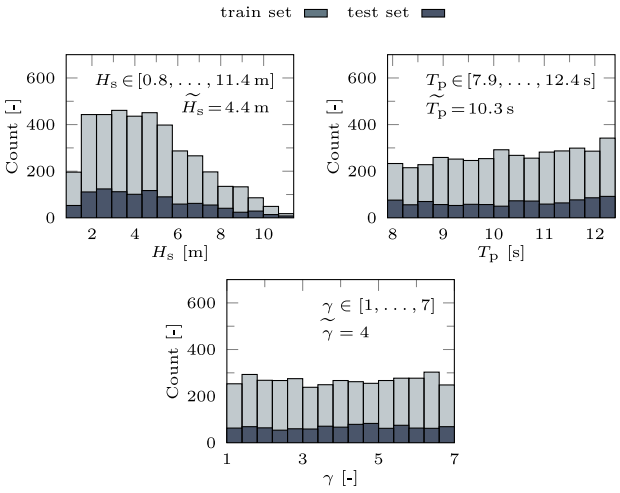
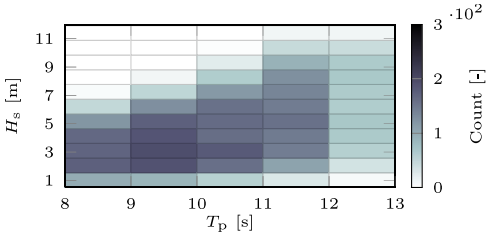
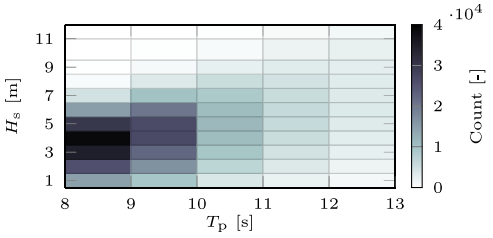


Fig. 3. Distribution of significant wave height H_s, peak wave period T_p and peak enhancement factor γ in training and test set. Data ranges and their median are given.



(a) Joint distribution of H_s and T_p in the data set



(b) Sea state characteristics of Northern Atlantic

Fig. 4. Comparison of significant wave height H_s and peak period T_p occurences in simulated data and hindcasted wave spectra obtained from measurements in the Northern Atlantic (Söding, 2001).

the distribution of the simulated response amplitudes. The values were extracted by splitting each time signal into 4 bins and taking the maximum value of each bin.

The model tests of the chemical tanker were conducted in a scale of 1 : 70 at the seakeeping basin of the Ocean Engineering Division of the Technical University of Berlin. The basin has a length of 110 m with a measuring range of 90 m. The width is 9 m and the water depth is 1 m. The basin is equipped with an electrically driven piston type wave generator able to generate deterministic irregular sea states with predefined characteristics and tailored wave sequences. Only the heave and pitch motions of the model were unrestrained with the other degrees of freedom being restricted by springs. The motions of the model were recorded by an optical tracking system. Nine wave probes were installed, one of them being located amidships. The vertical bending moment and longitudinal forces were obtained by three force transducers at $L_{pp}/2$, two being mounted on deck and the third one mounted underneath the keel. The measured model scale data was converted to full scale by applying Froude similarity law in a

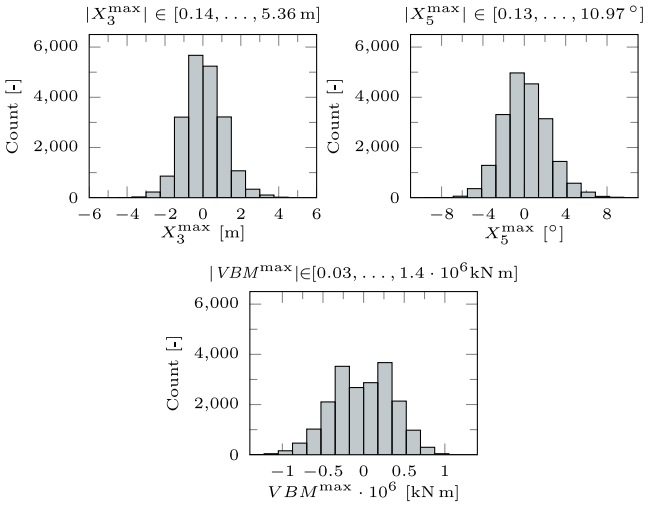


Fig. 5. Distribution of heave X_3, pitch X_5 and vertical bending moment *VBM* amplitudes in the simulated data set.

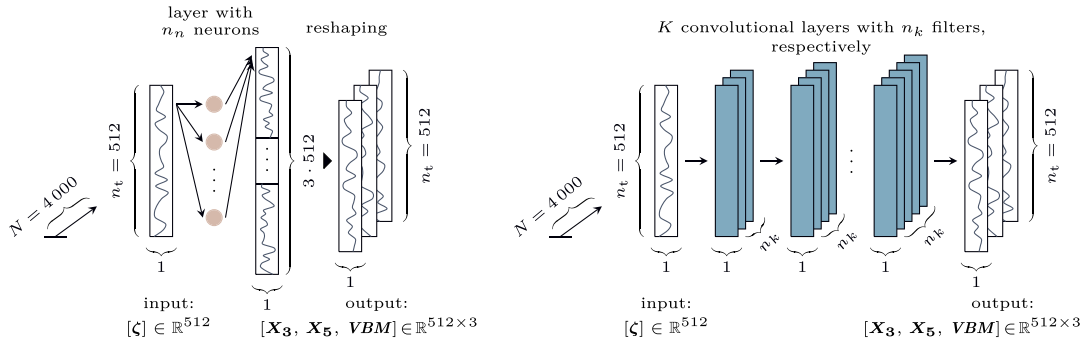
Table 2
Settings of the neural network models.

Parameter	Unit	MLP	FCNN	LSTM
Size of training set	[–]	4 000	4 000	4 000
Size of test set	[–]	1 000	1 000	1 000
No. of trainable parameters	[–]	131 264	57 113	4 624
No. of hidden layers	[–]	1	6	5
Batch size	[–]	64	64	64
Time elapsed per prediction of one sample	[s]	0.0131	0.0213	–

preprocessing step. The experimental setup is explained in more detail by Wang et al. (2024).

3.3. Presentation of neural network models

For all networks, a hyperparameter study was carried out in order to find the best fit regarding model complexity, see Appendix B. The search for appropriate hyperparameters comprises the variation of the depth of the networks, the number of filters and neurons in the layers, different activation functions as well as loss functions. The study is performed on a training and validation set using the four-fold cross validation method. The capability of generalization is ensured using a training and a fixed test set comprising 20% of the generated data, see Appendix C. In terms of data preparation, the data was scaled to a range between 0 and 1 using min–max normalization. The minimum and maximum values for normalization were extracted from the training set. For both models, the SSP as loss function was applied. The hyberbolic tangent was used as activation function, except for the last layers, which hold no activation. The Adam optimizer with exponential learning rate was employed. The initial learning rate was set to 0.01 and the decay rate to 0.98. The batch size was 64. Furthermore, for each layer batch normalization was applied with a momentum of 0.8. The MLP, shown in Fig. 6(a), consists of one densely connected hidden layer holding $n_n = 64$ neurons. Fig. 6(b) shows the architecture of the FCNN. $K = 6$ convolutional layers are applied and the dimensions of the feature maps are kept equal to the input layer dimension by applying zero padding. The number of feature maps of each layer increases with the depth of the model. The kernel size was set to 5 with exponentially growing dilation rate of base 2. The final settings of the models are listed in Table 2.



(a) MLP (b) FCNN

Fig. 6. Architectures of the multi-layer perceptron and fully connected neural network for mapping the input signal of the surface elevation ζ amidships and the output signals of heave X_3, pitch X_5 and the vertical bending moment VBM.

