Journal of Naval Architecture and Marine Engineering
July, 2026
https://doi.org/10.3329/jname.v23i2.72551 https://www.banglajol.info

A METHODOLOGICAL CASE STUDY OF MACHINE-LEARNING INTERPOLATION FOR CATAMARAN MOTION PREDICTION USING THE NPL EXPERIMENTAL DATASET

A. Saha*, S. A. Sumaiya and N. M. Golam Zakaria
Dept. of Naval Architecture and Marine Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh,
*argha728@gmail.com, alamsafia07@gmail.com, gzakaria@name.buet.ac.bd

Abstract:

The assessment of catamaran motion during early-stage design is of great significance, since it can influence both structural design and operational considerations. Conventional methods for estimating ship motion, such as theoretical calculations, numerical modeling, and simulations, may have lower accuracy, higher costs, constraints, and uncertainties. The use of machine learning methods has become popular recently due to the surge in data. In this study, a data-driven machine-learning methodology is employed as a methodological case study to interpolate the motion characteristics of NPL catamarans within a limited experimental dataset. This is achieved by using experimental data on the motion of NPL catamarans in long-crested head seas and applying regression methods. The regression models utilized in this study include polynomial regression, K-nearest neighbor regression, regression trees, random forests, and artificial neural networks. These methods were implemented in Python using optimized parameters. When comparing these algorithms, the principal aspect was their in-domain interpolation accuracy for the studied NPL cases. This work proposes a complete data preparation pipeline, including design parameter selection, data extraction, and preprocessing. The study also compares the predicted motion with results obtained from different numerical approaches, such as the 3D Green's function, strip theory, and a 3D pulsating source. The results indicate that some of the employed machine-learning models can reproduce heave and pitch trends with good accuracy within the parameter range studied in the NPL dataset. However, owing to the limited nature of the dataset, including constant hull-form coefficients, only two demi-hull spacing ratios, and restriction to long-crested head seas, the resulting models should be interpreted as interpolation tools within this dataset rather than as a general preliminary design tool for catamarans.

Keywords: Ship motion prediction, machine learning, regression analysis, data-driven modeling, artificial neural networks

NOMENCLATURE

L	Demi hull length between perpendiculars [m]	∇	Volume of displacement (demi-hull) [m^3]
B	Demi hull maximum breadth [m]	LCB	Longitudinal center of buoyancy
T	Demi-hull draught [m]	Fn	Froude number
NPL	National Physical Laboratory	S	Separation between catamaran (demi-hull) [m]
C_B	Block coefficient (demi-hull)	***Greek symbols***	
C_P	Prismatic coefficient (demi-hull)	λ	Wavelength [m]
C_M	Midship section area coefficient (demi-hull)		

1. Introduction

Multihulls demonstrate favorable attributes as vessel configurations owing to their notable expansive deck space and substantial transverse stability. Naval architects are attracted by multihulls, particularly catamarans, when confronted with the need to both accommodate a spacious deck and achieve high speeds in ship design by maintaining the large overall beam of the vessel as well as keeping the single hulls of the catamaran relatively slender. Contemporary maritime operations utilize catamaran designs for both commercial and naval purposes. But the drawback of the relatively slender catamaran is its low-restoring qualities when subjected to motion in a seaway. Thus, it may result in large relative motion, which increases the probability of green water, deck slamming, or even nose diving (van't Veer, 1998). Usually, the design of a vessel involves several distinct phases,

 Received on: April, 2024

including concept design, preliminary design, contract design, and detailed design. The lack of longitudinal restoring properties consistently makes the accurate prediction of the seakeeping characteristics of catamaran ships under seaway a subject of significant practical importance, as it impacts both vessel design and operational aspects. For example, the vertical movements experienced by such catamaran ships might put limitations on their operations, ultimately leading to a deterioration of working conditions on board. Moreover, these occurrences are linked to the dynamic forces applied to the vessel structure and equipment, which may result in water accumulation on the deck. Hence, the assessment of vessel motion, particularly for catamaran hull forms, is important during early design considerations.

Conventional approaches for assessing ship motion include theoretical computations, numerical modeling, and simulation, all of which may be costly and time-consuming. Initially, evaluation techniques focused on forecasting calm water resistance and seakeeping performance using potential flow theory. In parallel, various advanced computational fluid dynamics (CFD) tools have been developed utilizing the concept of the free surface RANSE (Reynolds-Averaged Navier-Stokes Equations) method, which has experienced significant advancements due to the rapid development of technology. The first-order Rankine panel method was used to predict the seakeeping properties of Du Toit Yacht Design (DUT) catamaran 372 by van't Veer (1998). Due to the omission of viscosity, potential flow theory-based approaches have certain limitations, particularly in accurately predicting the damping effects on the catamaran's motions in waves. As a result, motions influenced by viscous effects, including boundary-layer-related damping and drag, may not be captured accurately. In their study, Mittendorf and Papanikolaou (2020) devised a simplified thin-ship theory panel method to forecast resistance in calm water. This method was implemented in the optimization process of battery-powered, fast, zero-emission catamarans, which includes the NPL 5b catamaran and the Wigley hull. It is suggested that the approach be implemented during the initial phases of fast catamaran design. Doğrul et al. (2021) conducted a numerical investigation utilizing URANS (Unsteady Reynolds Averaged Navier-Stokes) to analyze the heave and pitch transfer functions of the Delft catamaran 372 for regular head waves, with a Froude number of 0.3. To predict the trim and sinkage of a fast displacement catamaran operating in shallow water, Gourlay (2008) introduced a slender body theory that utilizes the linear superposition of the flow fields produced by each hull. Using NPL high-speed round bilge displacement hull series demi-hulls, sinkage and trim were computed for subcritical, transcritical, and supercritical flows. The experimental findings of surge, heave, and pitch movements of a moored FPSO (floating production, storage, and offloading) unit were presented by Soares et al. (2014). These results were subsequently compared to the computed values obtained through the Green's function panel method and strip theory. In their study, Qian et al. (2015) devised a novel small-waterplane-area-twin-hull (SWATH) featuring inclined struts. To analyze the heave RAO, pitch RAO, and added resistance properties, they developed Salvesen-Tuck-Faltinsen (STF) strip theory. Subsequently, they integrated the wave generation module with the RANSE code. A comparative analysis of different methods for predicting added resistance and wave-induced ship motions was conducted by Lee et al. (2021). Several numerical methods for seakeeping computations were employed to compare the outcomes of two distinct experimental approaches (i.e., captive, and free-running model experiments). Accuracy and dependability, among other attributes, were examined for each analysis method.

More recent studies have continued to examine catamaran motion in waves using numerical and hybrid numerical-experimental approaches. Han et al. (2023) investigated the motion responses of a cabin-suspended catamaran in head waves using numerical calculation with experimental validation and showed that motion reduction through tuning of suspension parameters depends strongly on the encounter frequency. Nguyen et al. (2025) numerically studied the motions and wave-induced loads of a low-speed uncrewed surface vehicle catamaran in both regular and irregular waves and reported good agreement of heave, roll, and pitch with experimental data. The study also highlights the importance of vertical shear force and bending moment assessment. In a related head-wave study, Zhang et al. (2024) analyzed the hydrodynamic performance of catamarans in different formation configurations. The study showed that hull spacing, and relative arrangement can significantly influence resistance and motion responses. Additionally, in their study, Santoso et al. (2026) focused on fast catamaran ferry design optimization These studies further confirm that catamaran hydrodynamics in waves remain sensitive to configuration, operating condition, and analysis approach.

Various semi-empirical formulations were also formulated in the early days, drawing inspiration from experimental findings (Bhattacharya, 1978). Recently, significant advancements have occurred in the practical application of data science and machine learning to develop a variety of data-driven models. Today, with increasing availability of computational and experimental data, prediction approaches may be further investigated to support motion assessment during early-stage design, particularly for interpolation within constrained and well-defined datasets.

To estimate the added resistance introduced by regular head waves, Cepowski (2020, 2023) constructed an artificial neural network (ANN) model and a collection of five ANNs. These models utilized initial ship design parameters, including length, breadth, draft, and Froude number. In addition, a mathematical function representing the developed neural network was provided for use during the preliminary design phase. In their experimental investigation, Najafi et al. (2018) utilized model testing to assess the hydrodynamic performance of a hydrofoil-supported catamaran (HYSUCAT) under three distinct hydrofoils. Subsequently, ANN was employed to predict total resistance, effective power, sinkage, and trim under varying Froude numbers and hydrofoil types. The ANN model developed by Martić et al. (2023) was trained using added resistance in regular head wave data acquired through Boundary Integral Element Method calculations for a variety of container ships operating at different speeds. This model has the potential to assist in the preliminary design phase of ships by estimating the power demands under real-world sailing conditions. In their study, Liu et al. (2021) utilized an artificial neural network for predicting the amplitudes of the non-dimensional heave and pitch motion of several commercial ships. This prediction was based on 13 intrinsic factors specific to the ships, which play a crucial role in determining their motions while encountering head waves. The motion data of typical merchant ships obtained by the well-established frequency-domain 3D panel method software was used to train the model. The performance of the trained network was deemed adequate for general use when it was applied to predict the heave and pitch motions of two ships not included in the dataset. This resulted in a satisfactory agreement with experimental findings. Recent studies have further expanded the use of machine learning in ship hydrodynamics beyond single-model prediction. Yang et al. (2024) investigated ship resistance prediction using different sample settings. The research demonstrated that stacking ensemble learning improved prediction accuracy for container ships, while transfer learning enabled a resistance model trained on one ship type to be applied successfully to another. In a catamaran-specific study, Nazemian et al. (2024) employed regression tree, support vector machine, and artificial neural network models for calm-water resistance prediction and coupled the machine-learning predictor with a genetic algorithm. The study was focused on fast catamaran ferry design optimization. Their results showed that the ANN-based predictor correlated well with measured resistance and could be integrated into a rapid optimization framework. These studies indicate that recent machine-learning applications in naval hydrodynamics are moving not only toward improved prediction accuracy, but also toward design-oriented workflows and cross-domain model transfer.

Various ship design studies have found additional applications of artificial neural networks and machine learning. In their study, Agostino et al. (2022) conducted an evaluation and comparison of various recurrent-type neural network models, such as long short-term memory (LSTM) and gated recurrent units (GRU), for the purpose of real-time short-term prediction (nowcasting) of ship motions in challenging high sea conditions. The dataset used for this analysis was obtained from computational fluid dynamics simulations of a self-propelled destroyer-type vessel. Li et al. (2017) introduced a model that utilized the NARX network to examine the time-series prediction of ship motion. Tian and Song (2023) combined wave input with an LSTM neural network to predict the short-term motion of a ship. In their study, Liu et al. (2022) employed the Reservoir Computing (RC) model, a machine learning approach, to predict the six-DOF motions (surge, sway, heave, roll, pitch, and yaw) of the KVLCC2 ship during irregular waves. The researchers then compared the predicted results to those obtained from mooring model experiments. Experiments and visualizations validate the trained model's dependability and applicability in actual ocean conditions. To develop a center bulb for fast displacement catamarans, the artificial neural network (ANN) technique was used to replicate the responses of a computational flow solver that is based on low Froude number theory regarding changes in design variables (Danışman 2014). The outcomes derived from the model indicate that it accurately emulates the flow solver. In addition to numerical and experimental analysis, the hydrodynamic performance of the NPL catamaran was evaluated to determine the impact of the optimized center-bulb configuration. Bassam et al. (2022) conducted a performance analysis, comparing the prediction accuracy of several machine learning regression techniques typically used in the field. The assessment was conducted using an operational dataset gathered from the domestic ferry named 'M/S Smyril,' with a focus on predicting ship speed in actual operating conditions. By utilizing a data-driven machine learning methodology, the prediction of the optimal ship speed is enhanced, leading to improved operational efficiency. More recently, machine-learning applications specific to catamaran hydrodynamics have become more visible in the literature. Hassan et al. (2025) applied several machine-learning algorithms to experimental NPL catamaran data to predict the residual resistance coefficient and identify the optimum transverse separation gap between demi-hulls. Their study is particularly relevant because it also relies on NPL-based catamaran data and highlights the importance of spacing as a design variable. In another catamaran-focused motion study, Caramoy and Aryawan (2025) developed an ANN-based MISO framework for roll motion prediction using operational inputs such as ship speed, wave height, wave period, and heading, and reported that the model was able to capture nonlinear roll behavior with satisfactory accuracy. These recent studies suggest that current machine-learning research on catamarans is gradually extending from resistance prediction toward broader motion-related applications, although the specific response variables, datasets, and operating conditions still differ significantly from one study to another.

