Prediction of Wave Energy Spectrum Based on Ship Motions Using a Data-Driven Approach

Alessandro La Ferlita [1], Yan Qi [1], Emanuel Di Nardo [2], Simon Mewes [1], Ould El Moctar [1] Angelo Ciaramella [2]

[1] University of Duisburg-Essen, 47057 Duisburg, Germany, Institute of Ship Technology, Ocean Engineering and Transport Systems, Department of Mechanical and Process Engineering.
[2] University of Naples Parthenope, 80133 Naples, Italy, Department of Science and Technology

Abstract

Wave energy spectra are used to create sea states and to obtain ship motion transfer functions for different frequencies. These transfer functions are non-linear. Hence, the precise estimation is not straightforward. In this study, the spectral parameters, significant wave height and peak period, are obtained via a deep neural network (DNN) approach using the ship motions as input variables. The main advantage of such a method lies in its possibility to predict the spectral parameters without the use of ship specific properties.

Keywords

Sea state, DNN, Ship motions, Data-driven approach

1. Introduction

Accurate prediction of the sea spectrum is an important task for marine and engineering applications since the natural environment can expose the vessels to potential risks.
During the lifetime of the ship, waves and winds induce loads and related stresses to the hull structure, sometimes imperil the safety of the vessel.
Instantaneous loads, such as accelerating forces, slamming and sloshing loads, are those effects that the waves and the resulting ship motions impose on the ship's hull structure.
Therefore, the dynamic loads acting on the ship contribute to different effects such as fatigue, structural failure, corrosion, or crack propagation.
Thus, the precise estimation of significant wave height and of sea state parameters plays a relevant role since vessels may encounter adverse conditions during their route and experiencing an added resistance due to waves. This leads to an increase in the total resistance, and consequently, the ship fuel consumption.
To predict ship fuel consumption as well as fatigue and lifetime, a precise forecast of the environmental condition is imperative. Data of weather forecast and hindcast are expensive and often not complete. Therefore, alternatives are needed to fill the data gaps as well as to improve weather data.
The sea spectrum may be determined considering the so-called "wave buoy" analogy [1] approach often deducted in frequency-domain approach.
Thus, the ship is viewed as a buoy, therefore an inverse mathematical link between measured responses and the encountered directional wave spectrum is defined, i.e. the measured ship responses are used as input to estimate wave spectrum and associated sea state parameters. As a matter of course, the determination of the transfer functions is a fundamental requirement. The most relevant numerical methods based on potential theory to determine the transfer function may be classed [2] as follows:

ITADATA2023: The 2nd Italian Conference on Big Data and Data Science, September 11-13, 2023, Naples, Italy
alessandro.laferlita@uni-due.de (A. La Ferlita) yan.qi@uni-due.de (Y. Qi) emanuel.dinardo@uniparthenope.it (E. Di Nardo) simon.mewes@uni-due.de (S. Mewes) angelo.ciaramella@uniparthenope.it (A. Ciaramella) ould.el-moctar@uni-due.de (O. El Moctar)
0009-0007-1680-1741 (A. La Ferlita) yan.qi@uni-due.de (Y. Qi) 0000-0002-6589-9323 (E. Di Nardo) 0000-0002-0419-3423 (S. Mewes) 0000-0001-5592-7995 (A. Ciaramella) 0000-0002-5096-4436 (O. El Moctar)

