Proceedings of the ASME 2025 44th International Conference on Ocean, Offshore and Arctic Engineering
OMAE2025
June 22-27, 2025, Vancouver, British Columbia, Canada

OMAE2025-157346

COMPARATIVE ANALYSIS OF MACHINE LEARNING METHODS FOR PREDICTING STRUCTURAL RESPONSES IN SHIP HULL MONITORING

Simon Haberl[1,*], Moritz Braun[1], Sören Ehlers[1],

[1]German Aerospace Center, Geesthacht, Schleswig-Holstein ,Germany

ABSTRACT

Monitoring of ship hulls and their technical equipment is a critical component of ensuring ship safety and operational efficiency. It plays a key role in maintaining structural integrity under dynamic and often extreme loading conditions encountered at sea. Accurate prediction of structural responses not only helps in preventing structural failures but also enables the optimization of performance, reducing operational risks and costs. The structural responses of a ship subjected to wave-induced loads can be modeled through a combination of hydrodynamic simulations and Finite Element analyses. While these simulations offer deep insights into the behavior of ship structures under varying conditions, their computational intensity and complexity present significant challenges for real-time applications. To address these limitations, this research evaluates the efficiency and accuracy of machine learning methods, specifically Artificial Neural Networks and XGBoost, in approximating and predicting structural responses.

Keywords: machine learning, hull monitoring

1. INTRODUCTION

Maritime transportation underpins the global economy, with ships forming the essential machinery of international trade. These vessels are intricate systems, comprising interconnected mechanical and structural components designed for seamless operation. As the industry evolves, the development of new ship designs increasingly prioritizes cost efficiency and environmental sustainability. This often involves reducing structural weight while maintaining safety standards, striking a balance between minimizing material use and preserving the ship's integrity. Achieving this balance relies on detailed insights into the ship's structural health, which can be obtained through effective monitoring systems.

A ship's operational condition plays a critical role throughout its lifecycle, not only in ensuring safety but also in driving economic efficiency. Understanding a vessel's fatigue behavior allows for precise fleet management and targeted maintenance strategies. Current regulatory frameworks, such as those upheld by the International Association of Classification Societies (IACS), mandate rigorous inspections at regular intervals. Insights derived from monitoring the condition of ships have the potential to significantly reduce inspection efforts, enabling maintenance schedules tailored to the vessel's specific needs. Such condition-based approaches lower costs, optimize operational uptime, and enhance the overall reliability of maritime assets.

In addition to these considerations, foresight in ship operations is emerging as an essential dimension of modern maritime practices. This forward-looking perspective goes beyond existing approaches by emphasizing real-time prediction of structural responses. By integrating advanced wave radar technologies, machine learning algorithms, and predictive modeling, operators can anticipate how a ship's structure will perform under varying conditions. Such insights enable immediate adjustments to operations, mitigating risks before they escalate and ensuring that ships operate within safe and efficient parameters. Foresight promotes proactive decision-making, transforming how fleets are managed and maintained. By predicting structural behavior in real-time and aligning it with long-term operational strategies, ship operators can enhance sustainability, reduce costs, and improve safety.

1.1 State of the art

A holistic structural health monitoring, involving the continuous surveillance of every crucial detail of a ship, has not been fully achieved to date due to the inherent complexities. The literature offers a broad spectrum of approaches, with each contributing partially to the resolution of the comprehensive problem. Each of these applications is distinguished by its respective type of input data sources.

One approach involves the use of sensors explicitly measuring structural strain at predefined locations, which are then analytically processed to perform damage assessments. A practical application of strain measurement systems is described by

*Corresponding author
Documentation for `asmeconf.cls`: Version 1.38, March 24, 2025.

Drummen et al. [1]. Another method using strain sensor signals is Anyfantis et al. [2]. This method identifies damaged structural elements based on anomalies in strain sensor measurement signals. None of these approaches, however, addresses the dual challenges of foresight and real-time capability. The absence of foresight limits the ability to optimize operations proactively, and the lack of real-time processing capabilities further constrains the responsiveness of these systems.

Alternative methods exist that rely on implicit measurements. In this approach, the structural response is not explicitly measured in the form of strains but instead is determined implicitly. To achieve this, a readily measurable parameter, such as the movements and accelerations of the ship, is gauged on board. Statistical methods are then employed to infer the structural response based on these measurements. Lang et al. [3] propose a machine learning method to determine fatigue rates of a ship along its route. The foundation of the machine learning model is based on a three-year measurement campaign of a 2800 TEU container ship in the Atlantic, encompassing measurements of the ship's movements and corresponding strains. Neither the methods based on explicit measurements nor those based on implicit measurements directly consider the exciting wave environment of the ship, rendering the prediction of the structural response to future wave events impossible. Serr et al. [4] propose a method also based on machine learning to predict the wave-induced vertical bending moment in the time domain. The input of this method consists of time series of wave elevation. While this approach bridges the gap to the time domain and enables the prediction of future wave events, it predicts wave loads rather than the structural responses to these loads.