Although recent studies demonstrate the growing use of machine learning for ship resistance estimation, design optimization, and selected motion responses, the available literature remains fragmented in terms of response type, vessel class, and data source. Recent catamaran-related machine-learning studies have focused mainly on calm-water resistance, spacing optimization, or roll prediction under specific operational inputs, whereas comparative studies addressing heave and pitch transfer functions of NPL catamarans in long-crested head seas remain limited. This motivates the present work, which examines several conventional regression models within a constrained but well-known experimental catamaran dataset.

The aim of this research work is to evaluate several conventional machine-learning regression models as interpolation tools for estimating the heave and pitch transfer functions of NPL catamarans within the limited experimental dataset. The models are trained using experimental data from model testing and are examined within the restricted domain of high-speed round-bilge NPL catamarans in long-crested head seas. The models are compared with respect to their in-domain interpolation accuracy for the studied NPL cases, and their performance is evaluated under the present data constraints. Various data-preparation techniques and hyperparameter-tuning strategies are used to improve the interpolation performance of the machine-learning models. Additionally, various assessment methods are utilized to evaluate these models. The numerical results of motions for NPL catamarans in head-sea conditions obtained by three different approaches, namely the 3D Green function method, strip theory, and the 3D pulsating source method, are used for comparison. Comparison with these numerical results and the predictions from artificial neural network and K-nearest neighbor regression models shows that, for the present benchmark cases, the machine-learning models can display a higher level of agreement with the experimental data. It should be emphasized, however, that the present study is based on a limited experimental dataset and is therefore intended as a methodological case study rather than as a general-purpose motion prediction tool for catamaran design. The dataset is restricted to the NPL catamaran series, long-crested head seas, fixed values of the breadth-to-draft ratio and block coefficient, and only two demi-hull spacing ratios. Consequently, the developed models are only expected to interpolate within the studied parameter range and do not establish broad generalizability to other hull forms, hull coefficients, spacing configurations, or wave headings.

2. Methodology

2.1 Ship data description

The study employs experimental data about the seakeeping characteristics of several fast displacement catamaran models when subjected to long-crested head seas by Wellicome et al. (1995). Three hull configurations that were geometrically similar were evaluated, with length-to-displacement ratios ranging from 7.4 to 9.5. Based on the NPL round bilge series, the vessels are round bilge transom sterns. Various hull configurations were evaluated across a range of wavelengths and at three Froude numbers (Fn = 0.2, 0.53, and 0.8), including monohull and two catamaran demi-hull spacings. Measurements of heave, pitch, vertical acceleration, and added resistance in waves were conducted at two stations. Three hull forms were assessed in total. Molds derived from the foam vessels utilized in calm water experiments (Molland et al. 1994) were employed to fabricate the models. The primary characteristics of the monohulls are shown in Table 1.

Table 1: Hull form particulars (Monohull)

Model	4b	5b	6b
Length	1.6 m	1.6 m	2.1 m
$L/\nabla^{1/3}$	7.4	8.5	9.5
L/B	9.0	11.0	13.1
B/T	2.0	2.0	2.0
C_B	0.397	0.397	0.397
C_P	0.693	0.693	0.693

C_M	0.565	0.565	0.565
Waterplane area	0.338 m^2	0.267 m^2	0.401 m^2
LCB	-6.4%	-6.4%	-6.4%

The study encompassed experimentation on three distinct hull configurations, namely the monohull as well as the catamaran mode, with hull separation to length ratios of 0.2 and 0.4. The experimental setup involved performing measurements for every model configuration at three different Froude numbers (Fn = 0.2, 0.53, and 0.80) and a range of encounter frequencies extending from 6 rad/sec to 16 rad/sec. The calculation of the transfer functions obtained from the regular wave experiments was conducted in the following manner:

$$\text{Heave TF} = \frac{Heave\ Amplitude\ RMS}{Wave\ Amplitude\ RMS} \tag{1}$$

$$Pitch\ TF = \frac{\text{Pitch Amplitude RMS [rad]}}{Wave\ Amplitude\ RMS\ [m]} \times \frac{g[ms^{-2}]}{\omega_0^2[rads^{-1}]} \tag{2}$$

$$Acceleration\ TF = \frac{\text{Accel Amplitude RMS [ms}^{-2}]}{Wave\ Amplitude\ RMS\ [m]} \times \frac{1}{\omega_e^2[rads^{-1}]} \tag{3}$$

Where the frequency of encounter ω_e is related to the wave frequency ω_0 by Eq. (4).

$$\omega_e = \omega - \cos(\mu)\frac{\omega_0^2 u}{g} \tag{4}$$

2.2 Data preprocessing: feature selection

A key element of this research was to determine the function that would estimate the heave and pitch transfer functions based on selected hull, speed, spacing, and wavelength-related input parameters. The heave and pitch motion of a ship in head seas is influenced by various factors, including the ship's dimensions (such as length, breadth, and draft), different form coefficients (such as block coefficient and prismatic coefficient), ship speed, encountering frequency, and additional parameters like the radius of gyration and water density, according to Bhattacharya (1978). In this study, the utilization of the following parameters is employed to ascertain the motion transfer function of catamarans within the tested NPL catamaran interpolation domain.

$$[Heave\ amplitude, Pitch\ amplitude] = f(L, B, T, C_{B,} V, S, \lambda) \tag{5}$$

Where,

$L - Demihull\ length\ between\ perpendiculars\ [m]$
$B - Demihull\ maximum\ breadth\ [m]$
$T - Demihull\ draught\ [m]$
$C_B - Block\ coefficient\ (Demihull)$
$V - Ship\ speed\ [m/s]$
$Fn - Froude\ number$
$S - Seperation\ between\ catamaran\ (demihull)\ centerlines\ [m]$
$\lambda - Wavelength\ [m]$
$f - function\ for\ estimating\ the\ motion\ transfer\ function$

The non-dimensional form will be as follows:

$$[Heave\ TF, Pitch\ TF] = f(\frac{L}{B}, \frac{B}{T}, C_B, Fn, \frac{S}{L}, \frac{\lambda}{L}) \tag{6}$$

The idea of exploratory data analysis is to simply understand the relationships among the data, including the shape and spread of the data and the validity of using it for a variety of different predictive algorithms. The univariate numerical summary of the heave and pitch datasets is shown in Tables 2 and 3.

Table 2: Numerical summary of Heave dataset

Index	L/B	B/T	C_B	Fn	S/L	λ/L	Heave TF
Data count	222	222	222	222	222	222	222
Mean	11.11	2	0.397	0.5	0.3	1.43	0.76
Standard deviation	1.68	0	0	0.23	0.1	0.82	0.54
Min	9	2	0.397	0.2	0.2	0.38	0.02
25%	9	2	0.397	0.2	0.2	0.89	0.24
50%	11	2	0.397	0.53	0.4	1.25	0.79
75%	13.1	2	0.397	0.8	0.4	1.73	1.09
Max	13.1	2	0.397	0.8	0.4	5.68	2.02

Table 3: Numerical summary of the pitch dataset

Index	L/B	B/T	C_B	Fn	S/L	λ/L	Pitch TF
Data count	221	221	221	221	221	221	221
Mean	11.1	2	0.397	0.5	0.3	1.44	0.73
Standard deviation	1.68	0	0	0.23	0.1	0.83	0.45
Min	9	2	0.397	0.2	0.2	0.38	0.01
25%	9	2	0.397	0.2	0.2	0.89	0.32
50%	11	2	0.397	0.53	0.4	1.25	0.77
75%	13.1	2	0.397	0.8	0.4	1.81	1.13
Max	13.1	2	0.397	0.8	0.4	5.72	1.66

Feature selection is a commonly employed data analysis approach applied to identify the most important features within a given dataset. Therefore, the reduction in data dimensionality and computational cost of modeling leads to an enhancement in model performance. Based on the summary, it is evident that the breadth to draft ratio and the block coefficient (demi-hull) exhibited no changes across the whole dataset. Therefore, it is possible to exclude this factor when training the models. However, this also implies an important limitation of the present study: the resulting models cannot learn or assess the influence of these key hull-form parameters on catamaran motion. Similarly, since the dataset contains only two demi-hull spacing ratios and is limited to long-crested head seas, the trained models should be interpreted only within this constrained experimental domain. In other words, the present models are intended for interpolation within the tested NPL series rather than for general prediction across broader catamaran design spaces. A correlation analysis is carried out to demonstrate the relationships between the independent variable, motion transfer function, and the other retained input variables. The present research employed Pearson's correlation coefficient, mathematically represented by Eq. (7).

$$r_{x,y} = \frac{\sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{7}$$

Where,

$r_{x,y} = Correlation\ between\ x\ and\ y\ variables,$

$x_i = i^{th}\ value\ of\ x,$

$y_i = i^{th}\ value\ of\ y,$

$\bar{x} = average\ of\ x,$

$\bar{y} = average\ of\ y$

The correlation coefficient r changes from -1 to +1. A negative correlation signifies that an increase in one variable is accompanied by a decrease in the other, while positive correlations suggest that an increase in one variable results in an increase in the other. As the values get further from zero and closer to -1 and 1, those changes become more predictable with less error or variation. Pearson’s correlation of the heave and pitch transfer functions to other input variables is shown in Tables 4 and 5.

Table 4: Pearson's correlation of heave transfer function to other input variables

	L/B	Fn	S/L	λ/L	Heave TF
L/B	1.000	0.112	-0.133	-0.228	-0.062
Fn	0.112	1.000	0.057	0.343	0.638
S/L	-0.133	0.057	1.000	0.050	-0.028
λ/L	-0.228	0.343	0.050	1.000	0.514
Heave TF	-0.062	0.638	-0.028	0.514	1.000

Table 5: Pearson's correlation of pitch transfer function to other input variables

	L/B	Fn	S/L	λ/L	Pitch TF
L/B	1.000	0.111	-0.117	-0.248	-0.119
Fn	0.111	1.000	0.056	0.333	0.395
S/L	-0.117	0.056	1.000	0.052	0.007
λ/L	-0.248	0.333	0.052	1.000	0.722
Pitch TF	-0.119	0.395	0.007	0.722	1.000

However, it's possible that a non-linear relationship exists but goes unnoticed. Consequently, distance correlation is used to examine the non-linear correlation. This measure works equally well with linear and non-linear information. It benefits from not assuming the input vectors are normally distributed, and it is less affected by outliers. Correlation levels are reported between 0 and 2, with 0 indicating perfect correlation and 2 indicating complete negative correlation. The correlation between the input variable u and the output variable v is defined as

$$Cor_D(u,v) = 1 - \frac{(u-\bar{u}).(v-\bar{v})}{\sqrt{(u-\bar{u})^2.(v-\bar{v})^2}} \quad (8)$$

The distance correlation of the heave and pitch transfer functions with other input variables is presented in Table 6.

Table 6: Distance correlation of heave and pitch transfer function to other input variables

	Heave TF	Pitch TF
L/B	1.062	1.119
Fr	0.362	0.605
S/L	1.028	0.993
λ/L	0.486	0.278

The analysis of Tables 4 and 5 reveals that the ratio of wavelength to ship length, together with the Froude number, has a significant influence on the transfer functions of heave and pitch. Similar results are also observed for distance correlation in Table 6. This phenomenon is due to the fact that the ship's reaction to waves is dependent upon the wavelength in relation to the ship's length. It is feasible for a vessel to traverse waves with minimal disruption when the wavelength substantially surpasses the length of the vessel. The ship may, however, undergo significant pitching and heaving when confronted with the waves if the wavelength is shorter than or equal to its length. Shorter wavelengths can result in more pronounced ship motion. Moreover, higher speeds in head waves can lead to more significant heaving and pitching motions due to the increased frequency of wave encounters and the potential for more forceful impacts with wave crests. Therefore, Froude number remains an important retained input variable for interpolating the motion response within the present dataset. On the contrary, within the present dataset, the ratio of demi-hull spacing to ship length appears to show limited correlation with the heave and pitch transfer functions. This can be attributed to the limited scope of the dataset, which solely encompasses experimental findings from two distinct demi-hull separation configurations in head waves. The scatter plot depicted in Figure 1 demonstrates the correlation between the input parameters and the heave and pitch transfer functions.