4. Assessment of the response prediction horizon

This section is devoted to recreating the theoretical assumptions of the prediction horizon as described in Section 2.2. The goal is to understand and quantify the resulting temporal limitations of the mapping task. For this, a preliminary study has been carried out in two steps. First, the neural networks were tested in terms of their capability to represent the linear propagation of waves addressing the classical prediction zone, see Fig. 1(a). Second, the mapping is applied for the ship response prediction in order to assess the impact of the wave prediction zone on the mapping capability. The observation time of the input time series was chosen smaller compared to the time series to be predicted (output) in order to identify the prediction zone and its spatio-temporal development clearly. Thus, the time interval of the input data is restricted by dropping the first and last 45 s of the $\tau = 180\,\mathrm{s}$ time frame, resulting in a mapping problem with an input time interval of $\mathbf{T}_{\mathrm{Input}} = [45\,\mathrm{s},\ 135\,\mathrm{s}]$ and an output time interval of $\mathbf{T}_{\mathrm{Output}} = [0\,\mathrm{s},\ 180\,\mathrm{s}]$. It is noted here that this restriction is only applied in the studies carried out in this section and that for the final results, both models were trained using the whole data time interval.

The objective of this study is to investigate the mapping error over time. The results are therefore presented by calculating the mean absolute error MAE between ground-truth y and prediction y' of the N datasamples in the test set along the temporal axis

$$\mathrm{MAE}\,(t_i) = \frac{1}{N}\sum_{i=1}^{N} |y'(t_i) - y(t_i)|,\ \ i = 1, \dots, 512\,. \tag{9}$$

Furthermore, by norming the error to a range between 0 and 1, the MAE can be compared to the physical prediction error indicator $1 - P$ and therefore be used analogously as data-driven prediction error indicator. It is emphasized that the MAE is applied to show the progression of the error over time and not to make a qualitative statement about the individual mapping problems. The following investigation are carried out using both neural networks, however only the results of the MLP are presented for reasons of redundancy.

4.1. Prediction zone of linear waves

In the first study, the mapping task involves the reconstruction of the surface elevation $\zeta(x_{\mathrm{fwd}})$ and $\zeta(x_{\mathrm{aft}})$ based on $\zeta(x_{\mathrm{mid}})$. The setup replicates the case shown in Fig. 1(b), which illustrates that the predictability of the wave profile along the entire hull is limited in the beginning and end of the time signal. The architecture of the MLP itself has not been modified and only the input and output layers have been changed to the surface elevation data. The MAE representing the data-driven prediction error indicator and the theoretical error indicator are plotted in Fig. 7(a). Figuratively speaking, the curves inside of the white area represent how the predictability of the wave profile propagates in time. Overall, the curves of the MAE show agreement with the theoretical prediction indicator. The response prediction horizon, as illustrated in Fig. 1(b), is marked where $1 - P < 0.15$ holds simultaneously for both of the predictions of $\zeta(x_{\mathrm{fwd}})$ and $\zeta(x_{\mathrm{aft}})$, meaning when the wave profile is fully predictable. This condition is satisfied for $\mathbf{T}_{\mathrm{Pred}} = [45\,\mathrm{s} + \Delta\tau, \dots,\ 135\,\mathrm{s} - \Delta\tau] = [60\,\mathrm{s}, \dots,\ 120\,\mathrm{s}]$ with $\Delta\tau = 15\,\mathrm{s}$ as cut-off time interval. The results in terms of the SSP are evaluated inside of the established response prediction horizon and presented in Table 3. The median SSP lies in the range of $0.044 - 0.050$, which indicates accurate mapping results. One test set datasample is illustrated in Fig. 8, giving a more vivid impression of the mapping accuracy with close agreement between ground-truth and prediction where $1 - P < 0.15$ holds. The choice of 0.15 as the limiting value of predictability is therefore adequate in this case.

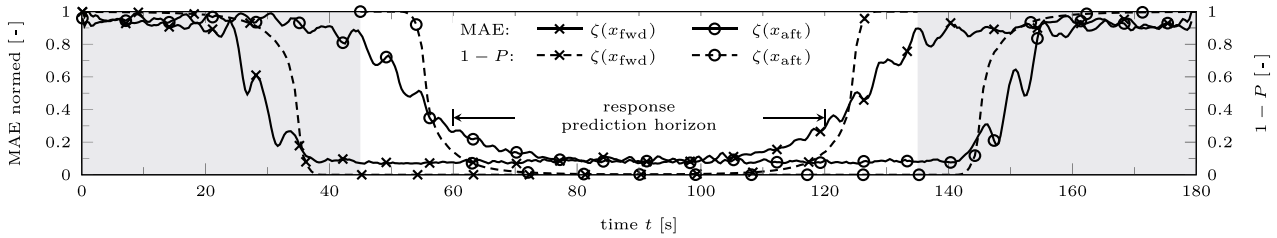
4.2. Response prediction zone

In the second step, the study shown above was repeated for mapping the vessel responses X_3, X_5 and VBM. The results are presented in Fig. 7(b) along with the results of the previous study from Fig. 7(a). At $t = 0\,\mathrm{s}$, the deviation between ground-truth and prediction for X_3 and X_5 is zero due to the vessel's initial position in rest. The MAE of the responses lie in between the ones of the surface elevation. The results show that the definition of the response prediction horizon is sufficient for cutting off the mapping error at the boundaries. Signals longer than 180 s can be evaluated by applying the model multiple times using the principle of a sliding window, as illustrated in Fig. 9. The window is shifted by the length of the prediction horizon, leaving a time segment of 15 s at the beginning and end of the signal that cannot be predicted by the model. It should be noted here that the surface elevation could also be given at other places than the midship section, for example at the forward perpendicular or in front of the ship, which would shift the response prediction horizon backwards in time. Another approach would be to provide the spatio-temporal wave field as input following Lee et al. (2023).

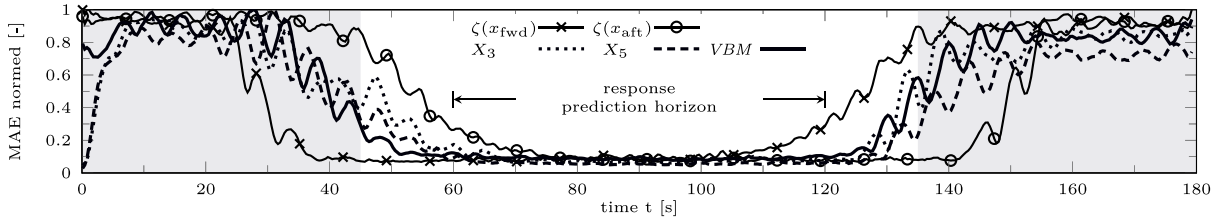
5. Results

The performance of the models are evaluated on simulation as well as experimental data. The results of the simulation data is going to be presented first.

In practice, measured data contain noise. In the first step, the models were investigated regarding noise sensitivity. Random zero-centered Gaussian noise was added to the simulation data of the training and test set features, meaning the surface elevation data. Note that the targets were left unmodified such that the models were trained to map the smooth signal as output. Three different standard deviations of Gaussian distribution $\sigma \in \{0.25, 0.5, 0.8\}$ applied to the 180 s long time series



(a) Normalized MAE for mapping surface elevation data in solid lines, compared to the theoretical prediction indicator $1 - P$ in dashed lines



(b) Normalized MAE for mapping vessel responses X_3, X_5 and VBM using the surface elevation amidships

Fig. 7. Assessment of the prediction zone for sea state mapping (top) and vessel responses (bottom) using the MLP. The normalized mean absolute error (MAE) is plotted along the temporal axis. The gray area marks the time frame outside of the time input interval. The response prediction horizon, see Fig. 1(b), is given.