As these methods rely on historical datasets including strain measuremens, they do not address the issue of the necessity for expensive and complex measurement systems. To completely eliminate the need for these, hull monitoring must transition to a fully virtual approach, entirely devoid of sensors. Methods of this nature are referred to in the literature as virtual hull monitoring. Schirmann et al. [5] propose a hull monitoring approach for assessing fatigue damage on various routes. This involves integrating AIS (Automatic Identification System) position data with weather reports to capture the wave environment of the ship. A component stochastic analysis in the frequency domain is then employed to assess fatigue damage using a beam formulation of a simplified ship cross-section to determine the stress transfer functions. Thompson et al.[6] propose a similar method based on numerical simulations. The innovation of this work involves the utilization of a full-stochastic approach in the frequency domain and the incorporation of a detailed global Finite Element model for stress determination. The stress spectra obtained with the method are statistically compared with measured signals from conducted measurement runs. In the context of virtual hull monitoring, a literature review (e.g Aarsnes et al.[7], Hulkkonen et al.[8], among others), reveals that the majority of publications center around simulations in the frequency domain. In all of these applications, the emphasis is placed on determining fatigue damage along a predefined route. The examination of wave conditions in the frequency domain can be conducted at various levels of accuracy. However, even the most precise representation using reconstructed 2D wave spectra (see Kumar et al. [9]) does not capture the full complexity of real sea conditions. While above mentioned studies demonstrate that the statistical analyses in the frequency domain provide sufficient accuracy for determining fatigue damage, they are unable to capture dynamic effects and non-linear load impacts from high steep waves due to the absence of phase information in the waves. Since these effects can trigger high stresses, they have an impact on fatigue damage, but can also have drastic consequences for the ultimate limit state. This limitation poses a challenge on the path towards comprehensive hull monitoring.

One approach to overcome these limitations is virtual hull monitoring in the time domain. A possible method to achieve this could look like the following: The near-field environment of the ship (wave environment) is recorded using wave radar. A machine learning method predicts the wave elevation at the time of impact with the ship, see Wedler et al. [10] as reference. With this information, numerical simulations are to be conducted to provide the structural responses of the ship in waves. Given the computational intensity of these simulations, they cannot be performed in real-time, hindering the aspect of foresight in maritime operations. Therefore, using the described numerical simulations, a dataset can created, featuring the wave elevation time series as features and the structural response as targets. This dataset trains a machine learning model, which later uses the wave radar ML model output as input to predict structural responses in real-time. This approach aligns closely with the foresight aspect of ship operations, as it enables proactive decision-making by providing predictions of structural behavior under varying conditions. By anticipating how the structure will respond to dynamic wave loads, operators can adjust operational parameters in advance, mitigating risks and ensuring the safety and efficiency of ship operations. The integration of machine learning-based real-time predictions thus bridges the gap between data collection and actionable foresight, offering a transformative solution to the challenges of operational guidance in complex maritime environments.

Given the extensive nature of comprehensively achieving this objective, this study concentrates on the applicability of machine learning methods as surrogate models to predict structural responses. The objective of the comparative analysis is to identify an appropriate machine learning method that yields optimal results. For the purposes of this study, the input is comprised of simplified wave parameters, like wave frequency, direction and phase angle as opposed to wave time series captured with wave radar.

The results of this study demonstrate that machine learning methods are well-suited as surrogate models for predicting the structural responses of ships in waves. Compared to direct simulations, these models offer significantly reduced computational time, enabling real-time application.

2. MACHINE LEARNING METHODS

This chapter provides a brief overview of the theory underlying the machine learning methods utilized.

An Artificial Neural Network (ANN) is a computational framework inspired by the structure and function of biological

neural networks. An initial description was provided from McCulloch and Pitts [11]. In a modern form it consists of interconnected layers of neurons, each performing simple mathematical operations to process data. The network is organized into three main types of layers: an input layer, which receives raw data; one or more hidden layers, and an output layer, which produces the final prediction or result. The neurons within each layer are connected to those in adjacent layers, with these connections assigned weights that determine the strength and influence of the signal passing through them. The neurons of ANNs employ activation functions to introduce non-linearity, enabling the network to model complex relationships in the data. The learning process involves adjusting the weights through optimization algorithms like backpropagation and gradient descent, minimizing the error between the network's predictions and the actual targets.

A specific implementation of a feedforward ANN, the TensorFlow Sequential API, provides a simple way to construct models where the data flows unidirectionally from the input layer, through the hidden layers, to the output layer. Each layer processes the output of the previous one, forming a direct pipeline of data transformation ([12]). The tensorflow Sequential API is used in this paper to predict the maximum stresses as target out of basic wave data as input.

XGBoost (eXtreme Gradient Boosting, see [13]) is a machine learning algorithm based on decision trees, designed to be very fast, accurate and efficient. At its core, XGBoost builds a series of decision trees, but instead of making all the trees at once, it creates them one by one. Each new tree focuses on fixing the mistakes made by the previous trees. This method of learning from errors is called gradient boosting. While one tree alone might not be very accurate, combining many of them into an ensemble creates a strong predictive model. XGBoost improves this process by making each tree better at addressing the specific errors made in earlier steps. It uses advanced techniques, such as second-order optimization, which improves how it learns from errors. Regularization prevents the model from overfitting (memorizing instead of generalizing). Sparsity handling, which works well with missing data or features. It works well even with very large datasets.

3. METHODOLOGY

To utilize machine learning methods for predicting ship responses, it is essential to have a dataset that includes both wave parameters and the corresponding ship responses. Therefore, an artificial dataset is constructed through numerical simulations. These simulations comprise hydrodynamic analyses, load mapping, and coupled structural simulations. The process, outlined in Figure 1, is detailed in the following section.

3.1 Dataset generation

To effectively depict the correlation between wave activity and the resulting structural responses, a coupled methodology is used, where hydrodynamic simulations are conducted to determine the hydrodynamic pressures exerted on the hull. In this study the boundary element (BEM) solver ANSYS AQWA is used. It is based on potential flow theory and assumes inviscid, incompressible, and irrotational flow. The pressures are subsequently mapped to a finite element model.