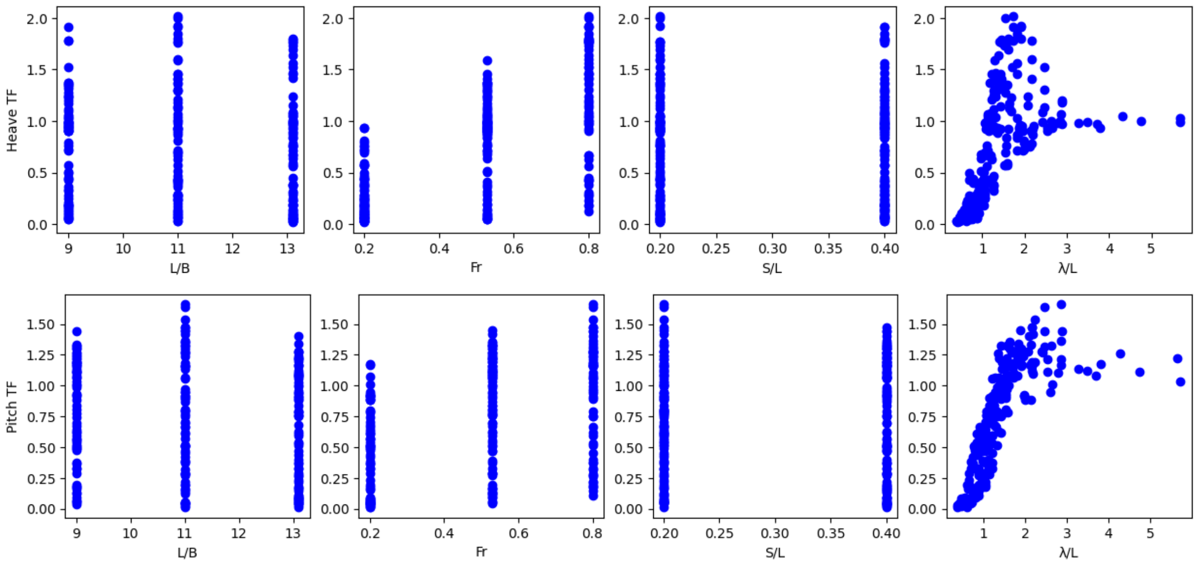


Figure 1: Scatter plots of heave and pitch transfer functions versus initial ship design parameters.

Since various early ship design parameters have distinct ranges, scaling features is an essential preprocessing phase. As a result, discrete characteristics can be contrasted and make equivalent contributions to the results of machine learning. Standardization and normalization are widely recognized scaling techniques, and this study employs both to investigate the impact they have on a range of regression models. The process of standardization involves transforming the feature data so that each value is centered on the mean and has a standard deviation of one unit. This causes the attribute's mean to be equal to zero, resulting in a distribution with a standard deviation of one as shown in Eq. (9). In the meantime, normalization adjusts the feature data from 0 to 1, which is also known as min-max scaling, as shown in Eq. (10).

$$x_s = \frac{x-\mu}{\sigma} \tag{9}$$

$$x_m = \frac{x-x_{min}}{x_{max}-x_{min}} \tag{10}$$

Where,

x_s = *Dimensionless standardized values of actual variable* x
x_m = *Dimensionless normalized values of actual variable* x
μ = *Mean value of variable* x
σ = *Standard deviation of variable* x
x_{min} = *Minimum value of variable* x
x_{max} = *Maximum value of variable* x

2.3 Prediction models

2.3.1 Polynomial regression (PR) model

Polynomial regression is a modified form of linear regression that incorporates feature interaction between independent variables and assumes a polynomial relationship between response (y_i) and predictor variables (x_{i1}, x_{i2}). Eq. (11) shows the feature interaction for a second order model where the number of independent variables is 2.

$$y_i = w_0 + w_1x_{i1} + w_2x_{i2} + w_3x_{i1}x_{i2} + w_4x_{i1}^2 + w_5x_{i2}^2 \tag{11}$$

Here, w_0 is the constant terms in the model and w_1 to w_5 represents the corresponding coefficients of the interacted predictor features, which are found by minimizing the error function. PR is one of the most common parametric models due to its simplicity and ease of comprehension. By adjusting the degree of feature interaction, the performance of the model can be improved. Various degrees of polynomials are thus examined in an effort to determine a suitable model configuration for the present dataset.

2.3.2 K – nearest neighbors (KNN) regression model

K-nearest neighbors (KNN) regression is a non-parametric methodology used to estimate the relation between independent variables and a continuous output. This approach does this by aggregating data from the local neighborhood. The data points are plotted on a multidimensional hypersurface, and Euclidean distance is utilized to calculate the distance between them. Euclidean distance can also be represented as the length of the straight line connecting the two points under examination, as shown in Eq. (12). (Pedregosa et al. 2011)

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2} \tag{12}$$

Cross-validation is used to determine the neighborhood size that minimizes the error.

2.3.3 Regression tree

The regression tree is a machine learning approach distinguished by its use of a hierarchical arrangement of decisions within a structure of branches and nodes. At every node in the structure, a decision is made based on the record's qualities. The anticipated result is evident at the termination of every branch of the tree. The execution of this approach is characterized by its simplicity and user-friendliness. A typical decision tree configuration is shown in Figure 2 (Pedregosa et al., 2011).

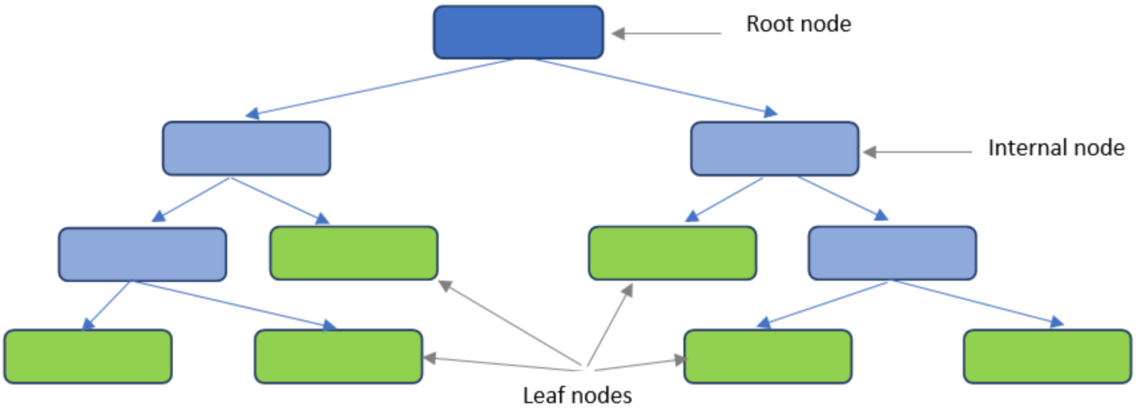


Figure 2: A regression tree configuration (Pedregosa et al., 2011)

In general, regression trees comprise two types of nodes. It consists of decision nodes (i.e., the root node and internal nodes) and leaf nodes. The decision nodes contain a condition for splitting the data, while the leaf nodes determine the value of each data point. The data is divided recursively using a binary tree until only pure leaf nodes remain. The division is performed so that the impurity of the nodes continues to decrease. For this, the variance reduction is computed. As a measure of impurity, variance is employed within the context of regression (Eq. 13)

$$Variance = \frac{1}{n}\sum(y_i - \bar{y})^2 \tag{13}$$

A higher variance means greater impurity. The variance reduction is the combined variance of the child node minus that of the parent node, as shown in Eq. (14).

$$Variance\ Reduction = Var_{(Parent)} - \sum w_i Var_{(Child_i)} \tag{14}$$

The weights w_i are the relative size of the child with respect to the parent. The split with the highest variance reduction is taken. The model assesses the variance reduction associated with each potential split and determines the optimal one. This process of selection happens recursively unless the model has reached its desired depth. Its precision, adaptability, and robustness are all determined by the depth of the tree. While a more refined regression tree may produce more accurate results, it also elevates the probability of overfitting, which could result in less precise predictions outside the training dataset. In contrast, a coarse regression tree exhibits greater robustness for out-of-sample predictions despite its reduced training accuracy. To determine the prediction, the mean value of all data elements present in the leaf node is calculated.

2.3.4 Random forest

Unless there are pure leaf nodes, the regression tree separates the dataset recursively. It maximizes variance reduction to identify the best split. If the training data is significantly changed by trading a data feature, the regression tree will be radically different. Thus, regression trees are sensitive to training data, which may increase variance. As a result, the model may not generalize. Random forest algorithm is used to mitigate this problem. Random forest is referred to as an ensemble method, which means a group of items working together for a specific purpose. It consists of multiple random decision trees and is significantly less sensitive to training data. The rows are randomly selected from the original data to build the new datasets. Every dataset will contain the same number of rows as the original one. This is called random sampling. Here, random sampling with replacement is used. So, one row can occur more than once in the new datasets. Which means that after selecting a row, it is put back into the dataset. The process of creating new data is called bootstrapping. After that, a decision tree is trained on each of the bootstrapped datasets independently. Although a single decision tree might be overfitted to the training data, this tendency is reduced when the overfitted trees work as a whole. Bootstrapping ensures that the same data is not used for every tree. So, in a way, it helps the model be less sensitive to the original data. So, the final prediction is much more robust and has a lot less variance than a single decision tree. The predictions of different trees $(\hat{f}^i(x))$ are then averaged as follows.

$$f_{random\ sampling}(x) = \frac{1}{n}\sum_{i=1}^{n}\hat{f}^i(x) \tag{15}$$

Not every feature is utilized to train the trees. A random subset of features is chosen for each tree, and only those features are utilized for training. This method is known as random feature selection. The random selection of features decreases the correlation between the trees. If every feature is utilized, then most trees will contain identical decision nodes. And they will behave identically. This will increase the variability. There is an additional advantage to the random feature selection. Here, some of the trees will be trained on features of lesser importance. Therefore, they have a propensity to make inaccurate predictions. However, some trees will provide inaccurate predictions in the opposite direction. Therefore, they become balanced. The number of trees, the number of features to consider in each split, the maximum depth of trees, and the numbers of samples are determined by cross-validation, which is explained in section 6. (Pedregosa et al. 2011)

2.3.5 Artificial neural network

Neural networks form the basis of deep learning, a subfield of machine learning whose algorithms draw inspiration from the structure of the human brain. Neural networks receive data, undergo self-training to identify patterns within the data, and subsequently generate predictions for a subsequent data set that possesses comparable patterns. Neural networks consist of neuron layers. The fundamental processing elements of the neural networks are these neurons. An input layer is the initial component that receives the input. The ultimate output is predicted by the output layer. Between these are the hidden layers, which are responsible for carrying out most of the network's computations. The structure of a neural network with two hidden layers is depicted in Figure 3.

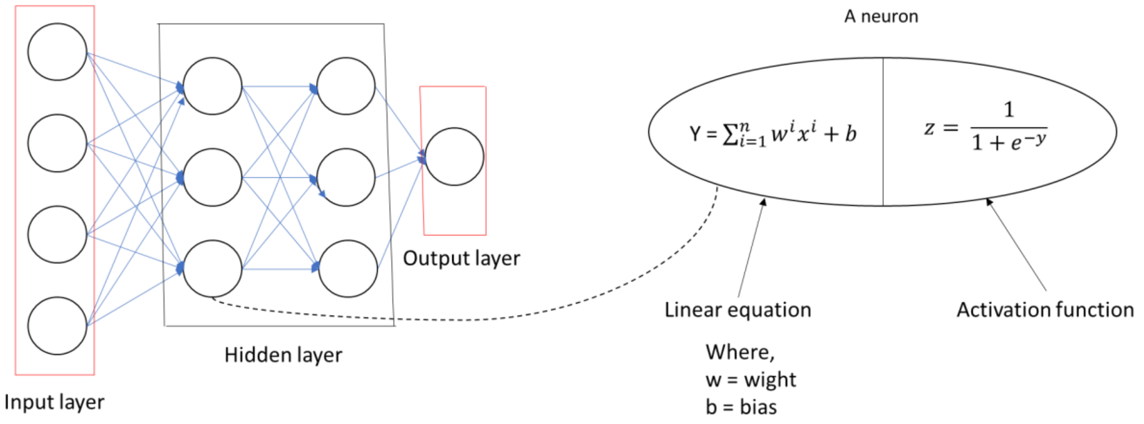


Figure 3: General structure of a neural network.