1. Strip method.
2. Unified theory.
3. High speed Theory
4. Green function method
5. Rankine source method

Usually, these approaches are linear and cannot capture further nonlinear aspects such as turbulence, wave dispersion, or water compressibility. In fact, often such kind of codes assume linearized small unit wave amplitudes hypothesis. In these approaches, the fluid is assumed to be inviscid, incompressible and irrotational without surface tension, such that the spatial flow velocity vector can be expressed as the gradient of a scalar 3D velocity [3].
Nevertheless, considering the improvement of computational power and parallel computing, the Reynolds-Averaged Navier–Stokes (RANS) approach is more frequently applied to solve also unsteady seakeeping problems [4].
However, the same task of obtaining the sea spectrum may also be realized using supervised machine learning. The fundamental idea is to learn the mapping from measured ship motion responses to the actual sea state from historical data. The cost in terms of computational time for the transfer function obtained with CFD (Computation Fluid Dynamics) is typically higher if compared to machine learning approaches. In addition, some problems regarding accuracy of free-surface due to the limited grid quality can be encountered [4].
The main advantage of data-driven methods is that it is not required to define specific vessel properties, such as radius inertia or hydrostatics of the vessels to discover the pattern between ship motions and sea states [5].
Regarding this aspect the literature does not offer many examples of such kind of approaches. For instance, Nielsen et al. [6] proposed a hybrid approach for wave spectrum estimation. Mittendorf et al. [7] proposed a prediction of sea states based on ship motions using a data-driven approach, considering the in-service data of a container vessel.
The application of machine learning and especially of DNN (Deep Neural Network) is used and spread in several naval architecture disciplines, e.g., in structural field ([8] [9]) or ship performances or engine break power and ship fuel consumption predictions [10], [11], [12].
The applicability and high performance of such methods lead to a high interest of many researchers when the nature of the problem is complex and when many nonlinear effects must be considered. Therefore, the application in the subject topic can lead to many advantages. Hence, the main objective of the study is to show the accuracy of a machine learning approach to predict the significant wave height and peak period.

2. Methodology

A simplified model for approaching the problem, which consists mainly of demonstrating the accuracy of a machine learning approach by using ship motions to predict significant wave height and peak period, was established. This consists of the following steps: at first a database with the measured data is generated, part of the data (90%) is used to train the DNN, and the remaining set of the data is used to perform the validation. Afterwards, unknown samples to the algorithm were used to test the model. Finally, the spectrums are determined and compared against the observed ones. The methodology is graphically presented in Figure 1.
A database of public domain data has been used [13]. The movement of 46 moored vessels for a total of 1609 hours in the duration from October 2015 to February 2020 were taken into consideration during five field campaigns. These data are used to train and test the model to predict the significant wave height and the peak period. The data is originally published by Alvarellos et al. [13]. The objective of their study was to predict the ship's movements in advance using an ANN (Artificial Neural Network). Here, the input data were the weather conditions, the ship characteristics and berthing location. Thus, more detailed input data than in the present study. This led to a high level of accuracy but requires more knowledge about the ship itself.

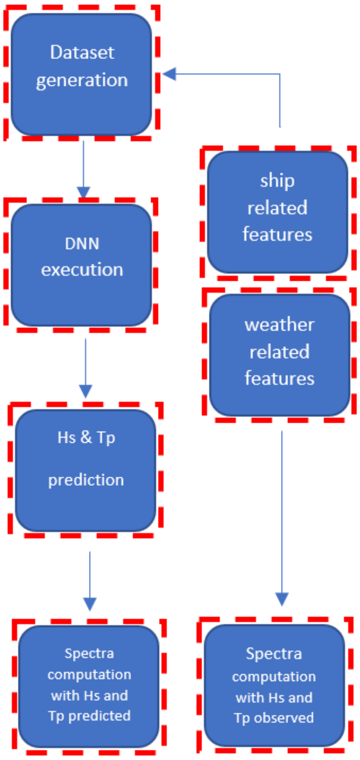

Figure 1: Schematic representation of the approach

A database of public domain data has been used [13]. The movement of 46 moored vessels for a total of 1609 hours in the duration from October 2015 to February 2020 were taken into consideration during five field campaigns measurements. These data are used to train and test the model to predict the significant wave height and the peak period.
The data is originally published by Alvarellos et al. [13]. The objective of their study was to predict the ship's movements in advance using an ANN (Artificial Neural Network). Thus, more detailed input data than in the present study. This led to a high level of accuracy but requires more knowledge about the ship itself. The input variables available for performing the training and validation tasks have been grouped into weather related features, such as:

- H_s [m]: significant wave height
- T_p [s]: peak wave period,
- θ_m [deg]: mean wave direction
- W_s [km/h]: mean wind speed.
- W_d [deg]: mean wind direction.
- H_0 [m]: sea level with respect to the zero of the port.
- H_{sm} [m]: significant wave height measured by the tide gauge.