For hydrodynamic diffraction analysis, a simplified geometry of a bulk carrier consisting of the outer surface of the hull is adapted. AQWA imposes some restrictions on the elements used to discretize the model regarding the elements' size with relation to neighboring elements and wave parameters. Following these restrictions, a mesh is generated with a maximum element size of $3.5m$, resulting in about 1500 elements. The discretized model is presented in Figure 2. In this study long-crested, regular waves are input to the machine learning model and therefore modeled in the BEM solver. To represent a broad spectrum of waves a number of 15 frequencies with equal intervals is inserted to be simulated. Furthermore, the wave directions are specified based on defined intervals from -180 to 180, where a range of directions is chosen at 45 intervals to represent head, following and oblique wave conditions. In addition for every wave frequency and direction combination a sequence of wave phase angles is defined, which enables a time series of the wave elevation. The simulations are obtained for every single wave parameter combination. The resulting pressures acting on the hull's outer surface are stored in the user files directory containing every combination of the selected frequencies, directions, and phase angles, along with information about the element coordinates, and the generated mesh. The result files generated from the hydrodynamic diffraction analysis are to be used by AQWA Wave to map the pressures resulting from the hydrodynamic analysis to the structural analysis environment. This process will generate load information files that contain the pressure information applied on each finite element mesh element for all of the predefined load cases consisting of the applied wave frequencies, directions, phase angles and heights. Using this approach, the pressures resulting from the hydrodynamic diffraction analysis are mapped onto the finite element model and the structural simulation environment is initialized.

The geometry used in the structural analysis is composed of a simplified model of the hull including first and second order elements like frames and stiffeners. The structure is discretized using shell elements, which are most suitable for representing thin walled structures, with an element edge length of 1 m. The material of the structure is structural steel with a density of 7850kg/m3, Young's modulus of 2.10e11Pa and Poisson's ratio of 0.3. Appropriate boundary conditions have to be imposed on the FE model to result in a solvable system of finite element equations. The boundary conditions are defined as specified by the American Bureau of Shipping (ABS), as shown in Figure 3. As per ABS guidelines, spring constraints have to be applied on both ends of the ship hull. The spring system constrains the longitudinal, vertical and transversal motions of the structure. Once the simulation setup is complete and the hydrodynamic pressures are imported, the structural analysis is initiated to generate the stress results. For the purpose of this work, a total of 675 test cases are implemented consisting of 15 frequencies, 9 directions and 5 phase angles assuming regular, long crested and stationary waves. Both the hydrodynamic and structural mechanics solvers are linear, so the structural response epends linearly on the wave height. Therefore, for the purposes of this study, the wave height is kept