In the input layer, every neuron receives each individual input. Neurons within a given layer are interconnected with neurons within the subsequent layer. The numerical value assigned to each of these connections is denoted by weight (w). The weights that are associated with the inputs are multiplied by them. And their sum is transmitted as input to the neurons in the hidden layer. The bias (b), a numeric value assigned to each of these neurons, is incorporated into the total of the inputs. This value is subsequently passed on via the activation function. Whether or not a neuron is activated is determined by the result of the activation function. An active neuron transfers information to the neurons of the subsequent layer via the connections. This is how data propagates throughout the network. The term for this is forward propagation. The output value is determined by the output neuron in the output layer. To start the neural network, certain initial values are assumed for the weights and biases (Chollet et al. 2015).

Both the input as well as the output are provided to the network throughout the training procedure. By comparing the expected output to the actual output, the prediction error can be ascertained. The magnitude of the error demonstrates how inadequate the existing network is. Moreover, whether the predicted values are greater or lower than anticipated is indicated by the sign assigned to the error. Then, this information is sent backwards through the network. The term for this phenomenon is reverse propagation. The weights are adjusted in accordance with this information. Iteratively, the process of forward and reverse propagation is carried out using several inputs. This procedure continues if the weights are assigned in a manner that guarantees the network a satisfactory level of accuracy in its predictions.

Activation functions play a crucial role in artificial neural networks, as they are one of the most significant components. The output of a neuron is determined by the activation function applied to its input. Different activation functions can be chosen based on the specific circumstances of a neuron, such as whether it is in the hidden layer or the output layer. Additionally, it is important to consider the specific problem that the artificial neural network is designed to address. The following are examples of commonly utilized activation functions in artificial neural networks.

i. Sigmoid:

$$sigmoid(z) = \frac{1}{1+e^{-z}} \tag{16}$$

ii. Tanh(x):

$$\tanh(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}} \tag{17}$$

iii. Rectified Linear Unit (ReLU):

$$ReLU(z) = \max(0, x) \tag{18}$$

iv. Leaky ReLU:

$$Leaky\ ReLU(z) = \max(0.1x, x) \tag{19}$$

A loss function, which is alternatively referred to as a cost function or as an error function, is employed to quantify the discrepancy between the algorithm's output and the specified target value. Loss function minimization is the goal of the learning procedure with respect to the determination of optimal weights and biases. As a result, various learning methods and neural network architectures can be employed to select an appropriate network architecture for the present dataset.

2.4 Prediction model evaluation

Building a statistical model capable of accurately predicting the result of an unobserved case is the primary aim of regression analysis. Post-model construction, a regression evaluation must be performed so that this objective can be fulfilled. Methods for evaluating models can be broadly classified into two categories: (1) utilizing the same dataset for both training and testing purposes; and (2) dividing the dataset into distinct subsets, one for model training and the other for post-training performance assessment.

The first evaluation approach involves training and testing on identical datasets, hence constituting a direct and uncomplicated procedure. The name of this method well reflects its essence. In this method, the model is trained on the whole dataset and then evaluated on a subset of the same dataset. Since the model has previously been exposed to all the testing data points during training, it seems likely that this assessment procedure will exhibit elevated training accuracy but decreased out-of-sample accuracy. Conversely, the second assessment methodology entails partitioning the dataset into distinct training and testing subsets that do not overlap. By implementing this approach, the precision of the model can be improved when it is utilized on data that was not incorporated during the training phase. The building of the model is derived from the training set. Afterwards, the collection of test features is sent to the model for the purpose of generating predictions. In the end, a comparison is made between the predicted and actual values of the test dataset. Furthermore, different measures of performance can be employed to assess and compare the efficacy of the model, as explained in the following subsequent sections.

2.4.1 K – fold cross-validation

A notable constraint of the train/test split methodology is its substantial dependence on the datasets employed for both data training and testing. The presence of variance in the data leads to a more accurate in-domain test performance when employing a train/test split, as compared to training, and testing on the same dataset. However, this approach is not without its limitations, as it is still influenced by the underlying dependency. The resolution of this issue is attained through the implementation of K-fold cross-validation, an alternative evaluation method. In its most fundamental form, K-fold cross-validation performs multiple iterations of train/test divides on a single dataset, ensuring that each split is unique. Subsequently, the outcome is subjected to averaging to generate a more uniform measure of estimated in-domain accuracy for held-out samples from the dataset. The dataset is split into K equal-sized groups, or folds. (K-1) of the groups are used to train the interpolation models, and the last group is saved for confirmation. As stated above, the process is run K times, with a different one of the K folds used as the confirmation group each time. The accuracy and computational complexity of a model can be influenced by varying values of K. This study considers a K value of 10 as an appropriate choice for balancing the estimated in-domain interpolation accuracy of the models with their computational complexity. In this study, since the dataset is structured and sparsely varied, random train/test splitting and K-fold cross-validation may place closely related NPL cases in both training and testing folds; therefore, the reported scores may overestimate performance for unseen hull forms, additional spacing ratios, different wave headings, or broader operating conditions.

2.4.2 Mean Absolute Error (MAE)

The metric under consideration is the arithmetic mean of the absolute values of the errors, which can be computed using the following formula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{20}$$

In this context, y_i denotes the actual output variable, while $\hat{y}_i$ represents the predicted value of the output variable generated by the trained model. Additionally, n refers to the total number of samples in the dataset.

2.4.3 Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error is frequently employed as a loss function for regression tasks and in the assessment of models due to its straightforward interpretation in regard to relative error.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i - \hat{y}_i}{y_i} \tag{21}$$

2.4.4 Mean Squared Error (MSE)

The index quantifies the average of the squared error. Mean squared error (MSE) is a more widely embraced metric compared to mean absolute error (MAE) due to its emphasis on capturing substantial mistakes. This phenomenon arises from the exponential amplification of bigger errors relative to smaller errors, which is attributable to the presence of the squared component.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{22}$$

2.4.5 Root Mean Squared Error (RMSE)

This criterion corresponds to the mathematical operation of taking the square root of the average of the squared differences between observed and predicted values. The root mean squared error (RMSE) is often regarded as a prominent assessment metric due to its interpretability in the same units as the output vector. This characteristic facilitates understanding and correlation of the information it conveys.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{23}$$

2.4.6 Coefficient of determination (R^2)

The coefficient of determination is extensively utilized within the data science field. The R-squared statistic is not considered an error in the strict sense but rather a widely used measure to assess the accuracy of a model. The metric quantifies the proximity of the data values to the regression line that has been fitted. A greater value of R-squared indicates a stronger degree of fit between the model and the observed data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{24}$$

Where, $\bar{y}_i$ is the average value of the observed output variable.

2.5 Implementation

Mathematical models are developed in a Python workspace for the interpolation models discussed in Section 4, along with the extracted motion dataset. The aim of this implementation is to assess the in-domain interpolation performance of the selected models for the extracted NPL catamaran motion dataset. Initially, experimental data is extracted and, afterwards, imported into the Python framework. The chosen parameters are then extracted, evaluated, and manipulated in accordance with the procedures outlined in the data preprocessing section. In the context of data scaling, a quantitative investigation is conducted to determine the sensitivity of the PR and KNN models to various scaling methods. In this situation, it was not essential to scale the data for the tree-based regression models, as they are inherently scale invariant. However, for the artificial neural network model, the appropriate scaling strategy is dependent upon the specific type of activation function utilized. This work employed the activation functions ReLU and linear, as well as normalizing the dataset for both training and testing purposes.

The datasets are subsequently divided into training and testing datasets to evaluate the in-domain interpolation accuracy of the models in terms of mean squared error, mean absolute error, and mean absolute percentage error. Specifically, 15% of the data was allocated for testing, while the remaining 85% was utilized for training.

However, the scores of various interpolation models may be significantly influenced by the composition of the training and testing datasets. To mitigate this problem, K-fold cross validation is employed to estimate the coefficient of determination for each of the polynomial regression, K-nearest neighbors, and tree-based machine learning models. The impact of the parameter on the performance of various models is examined using K-fold cross-validation, with K values of 5, 10, and 20. Nevertheless, employing a fold number of 10 yielded a more consistent in-domain validation result. Finally, the coefficient of determination for each model was calculated using a 10-fold cross-validation split on the entire dataset.

The analysis included evaluating the polynomial degrees and the number of closest neighbors (2, 3, and 4) in both the polynomial model and the K-nearest neighbor regression. For tree-based regression models, various configurations and sizes of regression trees are considered, and grid search is used to identify suitable hyperparameter combinations. The selected hyperparameter combinations are chosen based on their validation performance within the present dataset. The optimized hyperparameters include the maximum depth of the decision tree, the maximum number of features to evaluate at each split, the minimum number of samples necessary to initiate a split, and the number of trees inside a random forest.

The selected neural-network configuration is determined by evaluating several candidate network structures and learning settings within the present dataset. The phenomenon of overfitting is the primary challenge of neural network learning. Typically, a test set method is used to detect and prevent this phenomenon. So, 15% of the total dataset is separated as a test dataset, and the rest is used for training and validation. Optimizers define how neural networks learn. They find the values of parameters such that the loss function is lowest. There are different types of optimizers. In this study, an adaptive moment estimation (Adam) optimizer is used, where the learning rate of each parameter varies based on historical gradients and momentum.

2.6 Comparison of numerical results with the predicted results

Three different numerical results of motions for the NPL catamaran hull have been used for the comparison (Moinul, 2021). In the 3D Green function method, which is widely used for seakeeping problems, the underwater hull geometry of the vessel has been discretized in quadrilateral elements, and the 3D Green function in the form used by Inuoe and Makino (1989) has been used for the calculation. Quick and handy strip theory has also been used for motion prediction. Similar to the strip theory, the 3D pulsating source frequency domain Green function method, which includes the influence of forward speed through corrections to the zero-speed solution, has also been used for numerical calculation of the NPL catamaran hull.

3. Results and Discussion

Table 7: Performance measure for different machine learning approaches

		Heave transfer function					Pitch transfer function			
		R^2	MSE	MAE	MAPE		R^2	MSE	MAE	MAPE
Polynomial regression										
	Raw data	0.461	0.225	0.406	0.875		0.709	0.036	0.146	0.206
	Normalized data	0.459	0.154	0.316	0.511		0.706	0.032	0.149	0.476
	Standardized data	0.461	0.141	0.289	0.649		0.709	0.041	0.16	0.443
Degree 3	Raw data	-0.181	0.126	0.289	0.934		0.877	0.058	0.127	0.336
	Normalized data	0.084	0.081	0.239	0.931		0.836	0.815	0.389	1.238
	Standardized data	0.511	0.074	0.203	0.7		0.843	0.01	0.08	0.357

Degree 4	Raw data	-3.601	0.052	0.184	0.611		0.372	0.058	0.115	0.679
	Normalized data	-3.633	4.101	0.506	1.036		0.185	0.051	0.135	0.438
	Standardized data	-3.374	0.035	0.157	0.626		0.4	8.02	0.766	2.299
KNN regression										
k = 2	Raw data	0.936	0.022	0.11	0.3		0.941	0.007	0.061	0.155
	Normalized data	0.885	0.027	0.117	0.362		0.934	0.011	0.075	0.165
	Standardized data	0.891	0.024	0.108	0.352		0.935	0.009	0.067	0.158
k = 3	Raw data	0.907	0.019	0.114	0.315		0.941	0.007	0.068	0.166
	Normalized data	0.784	0.047	0.152	0.402		0.884	0.015	0.098	0.215
	Standardized data	0.815	0.041	0.137	0.386		0.885	0.013	0.089	0.208
k = 4	Raw data	0.89	0.03	0.143	0.349		0.937	0.008	0.066	0.163
	Normalized data	0.68	0.075	0.195	0.484		0.804	0.025	0.121	0.281
	Standardized data	0.717	0.064	0.114	0.463		0.813	0.022	0.114	0.275
Regression tree										
	Raw data	0.9	0.003	0.032	0.143		0.875	0.014	0.089	0.131
Random forest										
	Raw data	0.919	0.003	0.041	0.216		0.916	0.001	0.029	0.18
ANN										
	Normalized data	0.98	0.006	0.059	0.233		0.97	0.006	0.054	0.186

According to the findings presented in Table 7, it can be observed that among the many machine learning models examined, the K-nearest neighbors (KNN) model with a value of K set to 2, when applied to the raw dataset, yields the most favorable outcomes in terms of cross-validation scores. This is shown by a 10-fold cross-validation score of 0.936 and 0.941, as well as a mean squared error (MSE) of 0.022 and 0.007, respectively, for the prediction of the heave and pitch transfer functions. The strong performance of the K-nearest neighbor model, particularly at low neighborhood size, suggests that the method is highly effective for local interpolation within the present dataset. At the same time, such behavior indicates that the model may be relying strongly on similarity to nearby samples, and therefore its strong accuracy should not be interpreted as proof of a broadly learned hydrodynamic relationship.