And ship related features, such as:

- Surge [m]: linear longitudinal motion (bow-stern).
- Sway [m]: linear lateral motion (port-starboard).
- Heave [m]: linear vertical motion.
- Roll [deg]: tilting rotation of the vessel about its longitudinal axis (port/starboard)
- Pitch [deg]: up/down rotation of the vessel about its transverse axis
- Yaw [deg]: turning rotation of the vessel about its vertical axis.
- L [m]: ship length.
- B [m]: ship breadth.
- DWT [ton]: deadweight

Table 1
Sample's number used to train the DNN (Deep Neural Network).

Ship Motion	Roll	Pitch	Heave	Surge	Yaw	Sway
Number of Samples	1349	1349	365	365	1249	1452

A total number of 6129 samples (considering the six degree of freedom of the vessels) has been used for training the algorithm (see Table 1). The only two vessel topology available in the data set were bulk carrier and general cargo ship.
The DNN method was chosen since it quickly provided outputs that are less prone to overfitting and the computational durations are shorter [10], further the generalization ability of deep neural networks helps to obtain very satisfying results. Only two hidden layers are present in the network since the amount of data available is limited and the course of dimensionality aspect has to be taken into account. [14]. TensorFlow [8], a deep learning computational toolkit, and Google Colab [15] were used to create the DNN model, which was then trained on a nVidia Tesla K80 GPU (Graphic Processor Unit). In order to do this, we used the ADAM optimizer [16] with an initial learning rate of 0.001 and an exponential decay of 0.96 over a period of 10 epochs. As loss function the SmoothL1 is used to act as a L1 and L2 based on a threshold parameter, but preferring to act for the most as a L1 function.
Various idealized energy spectra exist to represent the sea state. The ITTC (International Towing Tank Conference) recommends the use of the JONSWAP (Joint North Sea Wave Observation Project) spectrum, which is based on the observations obtained in the North Sea westward from the Sylt Island (Westerland, Germany). The observation lasted for a period of 10 weeks during the year 1968–1969 [17]. The sea spectrum reads:

$$S_J(\omega_e) = \frac{C_1}{C_2} exp\left[-\frac{5}{4}\left(\frac{\omega_p}{\omega}\right)^4\right]\gamma^{exp\left[-\frac{5}{4}\left(\frac{(\omega-\omega_p)^2}{2\sigma^2\omega_p^2}\right)\right]}, \tag{1}$$

$$C_1 = \frac{5{H_S}^2}{16\omega_p} \tag{2}$$

$$C_2 = 1.15 + 0.1688\gamma - \frac{0.925}{1.909 + \gamma} \tag{3}$$

Where the ω_p denotes the peak frequency, ω the wave frequency, Hs the significant wave height, and γ the peak enhancement factor. σ is defined as follows:

$$\sigma = \begin{cases} 0.07, & \omega \le \omega_p \\ 0.09, & \omega > \omega_p \end{cases} \tag{4}$$

These values define the left and right sided widths of the spectral peak, respectively. The advantage of the JONSWAP model, compared to other model (for instance Pierson-Moskovitz model), is that it can consider effects of limited wind fetch length and water depth [18].
A theoretical estimate of the ship response spectrum can be obtained through a merge of the wave spectrum and the transfer function of the given motion response [6]. The formula below provides a mathematical representation of the problem:

$$S_{IJ}(\omega_e) = \int_{-\pi}^{\pi} \Phi_i(\omega_e,\beta)\,\overline{\Phi_j(\omega_e,\beta)}E(\omega_e,\beta)d\beta \tag{5}$$

Where:

- $S_{IJ}(\omega_e)$ is the response spectrum function of the frequency ω_e;
- $\int_{-\pi}^{\pi} \Phi_i(\omega_e, \beta)\, \overline{\Phi_J(\omega_e, \beta)}$ is the transfer function or RAO (Response Amplitude Operator), function of the frequency and of the spreading β, and
- $E(\omega_e, \beta)$ is the sea spectrum