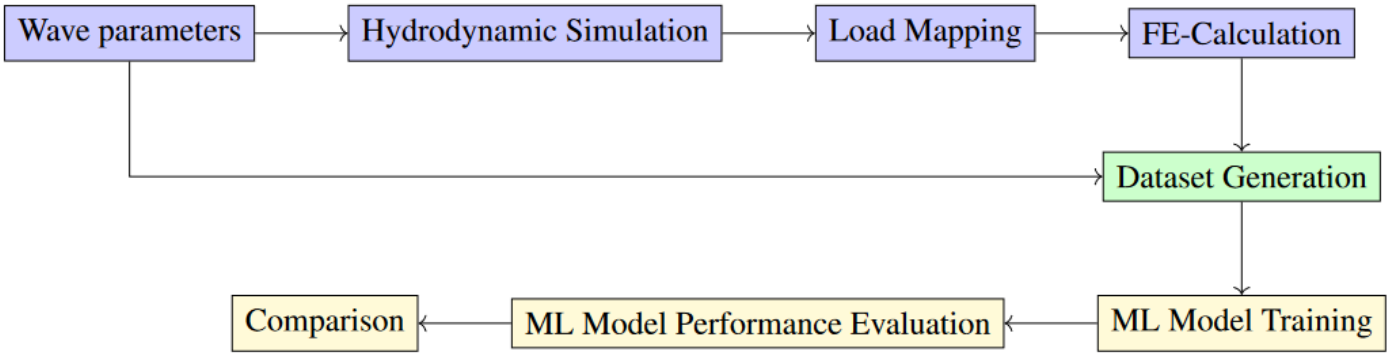


FIGURE 1: WORKFLOW TO CREATE A DATASET AND TRAIN THE ML MODEL

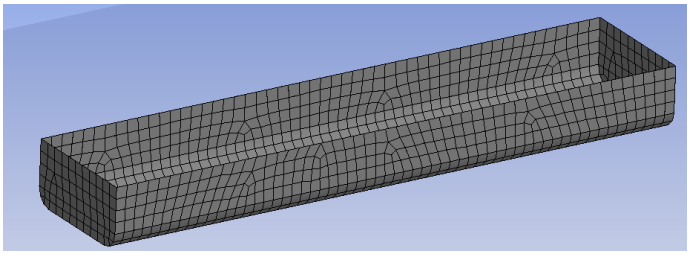


FIGURE 2: DISCRETIZED AQWA MODEL

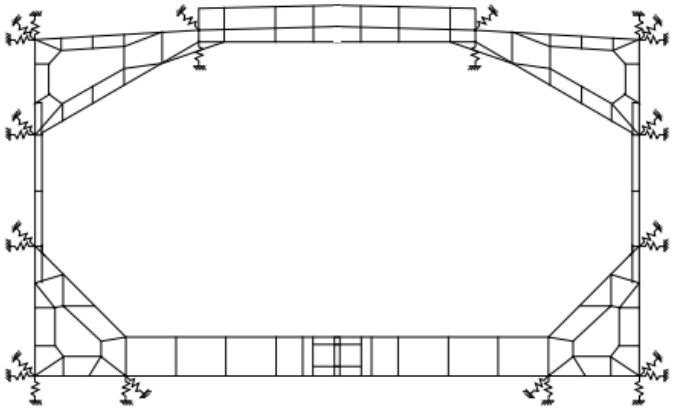


FIGURE 3: ABS BOUNDARY CONDITIONS SPECIFICATION

constant and serves not an input feature for the machine learning model. The results obtained from the simulations contain a list of all elements along with their coordinates and the equivalent von-Mises stresses in a comma separated format (.csv file).

Following the generation of result files from the structural analysis, it is imperative to transform the data into a functional format during the data preprocessing stage. The raw data exported from ANSYS Mechanical comprises the results for each test case in discrete files. Thus, a Python script is developed to extract the maximum stress values from each test, along with the element number where it occurs and its coordinates of location. Additionally, the relevant wave parameters are integrated and appended to a newly constructed dataset. The compiled dataset comprises 675 entries, corresponding to each load case (wave parameter combination). Each entry contains the wave parameters, namely frequency, direction, and phase, along with the maximum von Mises stress for the specific load case and the coordinates of the location where the maximum stress is observed.

3.2 Training of machine learning models

The following section explains the development and training of the two methods. The discussion begins with a description of the neural network employed. During its development, it was observed that the best results are achieved when two independent artificial neural networks are trained separately: one dedicated to predicting the maximum stress value and the other to determining the location of maximum stress.

The dataset generated as explained above is split into a training set to train the model and a test set to give an estimation of the model's performance. The training and test sets will be split into 80% and 20%, respectively. Furthermore, the data has to be normalized to ensure the convergence, stability and performance of the model by scaling the features and outputs to a common scale. To address this requirement, the min-max scaler from the scikit-learn library is employed.

After normalizing the data, a sequential model is used to initialize the ANN model. To achieve optimal results, both neural networks are subjected to hyperparameter tuning using a grid search. This process uses a brute force approach to optimize the number of epochs, the number of layers, the number of neurons per layer, and the learning rate. The resulting parameters define the final neural networks.

Consequently, the stress prediction model is composed of an input layer with three neurons, corresponding to the three features. Subsequently four hidden layers with 512 neurons in each layers and one output layer finalize the model. To minimize overfitting dropout layers with a value of 0.2 are inserted. The training with 1000 epochs and a learning rate of 0.001 resulting in a minimal training loss. In addition, the position prediction model is trained with 2500 iterations, four hidden layers with 612 neurons in each layer, and a learning rate of 0.001.

For the purpose of comparison two XGBoost regression models are initialized, to predict both the maximum stress and its location in ship structures, each tailored to its specific task through hyperparameter tuning. For the maximum stress prediction, the

tuning process results in a model configured with a squared error objective, 300 estimators, a learning rate of 0.2, and a maximum tree depth of 5. These parameters strike a balance between model complexity and training speed. The model is trained on scaled input data to enhance convergence and performance. For the prediction of the location of maximum stress, a separate XGBoost model is optimized using an alternative set of parameters. This model employs 400 estimators, a learning rate of 0.01 to allow finer adjustments during training, and a maximum tree depth of 9 to capture more complex relationships in the data. Additionally, subsampling and column sampling rates were set to 0.8 and 1.0, respectively, to prevent overfitting while maintaining sufficient feature diversity. Like the stress prediction model, it is trained on scaled data to ensure reliability in predicting the spatial location of the maximum stress.

To evaluate the performance of the ML models, two metrics are used, mainly the root mean square error and the R-square value. The smaller the root mean square error value, the better the performance. In addition, the R^2 value measures the proportion of the variance of the dependent variable that is explained by the independent variable [14]. Since the R^2 value is a ratio of the sum squared regression to the total sum of squares, the closer it is to one, the better the predictions.

4. RESULTS

This chapter focuses on presenting and comparing the results of the machine learning models. The comparison is structured as follows: First, the predictions of the maximum stress by both machine learning models will be compared with one another. Subsequently, the comparison of the predicted locations of the maximum stress will be presented. Each model is tested using the same test dataset.

A representative depiction of all data points in the test set can be seen in Figure 4. It shows scatter plots for predicting the maximum stress values for both, ANN and XGBoost. The x-axis represents the numerically determined value, while the y-axis represents the values predicted by the machine learning model. The bisector (x=y) represents perfect agreement between both values. The prediction of the maximum stress using artificial neural networks in Figure 4a demonstrates particularly good agreement between predictions and numerical values in the range up to 80 MPa. However, at higher stress levels, the scatter in the predictions increases.

In comparison, the predictions made using XGBoost show better performance in Figure 4b. The results align closer with the bisector across the entire stress range, indicating stronger agreement between the predictions and the numerical values.

The efficacy of stress prediction by the machine learning models can be quantitatively evaluated using the aforementioned performance metrics. For the ANN model, the Root Mean Square Error is reported as $13.20 MPa$, while the R^2 value reaches 0.71. Consistent with the findings from the scatter plots, the performance metrics of the XGBoost model demonstrate superiority. Specifically, the R^2 achieves 0.83, and the RMSE is recorded at $9.83 MPa$. An overview of the overall metrics can be found in Table 1.