The analysis reveals that the polynomial regression model achieved a notable R^2 value of 0.843 when used for pitch prediction, employing a regression degree of 3 on a standardized dataset. Nevertheless, in certain instances, the R^2 score obtained through cross-validation yielded a negative value. This phenomenon can occur when the model prediction is worse than simply using the mean value of the response variable. In the present study, these large negative values should not be interpreted merely as isolated numerical outcomes. Rather, they indicate the vulnerability of higher-degree global parametric models to high variance and overfitting when applied to a small and sparsely varying dataset. Several input variables are either constant or highly restricted, including the breadth-to-draft ratio, block coefficient, and the availability of only two demi-hull spacing ratios. As a result, higher-degree polynomial terms can introduce unstable global feature interactions that are not sufficiently supported by the available experimental feature space. The poor polynomial-regression performance therefore reflects a mismatch between model complexity and the restricted structure of the dataset.

The hyperparameters of tree-based regression models are optimized through grid search. This is necessary because fine trees have a propensity to overfit, while coarse trees tend to underfit. The hyperparameter tuning process involves adjusting the maximum number of levels in the tree, the number of features considered at each split, the minimum number of samples required to split a node, and the number of trees in the random forest, as shown in Table 8.

Table 8: List of hyperparameters tuned in regression trees and random forests.

Parameter	Experimented value
Maximum number of levels	2, 4, 8, 10, None
Maximum number of features considered at each split	1, 2, 4
Minimum number of samples required to split a node	2, 5
Number of trees in the random forest	20, 60, 100, 120

During the grid search procedure, all possible combinations of hyperparameters are evaluated using a 10-fold cross-validation method. In the experimentation phase, various hyperparameter values were tested for the purpose of predicting the heave transfer function using a regression tree. The optimal model was chosen from a pool of 30 candidates, each of which underwent 10-fold cross-validation, resulting in a total of 300 evaluations. The selection criterion had the highest cross-validation score. The optimal configuration was found to be a tree with 8 levels, considering a maximum of 4 features at each node and requiring a minimum of 2 samples to split a node.

In the prediction of the pitch transfer function, it was determined that the optimal depth for the model was 4. Additionally, during the construction of the decision tree, the maximum number of features examined at each node was set to 4, while the minimum number of samples required at each node was set to 2. The obtained R^2 scores for heaving and pitching configurations were 0.9 and 0.875, respectively. These scores indicate close alignment between the training and test results for the present split; however, given the limited dataset, this should not be taken as conclusive evidence that overfitting has been avoided. Subsequently, the models underwent evaluation using the test dataset, and the outcomes were evaluated in terms of mean squared error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The predicted values for the test dataset and their deviation from the observed values are visually depicted in Figure 5. Similar hyperparameter tuning was done for the random forest regression to reduce the tendency for overfitting. A total of 1200 fits were performed on 120 candidates, each of which underwent 10 folds. The combination of parameters was chosen based on the cross-validation score that yielded the highest value. The optimal configuration for the heave transfer function was determined to be 20 estimators. The maximum number of levels was set to "none," indicating that the trees can have as many levels as needed. At each split, a maximum of 4 features were considered, and a minimum of 2 samples were required to split a node. The model has high accuracy, with an R^2 value of 0.919 using this configuration. The ideal set of parameters for the pitch transfer function consisted of 100 estimators, a maximum depth of 8 for the estimators, a maximum number of features of 2, and a minimum number of samples of 2 in each split, which gives an R^2 value of 0.916. The regression tree and random forest models are also tested with a test dataset, and the test results are shown in terms of MSE, MAE, and MAPE in Table 7.

Artificial neural network models were employed to train and evaluate the prediction of heave and pitch transfer functions. Initially, the dataset was partitioned into training and testing datasets, with a split ratio of 85% for the training dataset and 15% for the testing dataset. The test datasets consisted of 35 data points, whereas the train datasets consisted of 187 and 186 data points for the heave and pitch transfer functions, respectively. The datasets underwent normalization for the purpose of training. In both models, a subset representing 15% of the training datasets was selected for the purpose of validation. The loss function used in the study was the mean squared error, and the Adam optimizer was used for training purposes. The structure of the neural network models is shown in Table 9.

Table 9: Artificial neural network model summary

Layer (type)	Output shape	Number of parameters

Hidden layer 1 (dense)	(None, 60)	300
Hidden layer 2 (dense)	(None, 45)	2745
Hidden layer 3 (dense)	(None, 30)	1380
Hidden layer 4 (dense)	(None, 20)	620
Hidden layer 5 (dense)	(None, 15)	315
Hidden layer 6 (dense)	(None, 10)	160
Output layer (dense)	(None, 1)	11
Total number of parameters		5531
Total number of trainable parameters		5531

The neural network models are comprised of six hidden layers, as well as one input layer and one output layer. In this study, a densely connected neural network architecture is employed, wherein every neuron in one layer is connected to every neuron in the next. Table 9 demonstrates that there are 60 neurons in the first hidden layer, 45 in the second, 30 in the third, 20 in the fourth, 15 in the fifth, and 10 in the sixth, with an input dimension of 4. The selected configuration involves training 5531 weights and biases. Each of the hidden layers utilizes the rectified linear unit (ReLU) activation function, while the output layer employs a linear activation function. The number of epochs for both training processes was 200. The training loss and validation loss were recorded after every epoch and are shown in Figure 4**.**

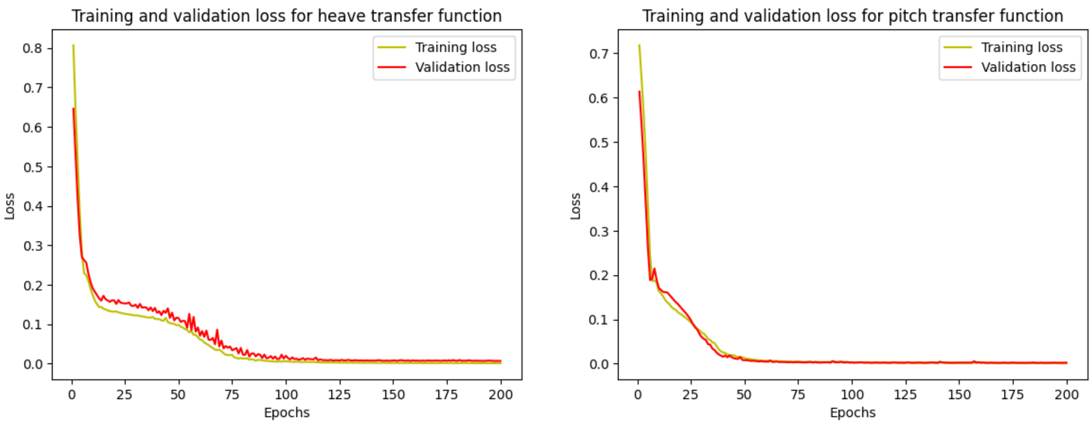
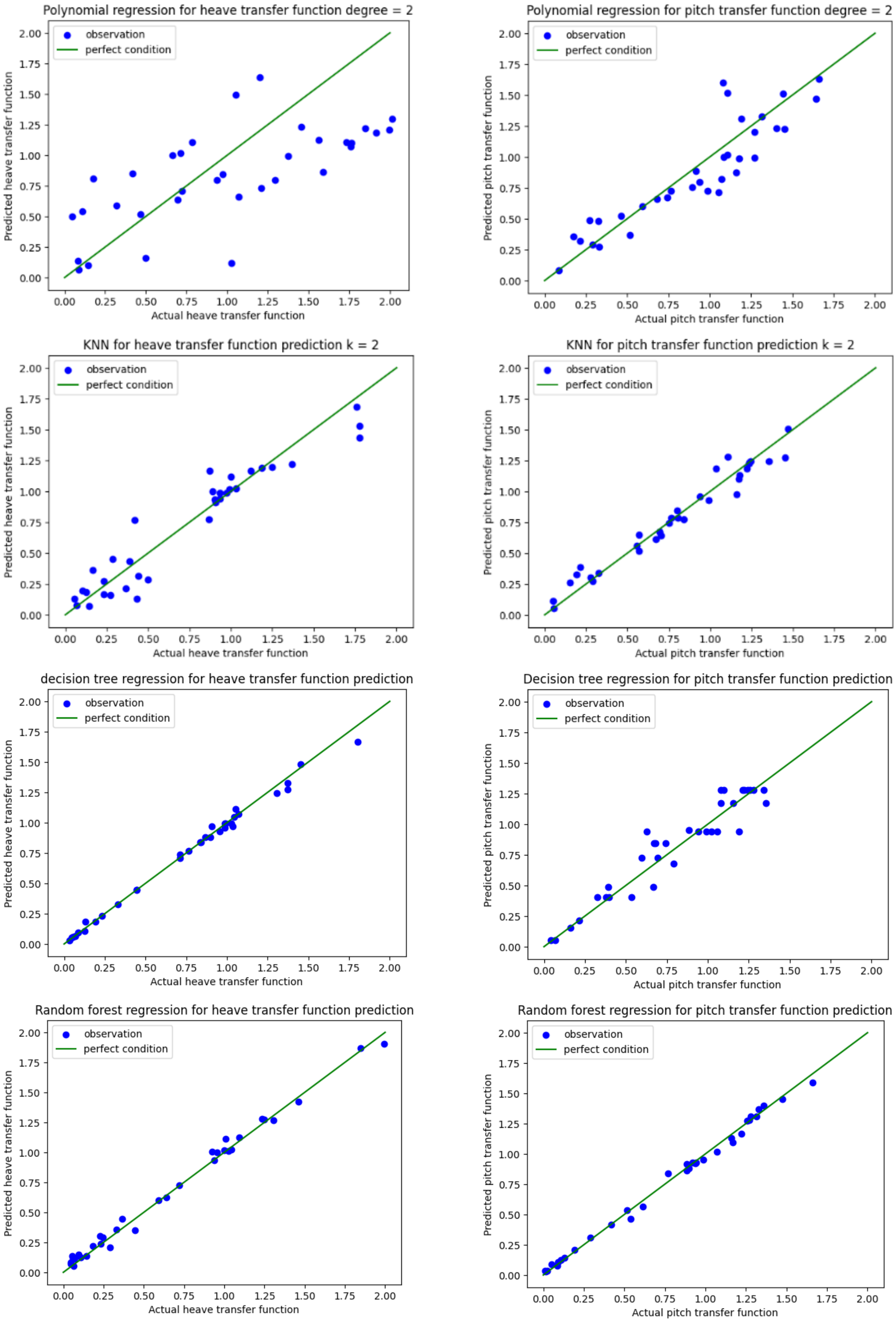


Figure 4: Training and validation loss record

The data reveals a clear trend of decreasing loss as the number of epochs increases. Specifically, at 200 epochs, the loss curve approaches a horizontal position relative to the x-axis, indicating that the loss has reached a stable value and further learning can be terminated. Following the completion of the training process, the models were evaluated using the test dataset. The resulting R^2 scores for the test dataset were found to be 0.98 and 0.97 for the prediction of heave and pitch transfer functions, respectively. The predicted results are also shown in Figure 5. This shows a good agreement with the accuracy of the models on train datasets, where the R^2 scores were 0.99 for both models. Although the train and test scores are close for the present split, this alone does not rule out overfitting given the small dataset and the complexity of the network. The ANN results should therefore not be interpreted as evidence of robust generalization, and practical use beyond the present dataset would require independent validation on additional hull forms and operating conditions, as well as comparison with simpler or regularized neural-network architectures.