3. Results

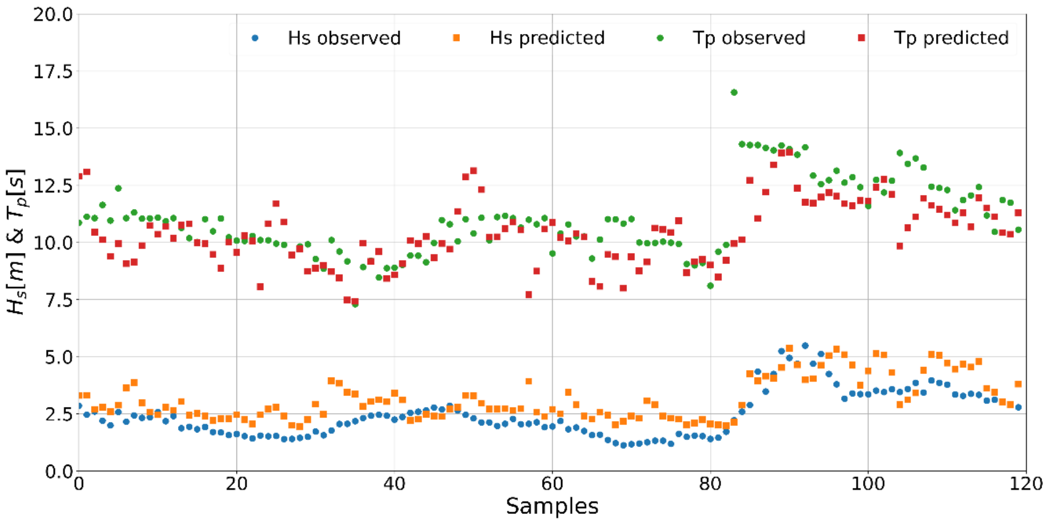

Figure 2: Comparison between Hs (significant wave height) and Tp (peak period) observed and predicted.

Figure 2 presents the comparison between predicted values of the significant wave heights and the peak period obtained from the data driven model (marked in orange and the red color) and from the measurements (marked in blue and green color).
The horizontal axis represents the number of samples. The vertical axis shows Hs and Tp, respectively expressed in meters and seconds. Exemplarily, 120 samples measurement for the sway motion. are considered as test cases (the so-called "wild data").
For the first 40 samples related to the Hs prediction, the DNN approach tends to slightly overestimate the observed values. From sample 40 to 60 the accuracy improves. Starting from sample number 60, the predictions derived from the DNN approach overestimate the measured values as observed in the first 40 samples.
The comparison for the prediction of Tp presents some peaks. As seen at sample numbers 83, the data driven model significantly underestimated the value compared to the corresponding measured value. However, it is observable that the trend of the prediction qualitatively follows the observed one.
The two standard loss functions which are often used to indicate in machine learning the prediction accuracy are the root mean squared error (RMSE) and mean absolute error (MAE) which yield respectively: 0.66 and 0.54.
After the Tp and Hs computation, the spectrums were obtained. Exemplarily, six different scenarios (a to f) to the unknown tested Hs and Tp cases are presented in Figure 3. The vertical axis represents the wave spectra density in $[\frac{\text{rad}}{\text{s}}]m^2$, while the horizontal axis shows the wave frequency in [rad/s].
The spectra were calculated for the measured and for the predicted through the formula previously shown in equation 1, assuming a fully developed sea.
A fixed peak enhancement factor of 3.3 has been chosen. The scenarios presented in Figure 3 reveal an overall good agreement with the processed spectrum obtained using the observed data (Hs, Tp).