In addition to predicting the maximum stress, this paper aims

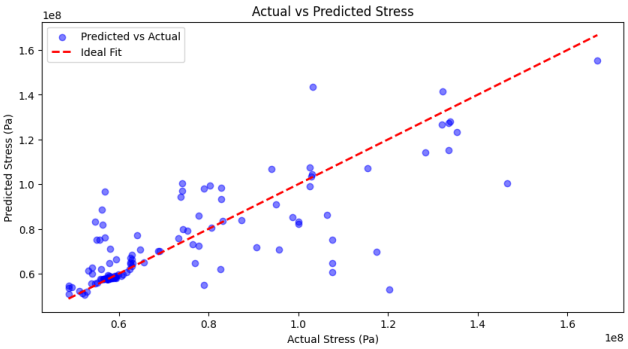
TABLE 1: PERFORMANCE VALUES FOR STRESS PREDICTION

ML Model	R^2 [-]	$RMSE$ [MPa]
ANN	0.7086	13.20
XG Boost	0.8383	9.83

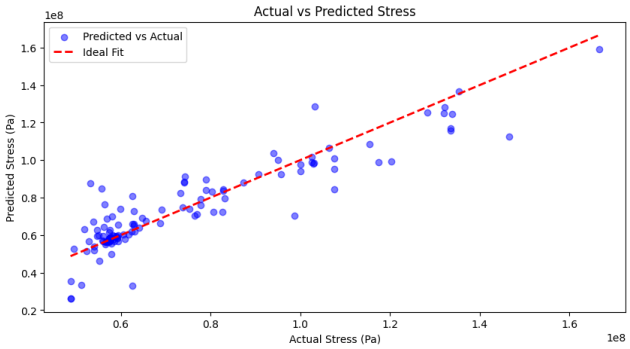
to predict the location of the maximum stress in the ship structure. This location is predicted as a vector with three components (x,y,z) using the two models under comparison. The residual plots in Figure 5 display the difference between the true values and the predicted values for each data point, highlighting how well the machine learning models (ANN and XGBoost) perform in predicting the x-, y-, and z-positions of maximum stress. Ideally, residuals should be centered around zero with minimal spread, indicating accurate and unbiased predictions. In Figure 5, the residuals for the x-position predictions are shown. For the ANN model, the residuals in Figure 5a are scattered around zero, with some significant outliers reaching up to $150m$, indicating high variability in prediction accuracy. This suggests that the ANN model struggles with the complexity or noise in the x-position data. The XGBoost residuals, while also centered around zero, exhibit a tighter distribution with fewer extreme outliers, suggesting slightly better performance in handling the prediction task. However, the substantial spread in residuals for both models highlights their overall inaccuracy and unreliability in predicting the x-coordinate.

In Figure 6, the residuals for the y-position predictions are displayed. For the ANN model, the residuals are concentrated around zero, with most values falling within $2.5m$. Outliners reach values of $15m$ This relatively narrow and uniform spread indicates consistent prediction accuracy across the dataset. Similarly, the XGBoost model shows a comparable pattern, with slightly less variation in the residuals, suggesting marginally better performance. Unlike the x-position predictions, the residuals for both models in this case demonstrate a reliable level of accuracy, with errors remaining minimal and evenly distributed. In Figure 7, the residuals for the z-position predictions are shown. For the ANN model, the residuals are tightly clustered around zero, with the majority falling within $0.25m$. Outliners rach values up to $0.75m$ This narrow spread indicates strong and consistent prediction accuracy in the z-dimension. Similarly, the XGBoost model displays a comparable residual distribution, with slightly lower variance and an even tighter clustering around zero, suggesting marginally better performance. The minimal and well-distributed residuals for both models highlight their reliability in predicting the z-coordinate.

The heatmap in Figure 8 presents the Root Mean Squared Error values, measured in meters, for predicting the location (x, y, z) of the maximum stress in the structure across both machine learning models: Artificial Neural Network and XGBoost. A clear disparity in performance between the models is evident, particularly in the prediction of the x-component. The ANN model exhibits the highest error for the x-component, with an RMSE of 52 meters, indicating a significant lack of precision. XGBoost outperforms the ANN with an RMSE of 47 meters,