Polynomial regression for heave transfer function degree = 2
Polynomial regression for pitch transfer function degree = 2
KNN for heave transfer function prediction k = 2
KNN for pitch transfer function prediction k = 2
decision tree regression for heave transfer function prediction
Decision tree regression for pitch transfer function prediction
Random forest regression for heave transfer function prediction
Random forest regression for pitch transfer function prediction
observation
perfect condition
Predicted heave transfer function
Actual heave transfer function
Predicted pitch transfer function
Actual pitch transfer function
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

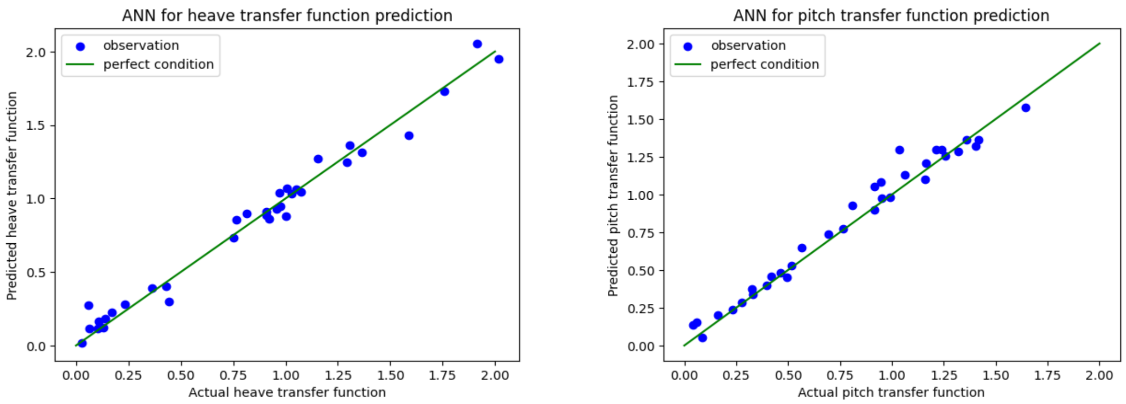


Figure 5: Predicted transfer function versus actual transfer function measurements for the test dataset for different approaches using raw data.

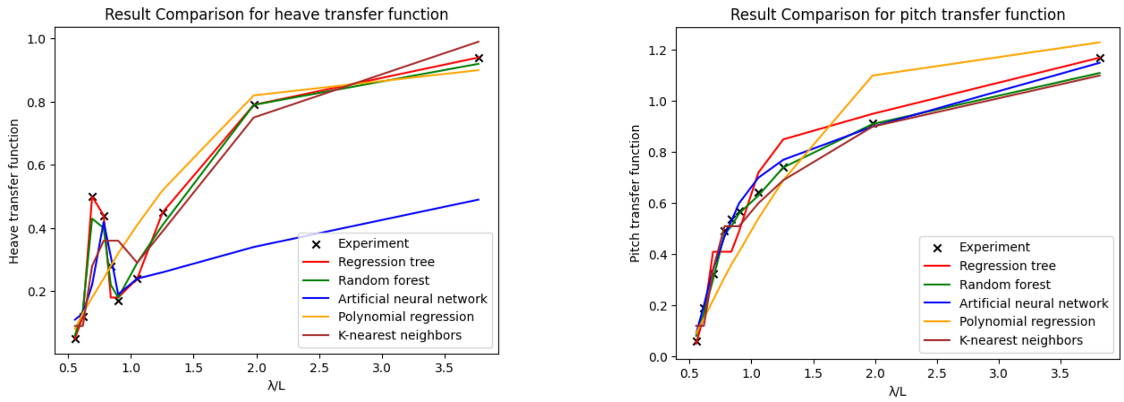


Figure 6: Experimental and predicted transfer functions from different models for NPL 4b, Fn = 0.2, S/L = 0.2

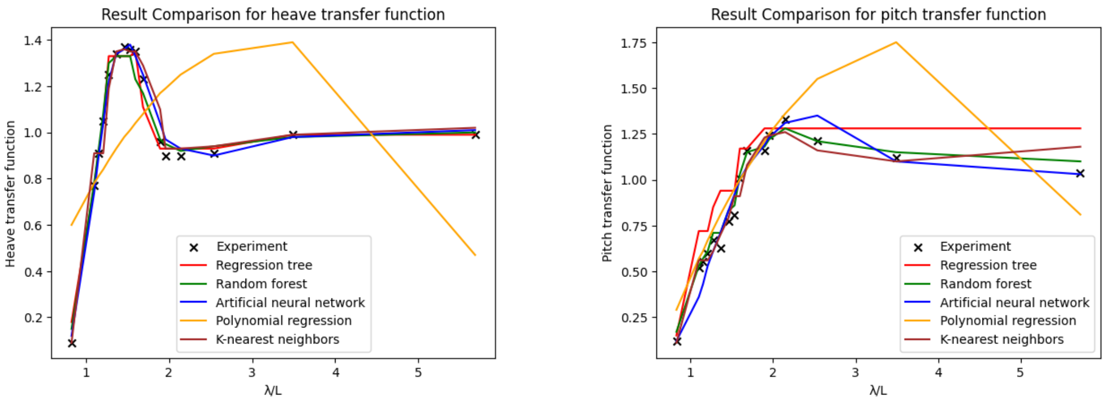


Figure 7: Experimental and predicted transfer function from different models for NPL 4b, Fn = 0.53, S/L = 0.2

Figures 6 – 22 compare the actual and predicted values of the motion transfer function for several machine learning models to show how well the suggested methodology performs. In this case, the transfer function for heave and pitch was estimated for a catamaran with NPL 4b, 5b, and 6b demi-hulls for different Froude numbers and different demi-hull spacings. The primary objective of this comparison was to evaluate how the models estimate the transfer function over different wavelength-to-ship-length ratios within the available NPL dataset. The figures indicate that the K-nearest neighbor, random forest, and artificial neural network models show closer in-domain agreement with the experimental transfer functions within the available NPL dataset. On the other hand, it can be observed from Figures 6 – 22 that the polynomial regression model yields the least accurate estimation. The models utilized in this context are the same as the models demonstrated in Figure 5.

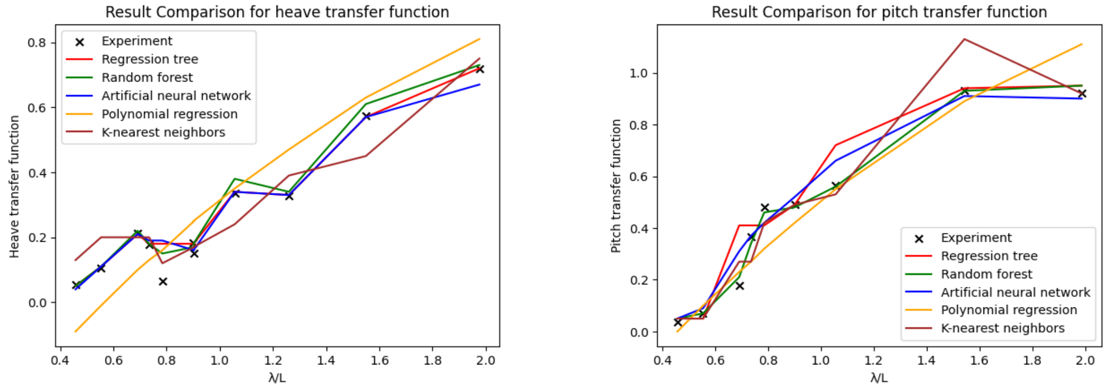


Figure 8: Experimental and predicted transfer function from different models for NPL 4b, Fn = 0.2, S/L = 0.4

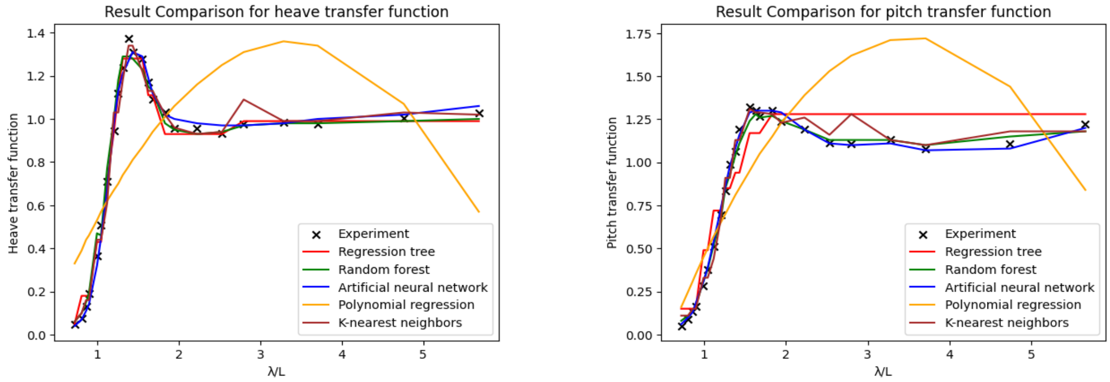


Figure 9: Experimental and predicted transfer function from different models for NPL 4b, Fn = 0.53, S/L = 0.4

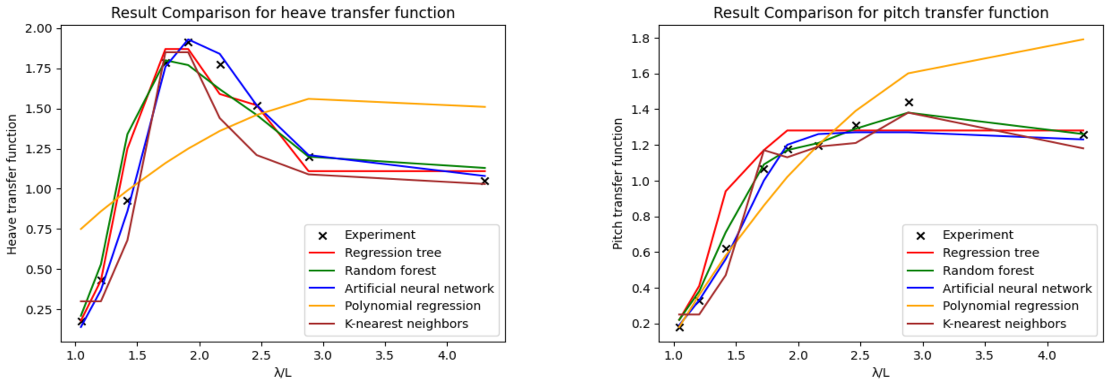


Figure 10: Experimental and predicted transfer function from different models for NPL 4b, Fn = 0.8, S/L = 0.4

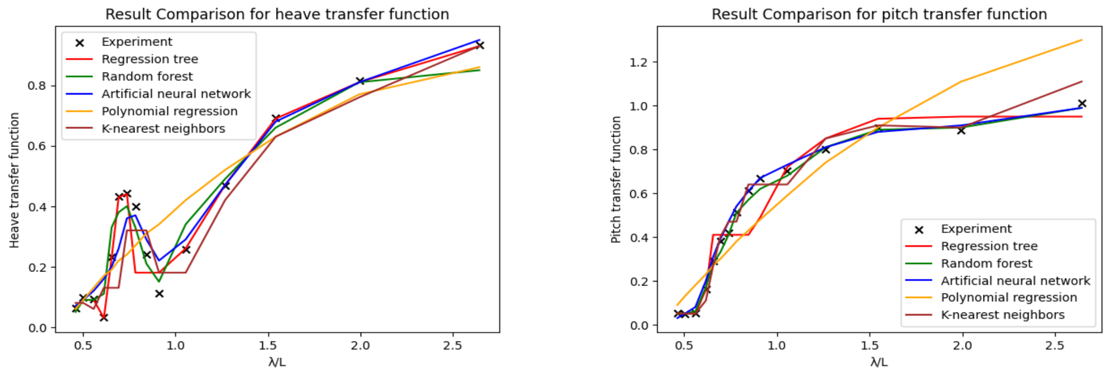


Figure 11: Experimental and predicted transfer function from different models for NPL 5b, Fn = 0.2, S/L = 0.2