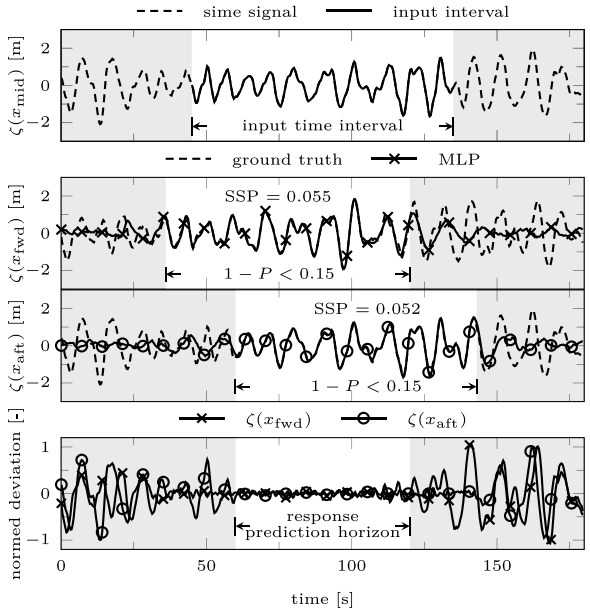


Fig. 8. Example of sea state mapping: Based on the surface elevation at midship in time interval $\mathbf{T}_{\text{Input}}$ (top), the surface elevation at forward perpendicular and aft perpendicular is determined (middle). Ground-truth is plotted in dashed lines and the mapped results are plotted in solid lines. The domain where $1 - P < 0.15$ holds is marked and values of SSP are given. At the bottom, the normalized deviation between ground-truth and prediction is plotted as well as the response prediction horizon.

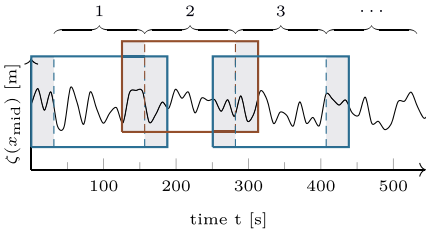


Fig. 9. Sliding window technique for predicting time signals longer than 180 s.

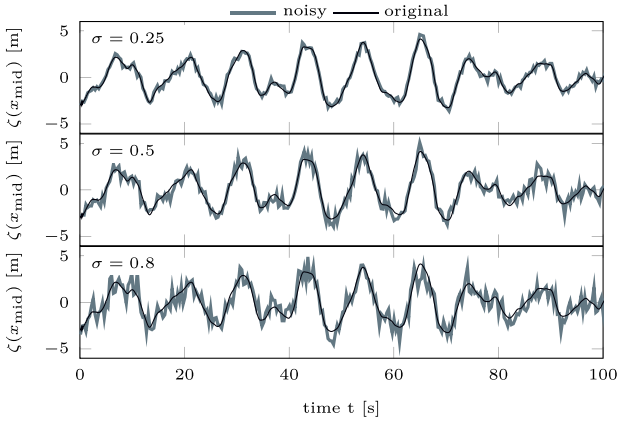


Fig. 10. Examples of applied Gaussian noise to the features of the simulation data set.

were investigated. Fig. 10 shows an example of the modified features with different noise levels. Both models were trained four times with the fixed test set and the results of the SSP loss are plotted regarding the mean and standard deviations in Fig. 11. The prediction accuracy of the models get worse with increasing noise level. Furthermore, the fluctuations in the results of the four runs increase with increasing noise level. Fig. 12 shows the mapping results of the vertical bending moment for the two cases of (i) no noise and (ii) $\sigma = 0.8$. Both models are able to remove the noise in the estimations. The prediction using noisy input data tends to under- and overestimate the ground-truth data, however is still able to make reliable estimations. In summary, the models are able to address noise in the input data to a certain level.

From the four runs carried out on the smooth input data, the best models were selected based on the performance on the test set. The results of the selected models are given in Table 3 in terms of the median SSP and the corresponding box plots are shown in Fig. 13. In addition to that, the results of a LSTM architecture inspired by the studies of del Águila Ferrandis et al. (2021) are given, serving as state-of-the-art. The MLP and FCNN perform similarly well compared to each other as well as regarding the three targets X_3, X_5 and VBM, as the

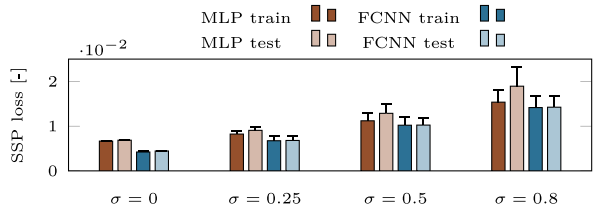


Fig. 11. Bar plots of the noise sensitivity study for the different noise levels. The SSP losses are plotted as mean and standard deviation.

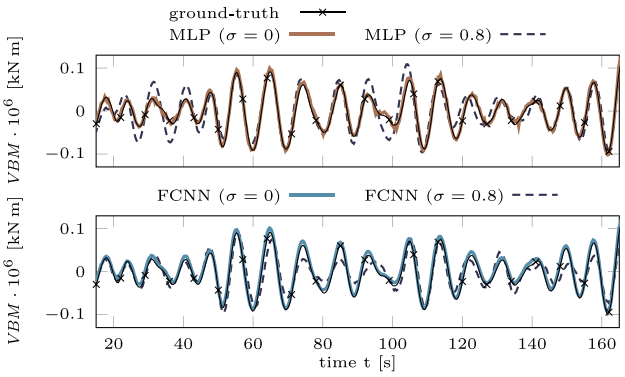


Fig. 12. Example of the mapping performances for noise levels $\sigma = 0$ and $\sigma = 0.8$. The MLP achieves results of SSP= 0.1 and SSP= 0.39, the predictions of the FCNN yields SSP= 0.12 and SSP= 0.18.

Table 3
Median SSP error for predictions on training and test set. The results of the preliminary study regarding the mapping of the surface elevation, see Section 4, are only given for the MLP.

Model	Data set	Sea state mapping		Vessel response prediction		
		$\zeta(x_{\text{fwd}})$	$\zeta(x_{\text{aft}})$	X_3	X_5	*VBM*
MLP	train	0.049	0.044	0.0749	0.0701	0.0675
	test	0.050	0.046	0.0764	0.0741	0.0709
FCNN	train	–	–	0.0352	0.0671	0.0253
	test	–	–	0.0343	0.0711	0.0255
LSTM	train	–	–	0.2529	0.2108	0.2531
	test	–	–	0.2505	0.2178	0.2561

median error is in the range of SSP $\approx 0.03 - 0.08$ and the variances are similar for all cases. In general, the FCNN achieves lower values for the SSP than the MLP for all targets. The FCNN is performing worst on the prediction of the pitch motion compared to heave and *VBM* with a median SSP of SSP ≈ 0.07. Compared to the LSTM, the proposed MLP and FCNN achieve more accurate mapping results. However, in the comparison of the three models it should be pointed out that they contain different amounts of trainable parameters. In particular, the LSTM has the lowest number of parameters, possibly affecting its possibility to represent the data. In Fig. 14, the performance of both models is evaluated in dependence of the sea state characteristics. The test set samples were sorted by splitting the axis parameters ratio of peak wave length to ship length r and wave steepness ϵ into 7 bins. For each bin, the average SSP of the corresponding predictions was determined. It can again be observed that the FCNN achieves lower values for the SSP as the MLP. No clear dependence of the prediction accuracy on the sea state conditions can be observed.