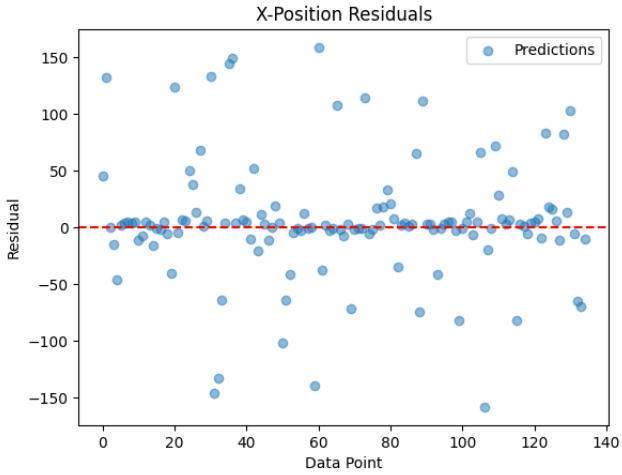


(a) Scatter of ANN Stress Prediction

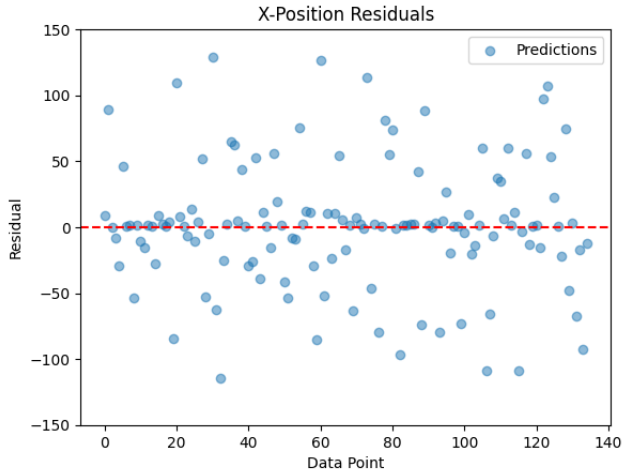


(b) Scatter of XG Boost Stress Prediction

FIGURE 4: SCATTER PLOTS FOR STRESS PREDICTIONS

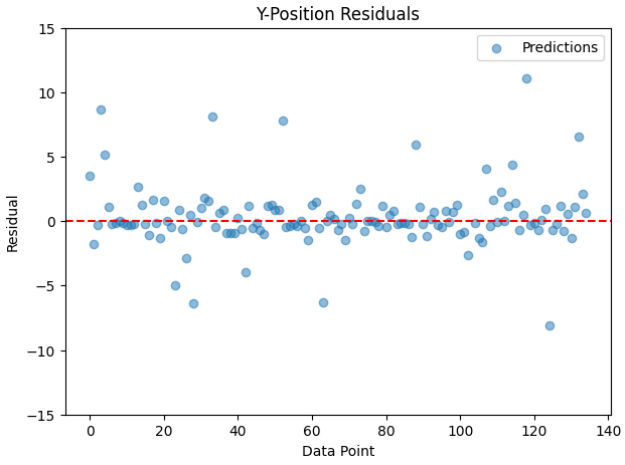


(a) Residuals of ANN x-Position Prediction

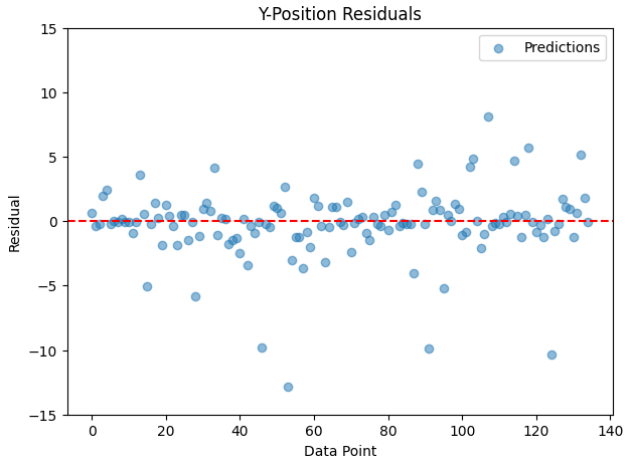


(b) Residuals of XGB x-Position Prediction

FIGURE 5: RESIDUALS OF X-POSITION PREDICTIONS



(a) Residuals of ANN y Position Prediction



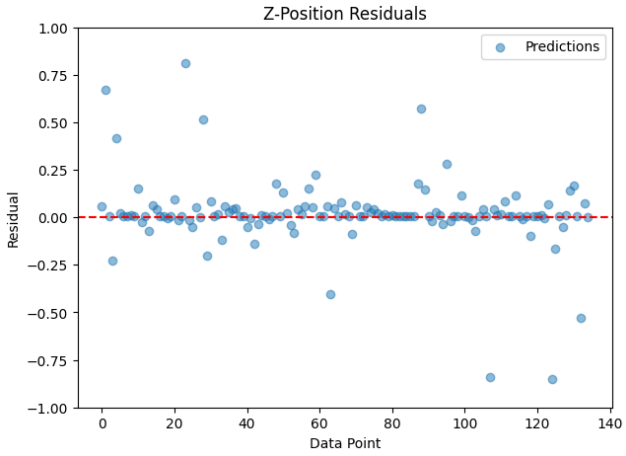
(b) Residuals of XGB y Position Prediction

FIGURE 6: RESIDUALS OF Y-POSITION PREDICTIONS

demonstrating slightly better predictive accuracy. Never the less the prediction on the x component of every model is not acceptable and therefore unreliable.

For the y-component, both models achieve much lower RMSE values compared to the x-component, highlighting their improved performance for this spatial dimension. Among the two

(a) Residuals of ANN z Position Prediction



(b) Residuals of XBG z Position Prediction

FIGURE 7: RESIDUALS OF Z-POSITION PREDICTIONS

models, XGBoost performs marginally better with an RMSE of 2.7 meters, closely followed by ANN at 2.8 meters.

The z-component exhibits the lowest RMSE values across all models, reflecting accurate predictions with minimal error. XGBoost once again achieves the best performance, with an RMSE of 0.18 meters, followed by ANN at 0.21 meters. Though the differences among the models are relatively minor.

Overall, the comparison shows that XGBoost tends to outperform ANN slightly in predicting the location of maximum stress, particularly for the x- and z-positions, while both models perform similarly for the y-position. The residual plots and RMSE heatmap collectively provide a clear depiction of the models' predictive accuracy and error distribution across the different components.