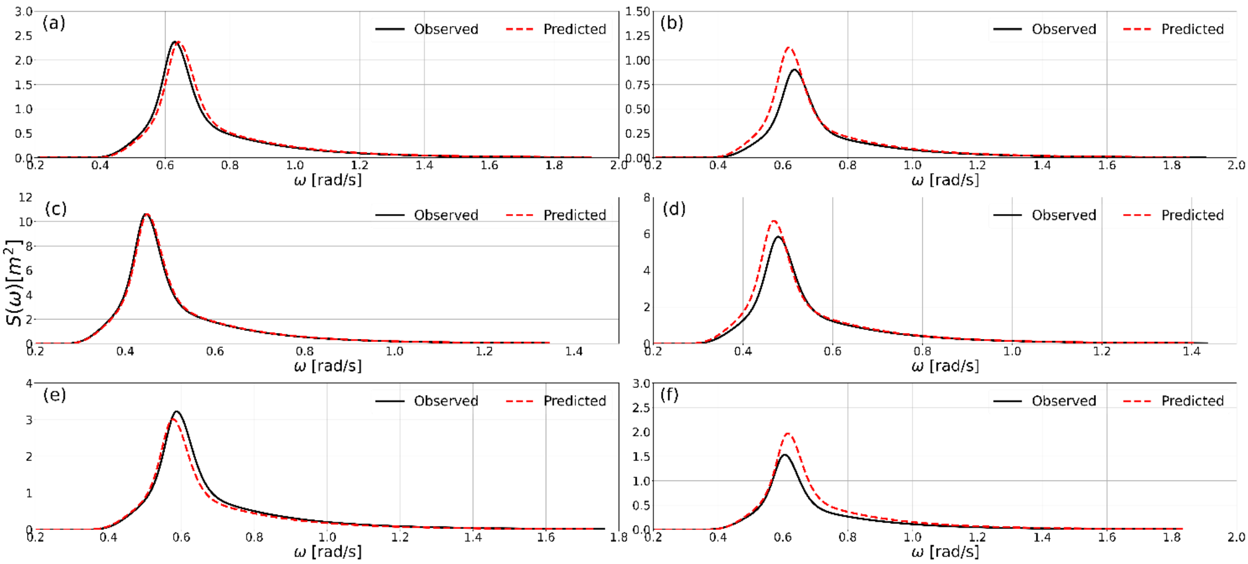

Figure 3: Observed and predicted spectra.