In the next step, the models were tested on the experimental data in order to assess their reliability. Two time series were selected from the experiments carried out and the models were applied to the measured

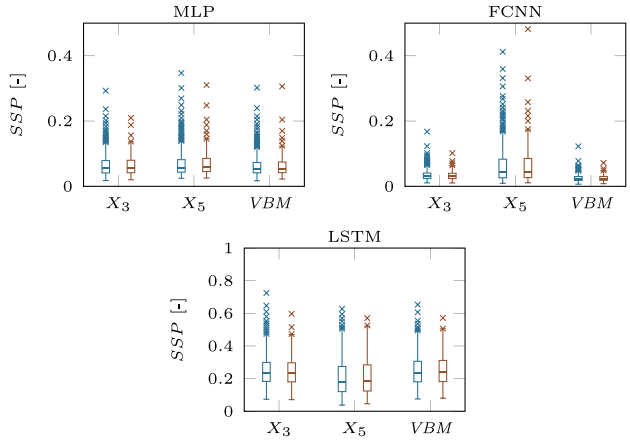
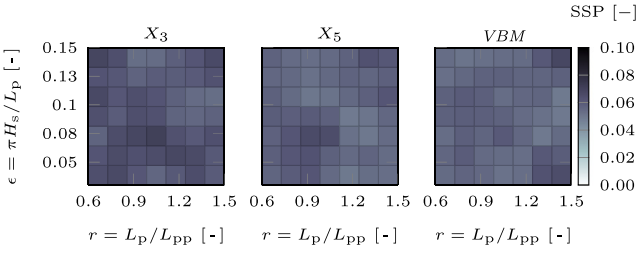
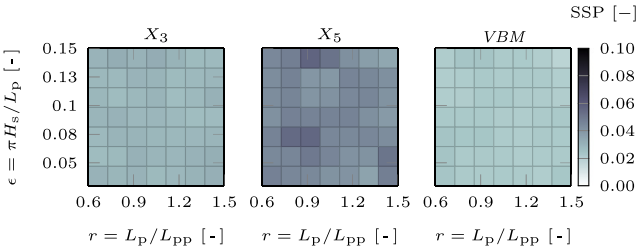


Fig. 13. Box plots of the SSP error from the results presented in Table 3. The left box represents the results of the training set, the right box the ones of the test set.



(a) MLP



(b) FCNN

Fig. 14. SSP error in dependence of wave steepness ϵ and ratio of peak wave length to ship length r for each target X_3, X_5 and *VBM*.

surface elevation from the wave gauge located at the midship section using the sliding window technique. The examples selected are, on the one hand, a sea state condition that is well represented in the training set of the models (scenario 1 with $H_s = 3\,\text{m}$, $T_p = 12\,\text{s}$, $\gamma = 3.3$) and, on the other hand, an extreme wave, the so-called 'new year wave' (scenario 2). In a separate step, the strip theory code was also applied to the experimental data and its results are included in the comparison shown in Figs. 15 and 16. The SSP given in the figures are determined for the strip theory code (STC) as well as the neural network models with respect to the experimental data as ground-truth. In the first example in Fig. 15, the models and the strip theory code give accurate predictions of the vessel responses. It is noticeable that some amplitude peaks are underestimated by all three models, as e.g. the heave motion around $t \approx 1\,850\,\text{s}$. Here the neural networks show more agreement with the results of the strip theory code. This is to be expected, as the neural networks will tend to follow the numerical mathematics of the strip theory code. The second example in Fig. 16 represents an extreme

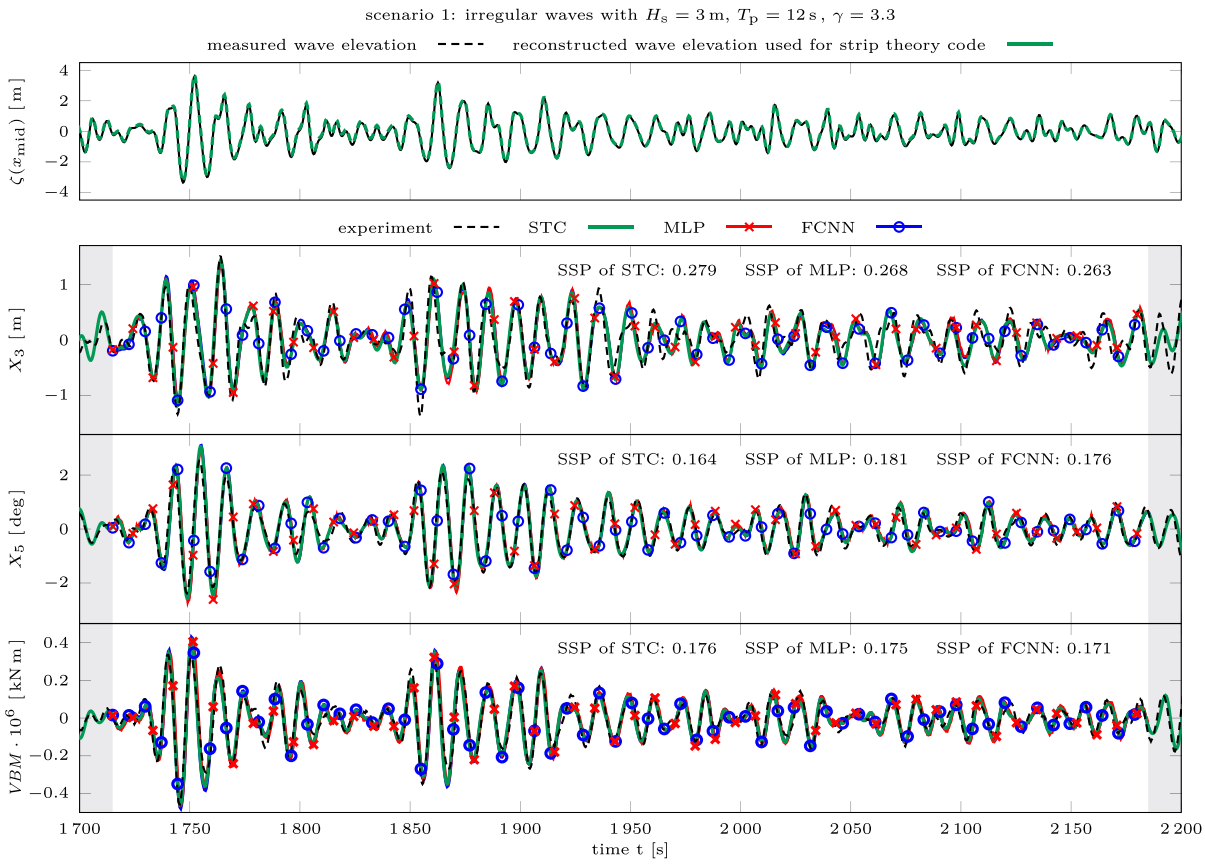


Fig. 15. Performance of the MLP and FCNN on experimental data for a sea state with $H_s = 3\,m$, $T_p = 12\,s$ and $\gamma = 3.3$. On the top, the measured surface elevation amidships is plotted. In the following three plots, the measurements of heave X_3, pitch X_5 and bending moment *VBM* are shown. The predictions of the MLP and FCNN are plotted in color and compared to the simulation results of the strip theory code (STC). The time frame outside of the response prediction horizon is marked in gray.

event reaching a maximal crest height of $\zeta_c^{max} = 18.5\,m$. Such extreme events are of particular interest as their prediction should be reliable, although being rare. Here, the strip theory code is able to estimate the vessel responses accurately. The neural networks, however, although making accurate predictions for the majority of the time signal, tend to underestimate the extreme amplitudes resulting from the freak wave, which can be observed at time $t \approx 850\,s$ for all three targets. Here the deviation between experimental data and prediction is approximately 30%. This behavior might be linked to the domain shift in this case, as the wave height as well as the response amplitudes are outside of the distribution of the training data, c.f. Figs. 3 and 5.

6. Conclusions

This study introduces two neural network models, a MLP and FCNN, as data-driven non-linear models for predicting the heave and pitch motions and the wave-induced loads of a vessel in irregular, long-crested seaway based on surface elevation data amidships. The on-board prediction of ship responses requires a model that is able to accurately capture non-linearities that strongly impact the vessel responses in severe sea state conditions and functions in close-to real time. The training data was generated using a seakeeping code in time-domain with zero forward speed of the vessel and one loading condition. The data set comprises a variety of sea state conditions including steep waves of large amplitude. The equations of motions are solved using strip theory and the solver computes the hydrostatic and Froude–Krylov forces over the instantaneous wetted surface in order to model non-linear effects in large-amplitude waves. The impact of the response prediction horizon on the mapping capability was highlighted and the resulting physical constraints of the models have been taken into account in the final evaluation. The proposed models were compared to a state-of-the-art LSTM architecture. The reliability of the predictions was investigated on simulation data as well as on experimental data obtained from model tests of the vessel in a seakeeping basin.

Both neural networks give promising results and perform similarly well in terms of prediction accuracy with average error below of SSP$\leq$ 0.08 for all three ship responses, surpassing the chosen state-of-the-art model. A strength of the method applied here is the ability of the neural networks to cover a large range of sea state conditions. The importance of the domain shift when applying neural networks to data outside of the distribution of the training data can be observed on the example of the 'new year wave'.