The objective of this paper is to demonstrate the capabilities of machine learning models in predicting the structural responses of ships subjected to wave conditions in real time. Emphasis is placed not only on the accuracy of the results but also on computational efficiency to enhance real time applicability. To illustrate this, a test scenario comprising a single wave event has been established. The methodologies described in the study are employed to predict the maximum stress and its location. The numerical simulation encompasses the hydrodynamic analysis, pressure mapping, and subsequent mechanical simulation. The machine learning approaches map the input wave parameters directly to the corresponding structural response. However, the machine learning models utilize two separate models, one for predicting maximum stress and another for determining its location. Thus, the computational time is the aggregate of these two predictions. All analyses are executed on a high-performance desktop computer. Table 2 indicates that the computational time for the numerical simulations is approximately $28s$. The prediction via an artificial neural network takes $0.04s$, while the XGBoost prediction requires $0.001s$. The findings clearly demonstrate that the machine learning models surpass the numerical simulations in this context, thereby facilitating the real-time application of the predictions.

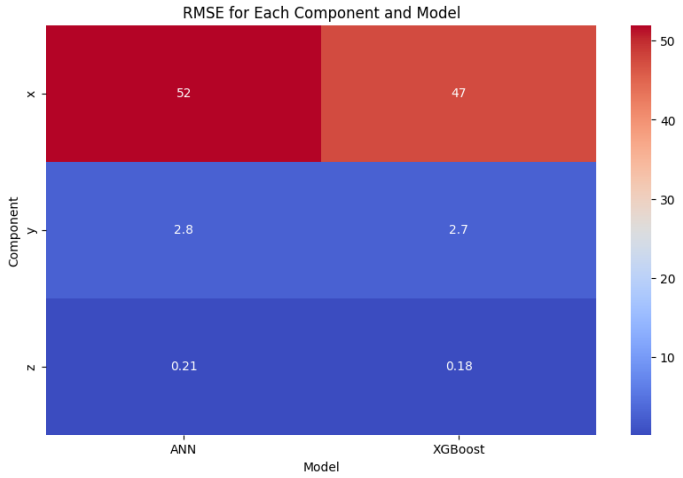


FIGURE 8: RSME HEATMAP LOCATION PREDICTION

TABLE 2: COMPUTATIONAL TIME FOR PREDICTING SHIP STRUCTURAL RESPONSES

ML Model	t [s]
Simulation	28
ANN	0.04
XG Boost	0.001

5. DISCUSSION

The results presented highlight distinct differences in the performance of the machine learning models for predicting both the maximum stress and its location. The ANN and XGBoost models demonstrate their strengths in stress prediction.

The stress prediction results underscore the superior performance of XGBoost, which achieves the lowest RMSE and the highest R^2 value, indicating its ability to capture the relationships inherent in the dataset effectively. XGBoost's gradient boosting framework allows it to iteratively correct errors from previous predictions, leading to better generalization and precision. ANN also performs reasonably well, particularly for stress levels below 80 MPa, but its increased scatter at higher stress levels suggests potential limitations in capturing the complexity of the data or an insufficient representation of extreme values in the training set.

When examining the predictions of the stress location, the results present a similar trend, with XGBoost emerging as the most accurate model overall. For the x-component, however, both models fail to deliver reliable predictions, with high RMSE values suggesting that this component may be more challenging to predict. The ANN and XGBoost models perform comparably well for the y- and z-components, with XGBoost showing marginally better accuracy. The disparity in x-component performance across models highlights potential limitations in the input data or feature engineering, suggesting a need for further exploration to improve predictions for this component.

The analysis reveals XGBoost as the most robust and versatile model for both stress prediction and stress location, excelling in most metrics. ANN proves competitive in many aspects but exhibits some weaknesses at higher stress levels and slightly less precision in spatial predictions. These findings emphasize the importance of selecting models tailored to the complexity and characteristics of the problem, as well as the potential value of integrating feature selection or engineering techniques to enhance performance further.

The significant inaccuracy in forecasting the location of the maximum stress underscores possible shortcomings in the dataset employed. Specifically, the x-component appears to be less strongly correlated with the provided wave parameters (frequency, direction, and phase), making it more challenging for the model to learn a clear predictive relationship. Additionally, numerical errors, such as singularities in the simulation data, may contribute to increased prediction variability. Furthermore, the dataset used in this study does not include variations in wave height, which may limit the model's ability to generalize across a broader range of structural responses. This indicates that the dataset may be lacking in information concerning stresses and their locations beyond a specific range, which could restrict the models' capacity to predict accurately in those areas.

In addition to assessing the accuracy of the results, it is imperative to elucidate the numerical limitations inherent in the dataset creation process. The described procedure initiates with hydrodynamic simulations of regular long-crested waves. The successive simulations proceed with incremental phase angles to represent hydrodynamic loads within the time domain. Nonetheless, the limitations of the employed hydrodynamic simulation environment preclude a comprehensive representation of reality. Initially, it only accounts for regular long-crested wave packets, whereby time-discretized simulations offer no substantial benefit. The inclusion of irregular waves would yield advantages over the frequency domain. Furthermore the Boundary Element Method in AQWA is based on potential flow theory and assumes inviscid, incompressible, and irrotational flow. This approach allows for efficient hydrodynamic simulations but comes with certain simplifications, such as neglecting viscous effects, nonlinear wave-structure interactions, and turbulence. These assumptions introduce limitations in accurately capturing highly nonlinear phenomena, such as wave breaking or extreme slamming events. Thereby constraining the study to the global ship response, as local effects such as slamming at the bow cannot be examined. Moreover, the simulations assume the absence of forward ship speed. While this assumption is not a critical limitation for frequency domain simulations, given that ship speed merely alters the encounter frequency, it would significantly affect nonlinear simulations in the time domain. The hydrodynamic loads thus determined are transferred individually to the mechanical model for each phase angle (load step). This transfer constitutes a one-way fluid interaction between the hydrodynamic model and the mechanical model. It is presumed that the application in hull monitoring does not necessitate a fully coupled simulation environment, thereby rendering the method suitable for future studies.