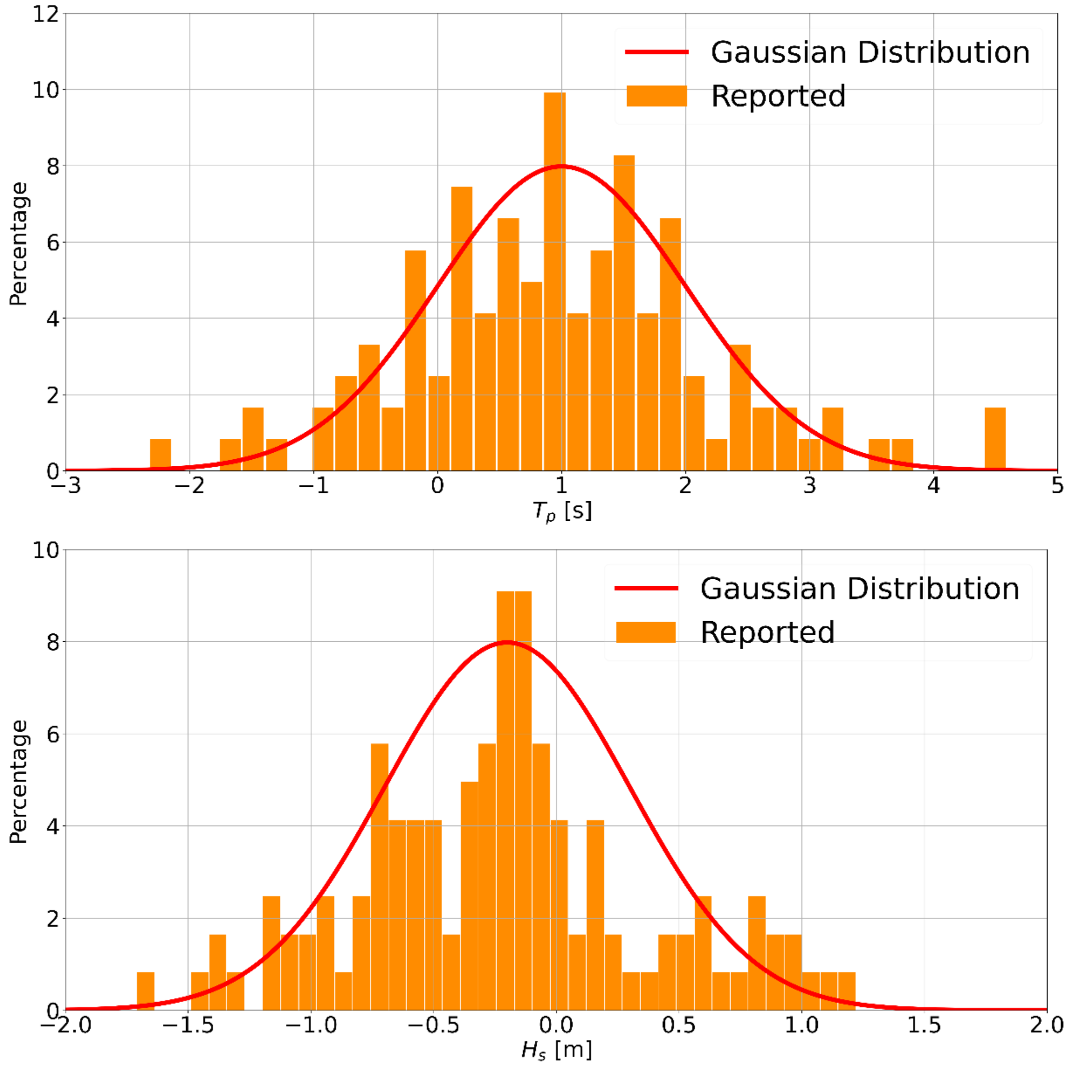

Figure 4: Difference distribution of Tp and Hs.

However, it must be said that scenario presented b, d and f show less agreement if compared to the actual spectrum. This can be observed even from the Hs and Tp obtained (Table 2). The reason can be related to the lack of data. The results obtained for the other scenarios (a, c, e) in term of Hs and Tp values are given as well. For these cases, the prediction is better, and this is reflected in the spectrum.
A statistical representation of the deviation between the predicted and observed Tp and Hs values is shown in figure 4.
With a limited sample size, the distribution can be approximated with the normal distribution. For the gaussian Hs curve, a mean value of -0.2 and a standard deviation of 0.5 has been set. For the gaussian Tp curve, a mean value of 1 and a standard deviation of 1 has been considered. In both cases, the spread around the central tendency is almost symmetric.
Almost 10 % of the Tp test cases present a difference of 1 second, on the other side the 9.3% of the Hs samples yield a difference of -0.1 m. For the Tp values only the 0.826% of samples has a deviation less than -2.732s. For the same percentage value, the Hs test cases present a deviation of -2.173. Approximately 9.917% of the Tp test cases predicted yield a difference of 0.445s with respect to the observed ones. The 9.091% of the Hs wild data presents a difference of -0.634 m if compared to the respective Hs values measured.
As for the great majority of machine learning algorithms, they cannot often quantify the error/uncertainty associated with their predictions or granting data convergence as shown above.
In this study the error data has the statistical distribution as shown in figure 4, but the dataset is built starting from specific ship features that can vary in different context, as it can happen with weather features.
A mix of these two features in uncovered situations can lead to inaccurate results. This situation is one of the possible problems that can happen with artificial neural networks due to a low robustness of the model based on the model architecture itself and on type of data. This condition can be reduced when a huge amount of data is available with coverage of multiple scenarios.
In the study case, the availability of approximately 6000 samples for training a simple DNN algorithm had probably led to perform sufficient accuracy prediction.
Therefore, in such scenarios with reduced amount of data, the capacity for robust and sample-efficient learning might be essential.

Table 2
Compared Hs and Tp

Scenario	Hs observed [m]	Hs predicted [m]	Tp observed [s]	Tp predicted [s]
a	2.78	2.80	9.98	9.80
b	1.72	1.90	9.9	10.16
c	4.95	4.96	14.09	13.99
d	3.8	4.03	13.14	13.41
e	3.13	2.99	10.71	10.9
f	2.19	2.50	10.39	10.21

4. Conclusions and Future Work

A data driven model was presented for the determination of the wave energy spectrum. The training of the model was performed with ship motion measurement data. This approach allows to circumvent the need for transfer functions and ship characteristic information. The results presented show sufficient agreement with the effective data measured despite the limited number of sample available for training the neural network.

The increase of data samples to train and validate the network must be considered for further development. Furthermore, parameters related to the spectrum, such as fetch length, might improve the accuracy of the prediction.

5. References

[1] U. Nielsen, "Sea state estimation based on measurements of wave-induced ship responses," orbit.dtu.dk, Doctoral thesis, 2018.