The drawbacks of deep learning models are (i) the large amount of data for training and testing required and (ii) that the models applicabilities are limited to the distribution within the training data, as seen on the example of the 'new year wave'. Therefore, further research and gathering of data is required. Most notably, the presented method is yet limited to one loading condition at zero forward speed. The problem can be extended by generating a data set containing the vessel's responses with regards to more degrees of freedom in short-crested sea state as well as forward speed and multiple loading conditions.

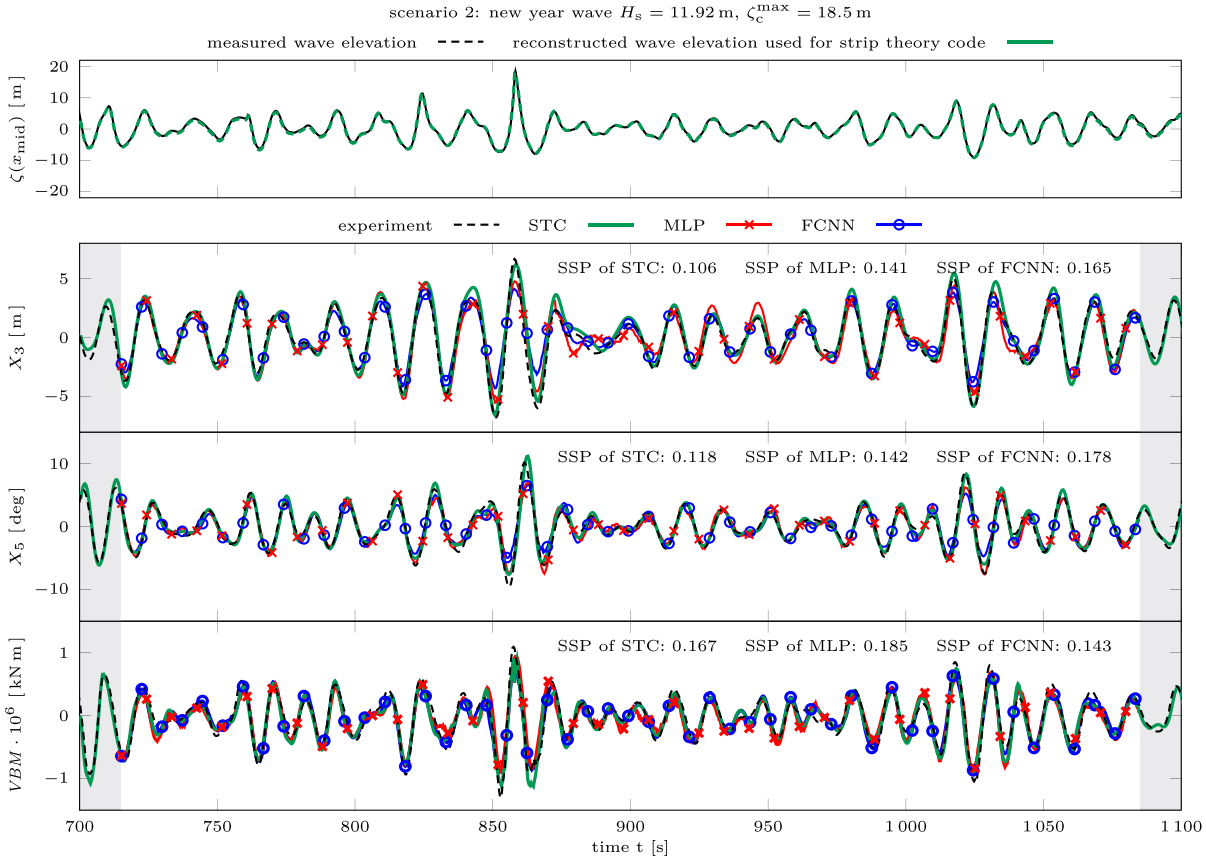


Fig. 16. Performance of the MLP and FCNN on experimental data of the 'new year wave' reaching a maximal crest height of $\zeta_c^{max} = 18.5\,m$. On the top, the measured surface elevation amidships is plotted. In the following three plots, the measurements of heave X_3, pitch X_5 and bending moment *VBM* are shown. The predictions of the MLP and FCNN are plotted in color and compared to the simulation results of the strip theory code (STC). The time frame outside of the response prediction horizon is marked in gray.

Lastly, the wave elevation is assumed to be given at high accuracy. In later application, however, the wave data will be determined from wave propagation models. Therefore, the accuracy of the ship response prediction will depend on the quality of the wave model. The coupling of wave propagation model and ship response prediction was not part of this study and has to be addressed in future.

CRediT authorship contribution statement

Johanna Serr: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Mathies Wedler:** Writing – review & editing, Validation, Supervision, Investigation. **Merten Stender:** Writing – review & editing, Validation, Supervision, Investigation. **Nuno Fonseca:** Software, Methodology. **C. Guedes Soares:** Software, Methodology. **Norbert Hoffmann:** Supervision, Resources, Methodology. **Sören Ehlers:** Supervision, Resources, Methodology. **Marco Klein:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This work was supported by Hamburg University of Technology (TUHH) [I^3-project "Predicting Ship Hydrodynamics to Enable Autonomous Shipping: Nonlinear Physics and Machine Learning"]; and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - grant number 277972093.

Appendix A. Literature research

Table A.4 summarizes the literature research given in Section 1.

Appendix B. Hyperparameter study

The neural networks were implemented and trained using TensorFlow (Abadi et al., 2015) version 2.11.0 and Keras (Chollet et al., 2015). The training was executed on a NVIDIA graphic card of type GP102 with CUDA version 11.7 (NVIDIA et al., 2020). The hyperparameter study was conducted using a four-fold cross-validation approach with an independent test set. First, the data set of $N = 5\,000$ samples was divided into a test set holding 20% or $N_{test} = 1\,000$ samples, cf. Fig. 3, and the remaining 4 000 samples being further split into four equal-sized parts using a stratified data split technique. In each cross-validation run, three parts functioned as training set and the remaining part as validation set. The test set was excluded from the investigation. The early-stopping approach monitoring the validation loss was implemented with a patience of 50 epochs starting after 700 epochs. The weights of the model corresponding to the lowest validation loss were restored. The results of the four-fold cross-validation hyperparameter study are presented below in Tables B.5 and B.6 for the MLP and FCNN. In addition to these two networks, a hyperparameter study of a LSTM architecture (del Águila Ferrandis et al., 2021) was carried out, whose results are presented in Table B.7. The SSP as loss function and hyberbolic tangent as activation function were applied. The best model was selected based on the performance on the validation set. In addition to varying the architecture in terms of depth, different loss and activation functions were employed in a second study for the final MLP architecture and the results are listed in Table B.8.

Table A.4
Examples of ship motion prediction models present in literature. The literature research includes three types of objectives; that are ship motion forecast, prediction of roll hydrodynamics as well as the mapping of wave and vessel response data. The table distinguishes between simulated (sim.) and experimental (exp.) data, as well as long-crested (l.-c.) and short-crested (s.-c.) in various sea state conditions (var. cond.).