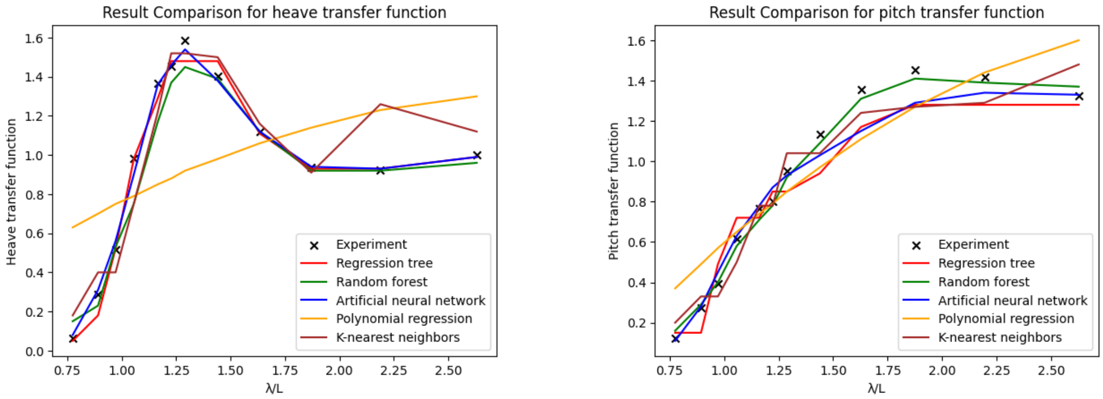


Figure 12: Experimental and predicted transfer function from different models for NPL 5b, Fn = 0.53, S/L = 0.2

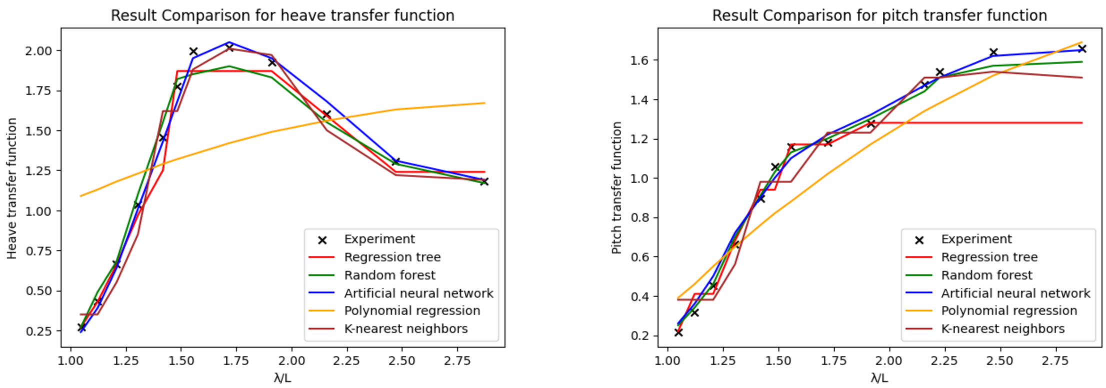


Figure 13: Experimental and predicted transfer function from different models for NPL 5b, Fn = 0.8, S/L = 0.2

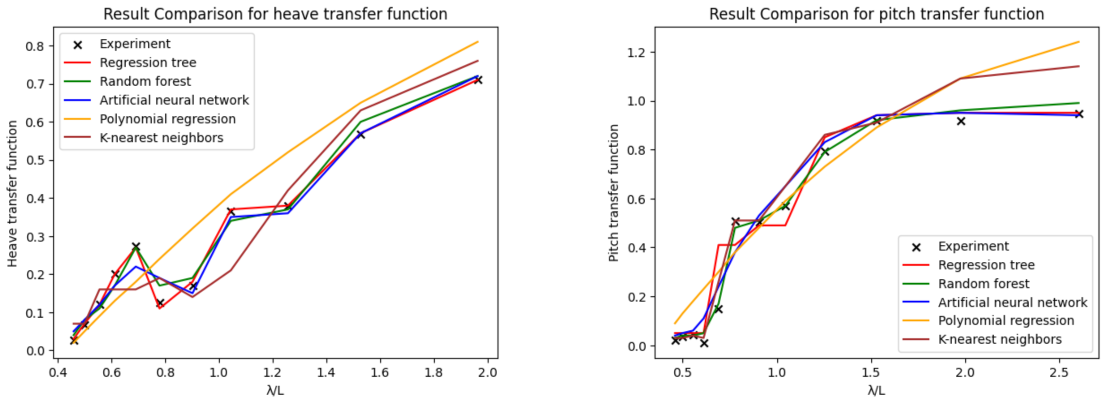


Figure 14: Experimental and predicted transfer function from different models for NPL 5b, Fn = 0.2, S/L = 0.4

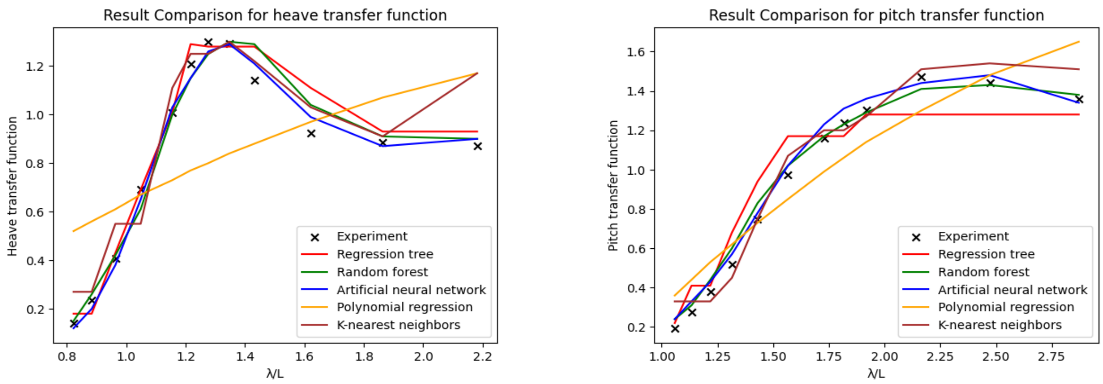


Figure 15: Experimental and predicted transfer function from different models for NPL 5b, Fn = 0.53, S/L = 0.4

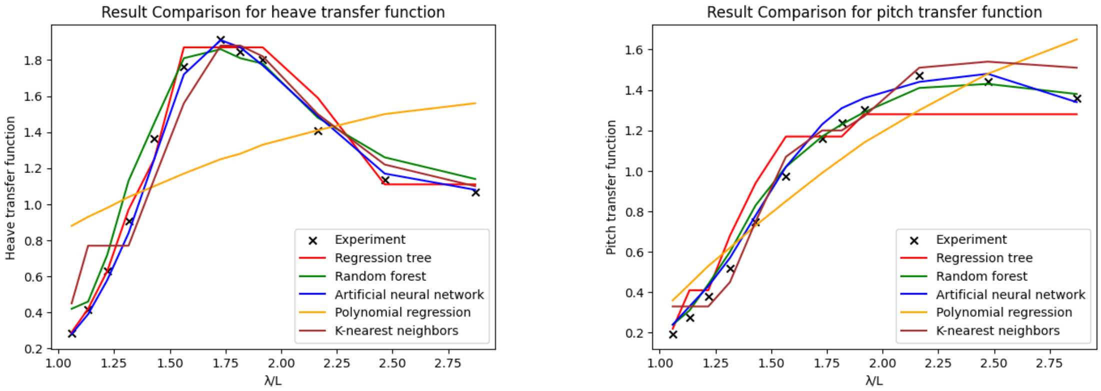


Figure 16: Experimental and predicted transfer function from different models for NPL 5b, Fn = 0.8, S/L = 0.4

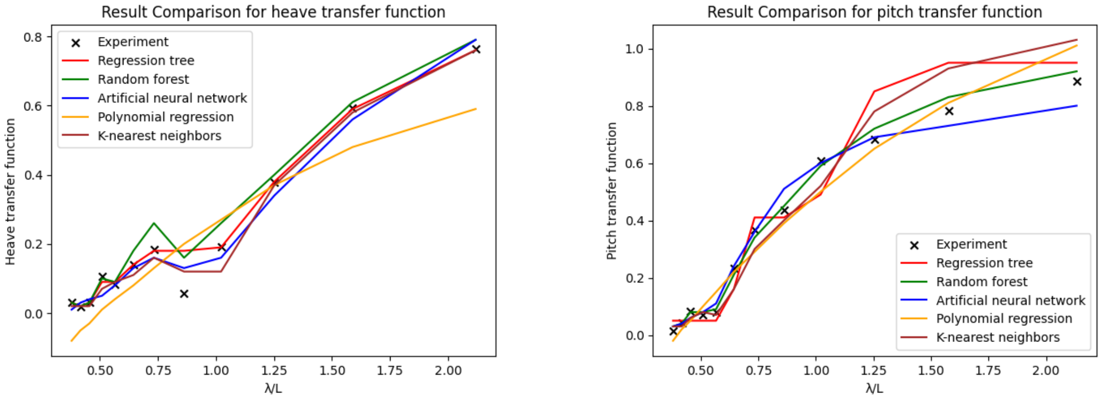


Figure 17: Experimental and predicted transfer function from different models for NPL 6b, Fn = 0.2, S/L = 0.2

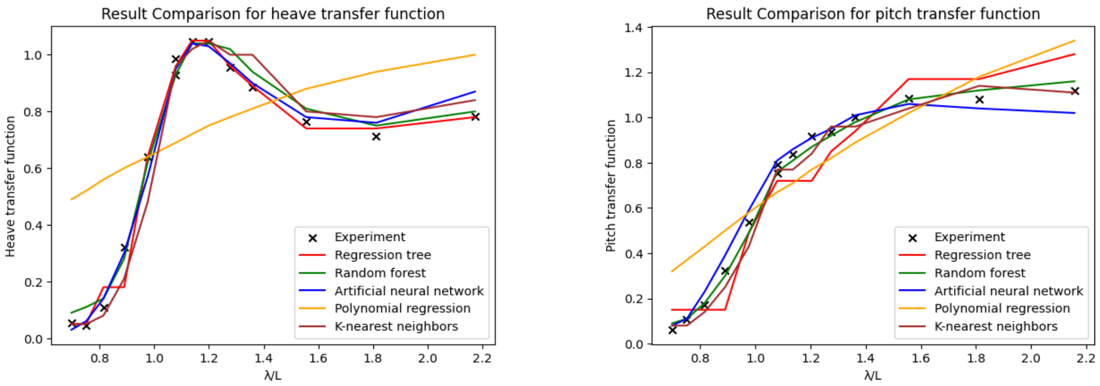


Figure 18: Experimental and predicted transfer function from different models for NPL 6b, Fn = 0.53, S/L = 0.2

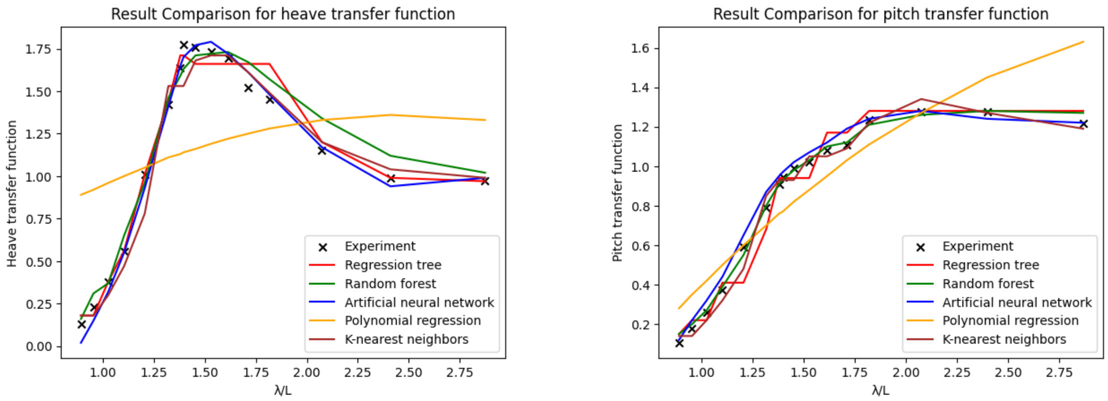


Figure 19: Experimental and transfer function from different models for NPL 6b, Fn = 0.8, S/L = 0.2