[2] V. Bertram, Practical Ship Hydrodynamics, Elsevier Ltd, 2012.

[3] A. La Ferlita, H. Rathje, T. Lindemann, R. Bronsart and P. Kaeding, "An Advanced Salvage Method for Damaged Ship Structures," in *SNAME Maritime Convention*, Providence RI (USA), 27-29 October, 2021.

[4] F. Pacuraru, L. Domnisoru and S. Pacuraru, "On the Comparative Seakeeping Analysis of the Full Scale KCS by Several Hydrodynamic Approaches," *Journal of Marine Science and Engineering,* vol. 8, p. 962, 2020.

[5] P. Han, G. Li, S. Skjong and H. Zhang, "Directional wave spectrum estimation with ship motion responses using adversarial networks," *Marine Structure,* vol. 83, 2022.

[6] U. D. Nielsen, R. E. Mounet and A. H. Brodtkorb, "Tuning of transfer functions for analysis of wave–ship interactions," *Marine Structures,* vol. 79, p. 103029, 2021.

[7] M. Mittendorf, U. Nielsen, H. B. Bingham and G. Storhaug, "Sea state identification using machine learning—A comparative study based on in-service data from a container vessel," *Marine Structures,* 2022.

[8] A. La Ferlita, E. Di Nardo , M. Macera, T. Lindemann, A. Ciaramella and N. Koulianos, "A Deep Neural Network to Predict the Residual Hull Girder Strenght," in *SNAME 2022*, Houston, Texas, 2022.

[9] A. La Ferlita, E. Di Nardo, M. Macera, T. Lindemann, A. Ciaramella and P. Kaeding, "Deep Neural Network (DNN) Method to predict the displacement behavior of neutral axis for ships in vertical bending," in *Proceeding of 20th International Conference on Ship and Maritime Research (NAV2022)*, Genoa, Italy, 15–17 June 2022..

[10] A. La Ferlita, Y. Qi, E. Di Nardo, T. Schellin, O. el Moctar and A. Ciaramella, "A Comparative Study to Estimate Fuel Consumption: A Simplified Physical Approach against a Data-Driven Model," *Journal of Marine Science and Engineering,* vol. 11, 2023.

[11] T. Uyanık, Y. Yalman, O. Kalenderli, Y. Arslanoğlu, Y. Terriche, C.-L. Su and J. M. Guerrero, "Data-Driven Approach for Estimating Power and Fuel Consumption of Ship: A Case of Container Vessel," *Mathematics,* vol. 10, 2022.

[12] W. Jabary, C. Liu, F. Sprenger, L. Kleinsorge, H. Baumfalk, M. Kaster, J. Neugebauer and O. el Moctar, "Development of machine learning approaches to enhance ship operational performance evaluation based on an integrated data model," in *33rd International Ocean and Polar Engineering Conference (ISOPE2023)*, Ottawa, Canada, 2023.

[13] A. Alvarellos, A. Figuero, H. Carro, R. Costas, J. Sande, A. Guerra, E. Peña and J. Rabuñal, "Machine Learning Based Moored Ship Movement Prediction," *Journal of Marine Science and Engineering,* vol. 9, 2021.

[14] M. Verleysen and F. Damien, "The curse of dimensionality in data mining and time series prediction," in *International work-conference on artificial neural networks.*, Berlin, 2005.

[15] E. Bisong, Building Machine Learning and Deep Learning Models on Google Cloud Platform, Apress, 2019.

[16] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980,* 2014.

[17] K. Hasselmann, T. Barnett, E. Bouws., H. Carlson, D. Cartwright, K. Enke, J. Ewing, H. Gienapp, D. Hasselmann and P. Kruseman, "Measurements of Wind-Wave Growth and Swell Decay during the Joint North Sea Wave Project (JONSWAP);," 1973.

[18] Y. Liu, S. Li, Q. Yi and D. Chen, "Wind Profiles and Wave Spectra for Potential Wind Farms in South China Sea. Part II: Wave Spectrum Model," *Energies,* 2017.

[19] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," *Neural networks: Tricks of the trade. Springer, Berlin, Heidelberg,* pp. 437-478, 2012.