Forecasting ship motions					
Input	Data type	Sea state	pred. zone	Prediction model	Remarks
motion data, ship hydrody-	sim.	l.-c.	5 s	Kalman filter Triantafyllou and Athans (1981), Triantafyllou et al. (1983)	noise-insensitive, estimation accuracy depending on
namics, H_s, ω_p					ω_p
motion data	exp.	s.-c, var. cond.	2 – 4 s	ARMA (Yumori, 1981)	applicable for motion compensation equipment and aircraft landing
wave data	exp.	l.-c. var. cond.	≡ theor. pred. zone	lin.wave theory, RAOs (Naaijen and Huijsmans, 2008)	
wave data	exp.	s.-c,, var. cond.	120 s	lin.wave theory, RAOs (Dannenberg et al., 2010)	tested on full-scale vessel in mild sea state conditions
motion data	exp.	l.-c., var. cond.	5 s	HODMD (Chen et al., 2023b)	prediction of ship maneuvering motion
motion data	exp.	s.-c., sea state code 5 – 6	7 s	FCNN (Khan et al., 2007)	prediction of roll angle
wave and motion data	exp.	l.-c., var. cond.	46.5 s	LSTM-based model (Guo et al., 2021)	noise-insensitive, tested using wave and solely motion data as input
motion data	sim.	s.-c., var. cond.	100 s	LSTM + CNN model (Lee et al., 2022)	able to consider motion memory effects
motion data	exp.	–	50 s	LSTM + GR model (Sun et al., 2022)	prediction of ship rolling and pitch angle, further testing in different sea state conditions necessary
motion data	CFD	l.-c., var. cond.	15 s	BiLSTM (Jiang et al., 2024)	pitch and heave prediction of KCS
motion data	exp.	l.-c., var. cond.	60 s	RBF (Yin et al., 2018)	ship roll prediction
motion data	exp.	l.-c.	30 s	RC (Yang et al., 2023a)	prediction of surge, heave and pitch motions
Prediction of roll hydrodynamics and wave propagation forecast					
Input	Data type	Sea state	pred. zone	Prediction model	Remarks
wave time series, roll hydrodynamics	exp.	l.-c.	5 s	SVM (Sclavounos and Ma, 2018)	able to predict roll hydrodynamics at different loading conditions
Mapping of wave data and ship responses (pitch, heave and roll)					
Input	Data type	Sea state		Prediction model	Remarks
wave time series	CFD	l.-c., sea state code 1&8		RNN, GRU, LSTM (del Águila Ferrandis et al., 2021)	LSTM superior, however limited to chosen sea state conditions

Appendix C. Learning curves

After determining appropriate hyperparameters, the models were further investigated regarding under- or overfitting. The data of the training and validation sets from the prior hyperparameter study were merged into a new training set. Next, the models were trained on the data from the new train set using different train set sizes and evaluated on the unseen test set. The test set was the same for the investigations. The training was carried out four times for each variation. Table C.9 depicts the results of the investigation. By lowering the amount of data samples in the train set, the MLP model tends to overfit as the train loss decreases while the test loss increases. For the FCNN, both the train loss as well as the test loss increases with decreasing train set size. This is an atypical behavior in a study of learning curves and a reason for this was not found. Fig. C.17 shows exemplary loss curves.

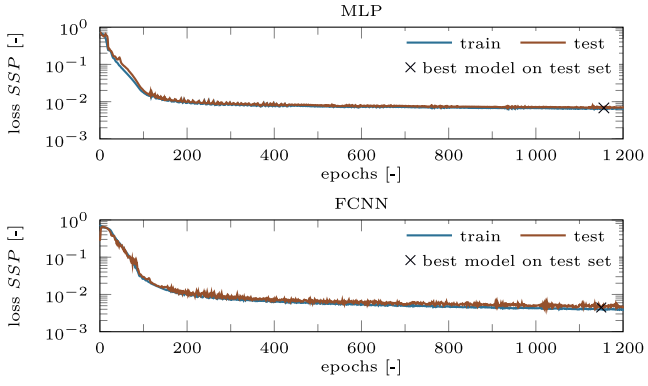


Fig. C.17. Loss curves of the selected model settings. The model achieving the lowest test loss is marked. The y-axes are plotted in logarithmic scale.

Table B.5
Results of the hyperparameter study for the MLP architecture. Two model depths n_d with varying number of neurons n_n are investigated. Two approaches are selected for the variation of the number neurons: first doupling the number of neurons in each layer and second keeping the number constant. The amount of model parameters are given as well as the mean number of epochs associated with the best model evaluated on the validation set. The performance of the models are measured by the SSP loss for the train and validation set in terms of mean and standard deviation of the four-fold cross-validation runs. The best model performing on the validation set is highlighted in the table.

MLP hyperparameters				SSP loss	
depth n_d	# neurons of each layer n_n	# weights	# epochs	train	val
1	[16]	32 848	1 148	0.02357 ± 0.00105	0.02387 ± 0.00152
	[32]	65 696	1 047	0.00926 ± 0.00027	0.00944 ± 0.00037
	[64]	131 392	1 026	0.00667 ± 0.00010	0.00689 ± 0.00034
	[128]	262 784	1 106	0.00619 ± 0.00018	0.00692 ± 0.00030
	[256]	592 384	1 082	0.00608 ± 0.00015	0.00731 ± 0.00012
2	[16, 32]	58 096	1 170	0.02330 ± 0.00074	0.02392 ± 0.00058
	[32, 64]	117 216	1 379	0.01077 ± 0.00042	0.01177 ± 0.00015
	[64, 128]	238 528	1 318	0.01023 ± 0.00083	0.01148 ± 0.00107
	[128, 256]	493 440	1 042	0.00994 ± 0.00050	0.01194 ± 0.00042
2	[16, 16]	33 184	1 071	0.00906 ± 0.00039	0.01188 ± 0.00037
	[32, 32]	66 880	1 182	0.02374 ± 0.00092	0.02430 ± 0.00114
	[64, 64]	135 808	1 410	0.00995 ± 0.00018	0.01046 ± 0.00038
	[128, 128]	279 808	1 487	0.00949 ± 0.00056	0.01039 ± 0.00054

Table B.6
Results of the hyperparameter study for the FCNN architecture. For three model depths n_d, different filter compositions as well as kernel sizes s_k are investigated. Two approaches are selected for the variation of the number of filters n_k: first doupling the number of filters in each layer and second keeping the number constant. The amount of model parameters are given as well as the mean number of epochs associated with the best model evaluated on the validation set. The performance of the models are measured by the SSP loss for the train and validation set in terms of mean and standard deviation of the four-fold cross-validation runs. The best model performing on the validation set is highlighted in the table.

FCNN hyperparameters					SSP loss	
depth n_d	# # filters in layers n_k	kernel size s_k	# weights	# epochs	train	val
3	[4, 8, 16]	3	887	885	0.01199 ± 0.00047	0.01241 ± 0.00044
		5	1 215	1 039	0.00749 ± 0.00044	0.00757 ± 0.00053
4	[4, 8, 16, 32]	3	2 823	1 076	0.00651 ± 0.00020	0.00649 ± 0.00033
		5	4 175	1 088	0.00483 ± 0.00016	0.00500 ± 0.00018
5	[4, 8, 16, 32, 64]	3	9 767	1 037	0.00532 ± 0.00017	0.00538 ± 0.00027
		5	15 215	1 084	0.00459 ± 0.00028	0.00483 ± 0.00033
3	[8, 16, 32]	3	2 719	881	0.01182 ± 0.00036	0.01218 ± 0.00052
		5	4 015	970	0.00746 ± 0.00026	0.00766 ± 0.00038
4	[8, 16, 32, 64]	3	9 663	972	0.00621 ± 0.00032	0.00640 ± 0.00049
		5	15 055	1 033	0.00467 ± 0.00013	0.00484 ± 0.0001
5	[8, 16, 32, 64, 128]	3	35 839	1 033	0.00506 ± 0.00016	0.00518 ± 0.00010
		5	57 615	1 078	0.00433 ± 0.00046	0.00462 ± 0.00028
3	[16, 16, 16]	3	2 079	861	0.01187 ± 0.00053	0.01209 ± 0.00065
		5	3 135	912	0.00752 ± 0.00040	0.00762 ± 0.00040
4	[16, 16, 16, 16]	3	2 927	1 187	0.00631 ± 0.00039	0.00636 ± 0.00040
		5	4 495	1 038	0.00502 ± 0.00024	0.00535 ± 0.00033
5	[16, 16, 16, 16, 16]	3	3 775	1 053	0.00488 ± 0.00011	0.00525 ± 0.00011
		5	5 855	1 003	0.00481 ± 0.00027	0.00492 ± 0.00005
3	[32, 32, 32]	3	7 215	813	0.01165 ± 0.0001	0.01200 ± 0.00016
		5	11 375	942	0.00738 ± 0.00041	0.00771 ± 0.00049
4	[32, 32, 32, 32]	3	10 447	999	0.00615 ± 0.00032	0.00628 ± 0.00034
		5	16 655	987	0.00458 ± 0.00019	0.00483 ± 0.00029
5	[32, 32, 32, 32, 32]	3	13 679	998	0.00488 ± 0.00047	0.00508 ± 0.00041
		5	21 935	1 028	0.00457 ± 0.00030	0.00476 ± 0.00018

Table B.7
Results of the hyperparameter study for the LSTM architecture. For varying model depths, different neuron arrangements are investigated. The hourglass, inverted hourglass structure as well as a constant distribution of neurons are considered. The amount of model parameters are given as well as the mean number of epochs associated with the best model evaluated on the validation set. The performance of the models are measured by the SSP loss for the train and validation set in terms of mean and standard deviation of the four-fold cross-validation runs.