It is important to note that the finite element model of the ship is simplified, and consequently, the structural responses do not entirely replicate those of an actual vessel. Nonetheless, the model is sufficiently detailed to capture local stress effects.

The machine learning models developed in this study are based on a simplified input format, considering only long-crested waves from a single direction. In reality, wave environments are significantly more complex, often consisting of multidirectional and short-crested waves. As a result, the trained models may not be directly applicable to real-world data without further adaptation. This limitation highlights the need for future research to enhance the models by incorporating a more diverse wave representation, ensuring improved generalizability and applicability to real-world conditions.

6. CONCLUSION

This study compared ANN and XGBoost for predicting maximum stress and its location in ship structures based on wave parameters. Both models demonstrated reliable performance in predicting maximum stress, with XGBoost showing slightly better accuracy overall. However, the prediction of the location of maximum stress remains critical, particularly due to significant errors in the x-coordinate predictions, which are the main driver of inaccuracies. While the models performed well for the y- and z-coordinates, the x-component's unreliability highlights the need for further refinement to improve location predictions.

REFERENCES

[1] Drummen, Ingo, Benhamou, Alexis, Hageman, RB. and Stambaugh, Karl. "Hull Structure Monitoring of a New Class of US Coast Guard Cutters." *ASNE Technology, Systems & Ships* (2019).

[2] Anyfantis, Konstantinos N. "An abstract approach toward the structural digital twin of ship hulls: A numerical study applied to a box girder geometry." *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment* Vol. 235 No. 3 (2021): pp. 718–736. DOI 10.1177/1475090221989188.

[3] Lang, Xiao, Zhang, Mingyang, Zhang, Chi, Ringsberg, Jonas W. and Mao, Wengang. "Physics-guided metamodel for vertical bending-induced fatigue damage monitoring in container vessels." *Ocean Engineering* Vol. 312 (2024): p. 119223. DOI 10.1016/j.oceaneng.2024.119223.

[4] Serr, Johanna, Wedler, Mathies, Stender, Merten, Fonseca, Nuno, Soares, C. Guedes, Hoffmann, Norbert, Ehlers, Sören and Klein, Marco. "Data-driven, non-linear ship response prediction based on time series of irregular, long-crested sea states amidships." *Ocean Engineering* Vol. 317 (2025): p. 119963. DOI 10.1016/j.oceaneng.2024.119963.

[5] Schirmann, Matthew, Collette, Matthew and Gose, James. "Ship Motion and Fatigue Damage Estimation via a Digital Twin." *Life-Cycle Analysis and Assessment in Civil Engineering: Towards an Integrated Vision* (2018): pp. 2075–2082.

[6] Thompson, Ian. "Virtual hull monitoring of a naval vessel using hindcast data and reconstructed 2-D wave spectra." *Marine Structures* Vol. 71 (2020). DOI 10.1016/j.marstruc.2020.102730.

[7] Aarsnes, Lars Holterud, Storhaug, Gaute and Radon, Maciej. "Utilization of Structural Design Models in Operation to Monitor Fatigue Strength Performance." Okada, Tetsuo, Suzuki, Katsuyuki and Kawamura, Yasumi (eds.). *Practical Design of Ships and Other Floating Structures*. Vol. 64 of *Lecture Notes in Civil Engineering*. Springer Singapore, Singapore (2021): pp. 622–636. DOI 10.1007/978-981-15-4672-3_38.

[8] Hulkkonen, Tapio, Manderbacka Teemu and Sugimoto Kei. "Digital Twin for Monitoring Remaining Fatigue Life of Critical Hull Structures." *Digital Twin for Monitoring Remaining Fatigue Life of Critical Hull Structures* (2019): pp. 415–427.

[9] Kumar, Nirnimesh, Cahl, Douglas L., Crosby, Sean C. and Voulgaris, George. "Bulk versus Spectral Wave Parameters: Implications on Stokes Drift Estimates, Regional Wave Modeling, and HF Radars Applications." *Journal of Physical Oceanography* Vol. 47 No. 6 (2017): pp. 1413–1431. DOI 10.1175/JPO-D-16-0203.1.

[10] Wedler, Mathies, Stender, Merten, Klein, Marco and Hoffmann, Norbert. "Machine learning simulation of one-dimensional deterministic water wave propagation." *Ocean Engineering* Vol. 284 (2023): p. 115222. DOI 10.1016/j.oceaneng.2023.115222.

[11] Mcculloch, Warren and Pitts, Walter. "A logical calculus of the ideas immanent in nervous activity." *Bulletin of mathematical biophysics* No. 5 (1943): pp. 115–133.

[12] Tensorflow. "The Sequential Model." (2023). URL https://www.tensorflow.org/api_docs/python/tf/keras/Sequential.

[13] Chen, Tianqi and Guestrin, Carlos. "XGBoost." Krishnapuram, Balaji, Shah, Mohak, Smola, Alex, Aggarwal, Charu, Shen, Dou and Rastogi, Rajeev (eds.). *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: pp. 785–794. 2016. ACM, New York, NY, USA. DOI 10.1145/2939672.2939785.

[14] Tensorflow. "Metrics Documentation." (2023). URL https://www.tensorflow.org/api_docs/python/tf/keras/metrics.