LSTM hyperparameters				SSP loss	
depth n_d	# neurons of each layer	# weights	# epochs	train	val
3	[8, 4, 8]	1 168	1 350	0.03189 ± 0.00201	0.03209 ± 0.00159
5	[16, 8, 4, 8, 16]	4 624	940	0.02162 ± 0.00022	0.02166 ± 0.00068
7	[32, 16, 8, 4, 8, 16, 32]	17 680	894	0.01918 ± 0.00010	0.02488 ± 0.00346
3	[4, 8, 4]	880	1 188	0.03250 ± 0.00154	0.03280 ± 0.00221
5	[4, 8, 16, 8, 4]	3 376	1 030	0.02333 ± 0.00023	0.02423 ± 0.00800
7	[4, 8, 16, 32, 16, 8, 4]	12 976	1 026	0.021720 ± 0.0003	0.02184 ± 0.00026
1	[5]	268	1 374	0.03525 ± 0.00302	0.03536 ± 0.00235
2	[5, 5]	508	1 295	0.03370 ± 0.00718	0.03452 ± 0.00864
3	[5, 5, 5]	748	1 003	0.02531 ± 0.01465	0.02546 ± 0.01477
4	[5, 5, 5, 5]	988	983	0.02337 ± 0.01363	0.02339 ± 0.01362
5	[5, 5, 5, 5, 5]	1 228	1 432	0.03506 ± 0.00382	0.03538 ± 0.00508
1	[20]	2 128	1 055	0.02657 ± 0.00137	0.02662 ± 0.00122
2	[20, 20]	5 488	1 030	0.02673 ± 0.00118	0.02732 ± 0.00174
3	[20, 20, 20]	8 848	1 091	0.02534 ± 0.00130	0.02567 ± 0.00154
4	[20, 20, 20, 20]	12 208	1 032	0.02551 ± 0.00106	0.02582 ± 0.00185
5	[20, 20, 20, 20, 20]	18 928	1 632	0.02540 ± 0.00320	0.02567 ± 0.00321

Table B.8
Results of the MLP applying different loss functions and activation functions. For the final MLP architecture resulting from the hyperparameter study, the Huber loss, mean absolute error (MAE) and mean of squares (MSE) are tested in addition to the SSP with the hyberbolic tangent as activation function. Besides the tangent, the sigmoid and rectified linear unit (ReLU) are tested. The results are given as loss values in terms of mean and standard deviation of the four-fold cross-validation.

	SSP loss		MAE loss		MSE loss $[10^{-5}]$	
Settings	train	val	train	val	train	val
Loss function						
SSP	0.00667 ± 0.0001	0.00689 ± 0.0004	0.00486 ± 0.0001	0.00496 ± 0.0003	6.6 ± 0.3	7.3 ± 0.8
Huber	0.00677 ± 0.0002	0.00783 ± 0.0003	0.00489 ± 0.0001	0.00581 ± 0.0002	6.3 ± 0.2	9.3 ± 0.9
MAE	0.00698 ± 0.0005	0.00702 ± 0.0004	0.00502 ± 0.0001	0.00506 ± 0.0003	6.5 ± 0.3	7.8 ± 0.9
MSE	0.00672 ± 0.0002	0.00782 ± 0.0004	0.00483 ± 0.0001	0.00578 ± 0.0002	6.1 ± 0.3	9.4 ± 0.6
Activation function						
tanh	0.00667 ± 0.0001	0.00689 ± 0.0004	0.00486 ± 0.0001	0.00496 ± 0.0003	6.6 ± 0.3	7.3 ± 0.8
sigmoid	0.01342 ± 0.0007	0.01394 ± 0.0006	0.00964 ± 0.0004	0.01005 ± 0.0003	17 ± 1.5	19 ± 1.3
ReLU	0.00712 ± 0.0005	0.00802 ± 0.0005	0.00512 ± 0.0004	0.00579 ± 0.0003	5.9 ± 6.6	9.5 ± 1.4

Table C.9
Results of the learning curves. All three models were trained for different training set sizes. The test set remains the same for all runs carried out. The number of epochs represent the mean amount of epochs associated with the best model on the test set. The loss values are given in terms of mean and standard deviation of the four repetitions carried out. The setting belonging to the best model performing on the test set is highlighted.

Data set sizes		MLP			FCNN		
train	test	# epochs	train loss	test loss	# epochs	train loss	test loss
250	1 000	1 174	0.00557 ± 0.00026	0.03074 ± 0.00528	1 097	0.00475 ± 0.00013	0.00543 ± 0.00027
500	1 000	1 137	0.00627 ± 0.00020	0.01601 ± 0.00216	1 012	0.00495 ± 0.00022	0.00532 ± 0.00025
1 000	1 000	1 120	0.00651 ± 0.00009	0.00883 ± 0.00050	1 034	0.00453 ± 0.00015	0.00473 ± 0.00029
2 000	1 000	1 129	0.00665 ± 0.00076	0.00714 ± 0.00036	1 024	0.00422 ± 0.00020	0.00475 ± 0.00016
4 000	1 000	1 038	0.00666 ± 0.00006	0.00686 ± 0.00005	945	0.00422 ± 0.00028	0.00464 ± 0.00022
		LSTM					
250	1 000	983	$0.02128 \pm 9e-05$	$0.0269 \pm 0.2e-05$			
500	1 000	959	0.02338 ± 0.00072	0.02533 ± 0.00106			
1 000	1 000	1 111	0.02112 ± 0.00102	0.02324 ± 0.00102			
2 000	1 000	1 044	0.02092 ± 0.00064	0.02272 ± 0.00076			
4 000	1 000	988	0.02111 ± 0.00063	0.02252 ± 0.00101			

References

Abadi, M., et al., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. URL https://www.tensorflow.org/.

Adegeest, L., 1996. Nonlinear hull girder loads in ships. 156, https://repository.tudelft.nl/record/uuid:43fc0165-ff14-48a6-b444-feb699df1e2a.

Chen, C.-Z., Liu, S.-Y., Zou, Z.-J., Zou, L., Liu, J.-Z., 2023a. Time series prediction of ship maneuvering motion based on dynamic mode decomposition. Ocean Eng. 286, 115446. http://dx.doi.org/10.1016/j.oceaneng.2023.115446.

Chen, C.-Z., Zou, Z.-J., Zou, L., Zou, M., Kou, J.-Q., 2023b. Time series prediction of ship course keeping in waves using higher order dynamic mode decomposition. Phys. Fluids 35 (9), http://dx.doi.org/10.1063/5.0165665.

Chollet, F., et al., 2015. Keras. https://keras.io.

Clauss, G., Klein, M., Guedes Soares, C., Fonseca, N., 2012. Response based identification of critical wave scenarios. J. Offshore Mech. Arct. Eng. 135, http://dx.doi.org/10.1115/OMAE2012-83861.

Dannenberg, J., Hessner, K., Naaijen, P., van den Boom, H., Reichert, K., 2010. The on board wave and motion estimator OWME. In: Proceedings of the International Offshore and Polar Engineering Conference. Vol. 3, pp. 424–431.

del Águila Ferrandis, J., Triantafyllou, M., Chryssostomidis, C., Karniadakis, G., 2021. Learning functionals via LSTM neural networks for predicting vessel dynamics in extreme sea states. Proc. R. Soc. A http://dx.doi.org/10.1098/rspa.2019.0897.

Fonseca, N., Guedes Soares, C., 1970. Time domain analysis of vertical ship motions. WIT Trans. Built Environ. 5.

Fonseca, N., Guedes Soares, C., 1998. Time-domain analysis of large-amplitude vertical ship motions and wave loads. J. Ship Res. 42, 139–153. http://dx.doi.org/10.5957/jsr.1998.42.2.139.

Goodfellow, I.J., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, Cambridge, MA, USA, http://www.deeplearningbook.org.

Guo, X., Zhang, X., Tian, X., Li, X., Lu, W., 2021. Predicting heave and surge motions of a semi-submersible with neural networks. Appl. Ocean Res. 112, 102708. http://dx.doi.org/10.1016/j.apor.2021.102708.

Jiang, Z., Ma, Y., Li, W., 2024. A data-driven method for ship motion forecast. J. Mar. Sci. Eng. 12, 291. http://dx.doi.org/10.3390/jmse12020291.

Khan, A., Marion, K., Bil, C., 2007. The prediction of ship motions and attitudes using artificial neural networks. Asor Bull. 26 (1), 2.

Klein, M., Dudek, M., Clauss, G.F., Ehlers, S., Behrendt, J., Hoffmann, N., Onorato, M., 2020. On the deterministic prediction of water waves. 5, http://dx.doi.org/10.15480/882.2565.

Kusters, J., Cockrell, K., Connell, B., Rudzinsky, J., Vinciullo, V., 2016. FutureWaves™: A real-time ship motion forecasting system employing advanced wave-sensing radar. pp. 1–9. http://dx.doi.org/10.1109/OCEANS.2016.7761478.

Law, Y., Santo, H., Lim, K., Chan, E., 2020. Deterministic wave prediction for unidirectional sea-states in real-time using artificial neural network. Ocean Eng. 195, 106722. http://dx.doi.org/10.1016/j.oceaneng.2019.106722.

Lee, J.-H., Lee, J., Kim, Y., Ahn, Y., 2023. Prediction of wave-induced ship motions based on integrated neural network system and spatiotemporal wave-field data. Phys. Fluids 35, http://dx.doi.org/10.1063/5.0163795.

Lee, J.-H., Nam, Y.-S., Kim, Y., Liu, Y., Lee, J., Yang, H., 2022. Real-time digital twin for ship operation in waves. Ocean Eng. 266, 112867. http://dx.doi.org/10.1016/j.oceaneng.2022.112867.

Morris, E., Zienkiewicz, H., Belmont, M., 1998. Short term forecasting of the sea surface shape. 45, pp. 383–400.

Naaijen, P., Dijk, R., Huijsmans, R., El-Mouhandiz, A., 2009. Real time estimation of ship motions in short crested seas. In: Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE. Vol. 4, http://dx.doi.org/10.1115/OMAE2009-79366.

Naaijen, P., Huijsmans, R., 2008. Real time wave forecasting for real time ship motion predictions. In: Proc. OMAE 2008. Vol. 4, http://dx.doi.org/10.1115/OMAE2008-57804.

Naaijen, P., Huijsmans, R., 2010. Real time prediction of second order wave drift forces for wave force feed forward in DP. In: Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE. http://dx.doi.org/10.1115/OMAE2010-20618.

Naaijen, P., Trulsen, K., Blondel-Couprie, E., 2014. Limits to the extent of the spatio-temporal domain for deterministic wave prediction. Int. Shipbuilding Progress 61 135, 203–223. http://dx.doi.org/10.3233/ISP-140113.

Naaijen, P., van Oosten, K., Roozen, K., van't Veer, R., 2018. Validation of a deterministic wave and ship motion prediction system. In: International Conference on Offshore Mechanics and Arctic Engineering, Volume 7B: Ocean Engineering, http://dx.doi.org/10.1115/OMAE2018-78037, V07BT06A032.

NVIDIA, Vingelmann, P., Fitzek, F., 2020. CUDA, release: 10.2.89. URL https://developer.nvidia.com/cuda-toolkit.

Ochi, M.K., 1990. Applied Probability and Stochastic Processes. Wiley New York.

Salvesen, N., Tuck, E.O., Faltinsen, O.M., 1970. Ship motions and sea loads. pp. 250–287, URL https://api.semanticscholar.org/CorpusID:109444905.

Sclavounos, P.D., Ma, Y., 2018. Artificial intelligence machine learning in marine hydrodynamics. In: International Conference on Offshore Mechanics and Arctic Engineering, Volume 9: Offshore Geotechnics, http://dx.doi.org/10.1115/OMAE2018-77599, V009T13A028.

Söding, H., 2001. Global Seaway Statistics. TU Hamburg-Harburg Schriftenreihe Schiffbau, Bericht Nr. 610, April 2001, Germany.

Sun, Q., Tang, Z., Gao, J., Zhang, G., 2022. Short-term ship motion attitude prediction based on LSTM and GPR. Appl. Ocean Res. 118, 102927. http://dx.doi.org/10.1016/j.apor.2021.102927.

Triantafyllou, M., Athans, M., 1981. Real time estimation of the heaving and pitching motions of a ship using a Kalman filter. In: OCEANS 81. http://dx.doi.org/10.1109/OCEANS.1981.1151582.

Triantafyllou, M., Bodson, M., Athans, M., 1983. Real time estimation of ship motions using Kalman filtering techniques. Oceanic Eng., IEEE J. 8, 9–20. http://dx.doi.org/10.1109/JOE.1983.1145542.

Wang, S., Klein, M., Ehlers, S., Clauss, G., Guedes Soares, C., 2024. Analysis of the behavior of a chemical tanker in extreme waves. J. Mar. Sci. Appl. http://dx.doi.org/10.1007/s11804-024-00508-0.

Wedler, M., Stender, M., Klein, M., Ehlers, S., Hoffmann, N., 2022. Surface similarity parameter: A new machine learning loss metric for oscillatory spatio-temporal data. Neural Netw. 156, 123–134. http://dx.doi.org/10.1016/j.neunet.2022.09.023.

Wu, G., 2004. Direct simulation and deterministic prediction of large-scale nonlinear ocean wave-field. Department of Ocean Engineering.

Xue, J., Gao, Z., Yi, W., 2024. An APSO-based TPA-BLSTM model for predicting ship motion in irregular waves using wave-series input. Ships Offshore Struct. 1–12. http://dx.doi.org/10.1080/17445302.2024.2336675.

Yang, Y., Peng, T., Liao, S., 2023a. Predicting 3-DoF motions of a moored barge by machine learning. J. Ocean Eng. Sci. 8 (4), 336–343. http://dx.doi.org/10.1016/j.joes.2022.08.001, Ocean Engineering for Future Marine Structures and Applications.

Yang, Y., Peng, T., Liao, S., 2023b. Predicting future mooring line tension of floating structure by machine learning. Ocean Eng. 269, 113470. http://dx.doi.org/10.1016/j.oceaneng.2022.113470, URL https://www.sciencedirect.com/science/article/pii/S0029801822027536.

Yang, Y., Peng, T., Liao, S., Li, J., 2024. Real-time prediction of ship motions based on the reservoir computing model. J. Ocean Eng. Sci. http://dx.doi.org/10.1016/j.joes.2024.07.001, URL https://www.sciencedirect.com/science/article/pii/S2468013324000391.

Yin, J.-C., Perakis, A.N., Wang, N., 2018. A real-time ship roll motion prediction using wavelet transform and variable RBF network. Ocean Eng. 160, 10–19.

Yumori, I., 1981. Real time prediction of ship response to ocean waves using time series analysis. In: OCEANS 81. pp. 1082–1089. http://dx.doi.org/10.1109/OCEANS.1981.1151574.

Zhang, T., Zheng, X.-Q., Liu, M.-X., 2021. Multiscale attention-based LSTM for ship motion prediction. Ocean Eng. 230, 109066. http://dx.doi.org/10.1016/j.oceaneng.2021.109066.