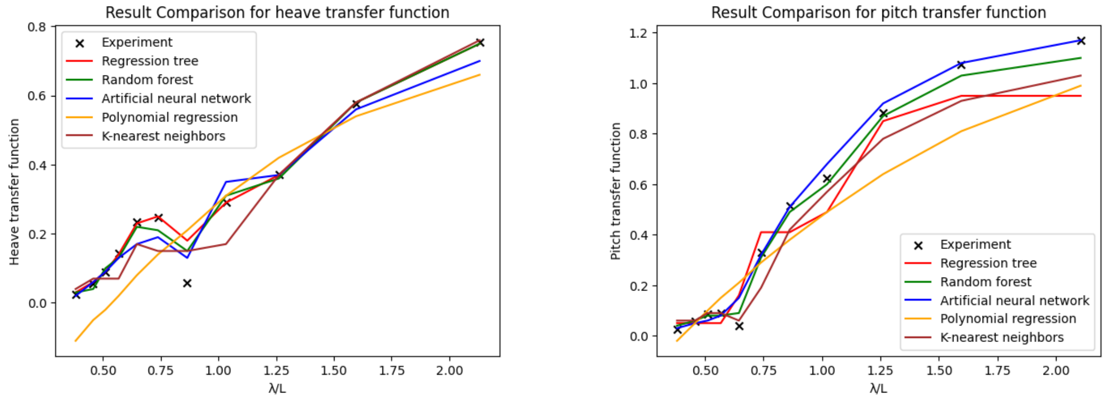


Figure 20: Experimental and predicted transfer function from different models for NPL 6b, Fn = 0.2, S/L = 0.4

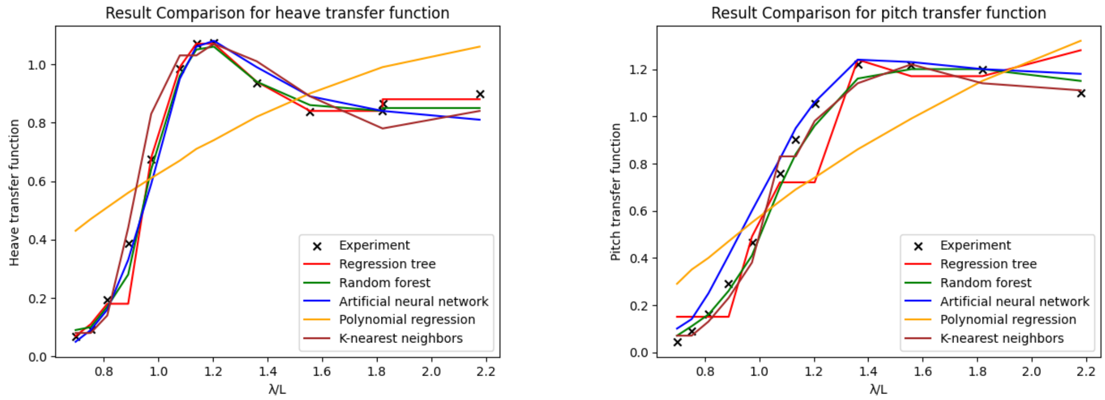


Figure 21: Experimental and predicted transfer function from different models for NPL 6b, Fn = 0.53, S/L = 0.4

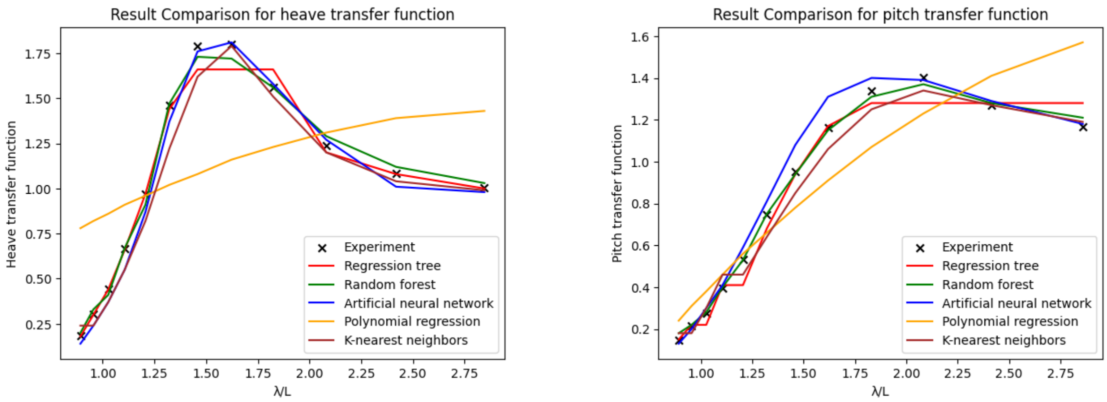


Figure 22: Experimental and predicted transfer function from different models for NPL 6b, Fn = 0.8, S/L = 0.4

The interpolated results obtained from the artificial neural network and K-nearest neighbors regression models for the NPL 4b catamaran at Froude numbers 0.2 and 0.5 and demi-hull spacings 0.2 and 0.4 were compared with the numerical technique described in Section 6. Figures 23-26 show that for lower Froude numbers, the 3D Green function accurately represents the motion response scenario, capturing all peaks for the two vessel designs. Higher Froude numbers indicate anomalies with the model test results. The current calculation shows a similar trend to the numerical results from the Strip theory and the 3D pulsating source methods (Moinul 2021). Although the results overestimate the experimental data, the 3D Green function method is still more accurate than those obtained from the strip theory and 3D pulsating source methods. However, the computational time of the 3D Green function method depends on the number of elements used for the calculation, and of course it takes more time than strip theory and the 3D pulsating source methods for numerical results.

For the benchmark cases considered here, the artificial neural network and K-nearest neighbors regression models showed closer agreement with the experimental data than the selected numerical approaches described in Section

6, as seen in Figures 23–26. However, this observation should be interpreted with caution and only within the scope of the present dataset. The comparison does not imply that the machine-learning models are broadly superior to conventional numerical methods since the present study is limited to the NPL catamaran series in long-crested head seas.

The artificial neural network's performance declines while predicting the heave transfer function for NPL 4b with a Froude number of 0.2 and demi-hull spacing of 0.2. Several factors may contribute to this issue. Artificial neural networks may have challenges when dealing with outliers in the training dataset due to the data distribution. Outliers may not follow the typical trends shown in most of the data, perhaps causing the model to generalize poorly (Hawkins 1980). The neural network design significantly impacts its capability to manage extreme values. Elaborate designs may better capture detailed data patterns but are more prone to overfitting, especially when outliers are present (Goodfellow et al., 2016). Therefore, the behavior observed in this case should be interpreted as a limitation of the present ANN model within a small and constrained dataset, rather than as a fully general conclusion about ANN-based motion prediction.

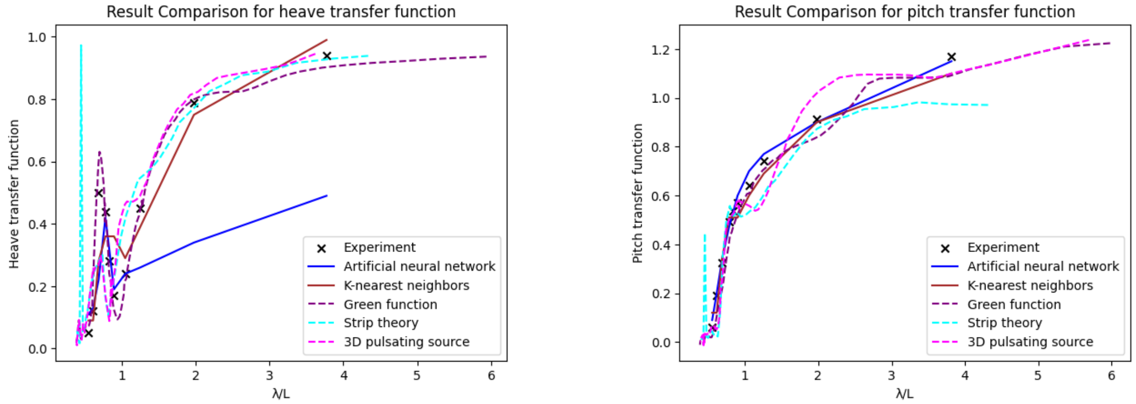


Figure 23: Comparison with numerical approach for NPL 4b at Fn = 0.2 and S/L = 0.2

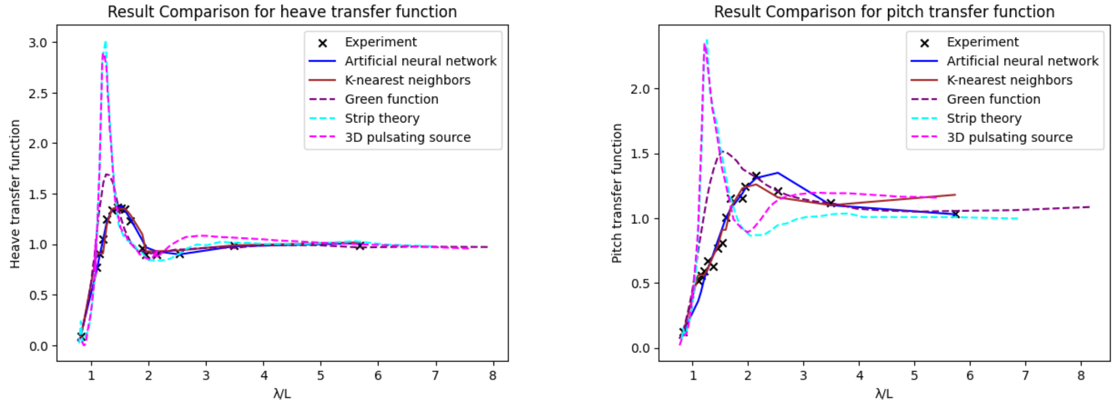


Figure 24: Comparison with numerical approach for NPL 4b at Fn = 0.53 and S/L = 0.2

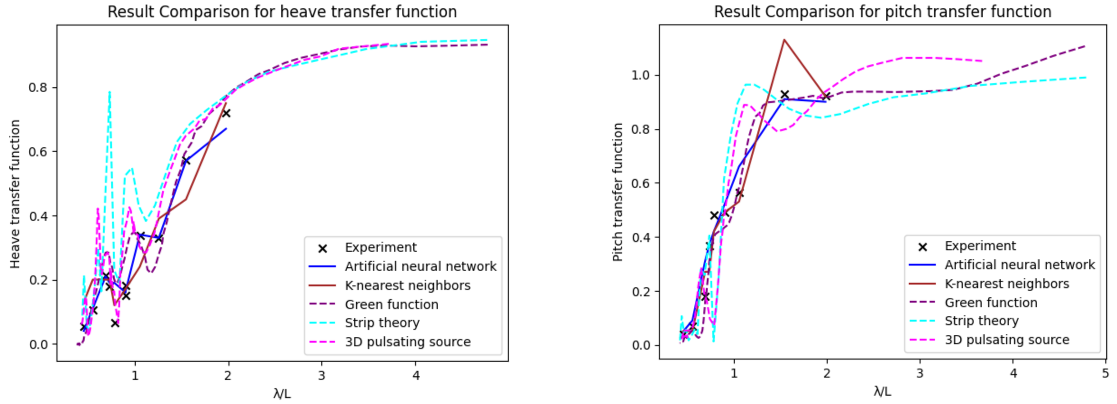


Figure 25: Comparison with numerical approach for NPL 4b at Fn = 0.2 and S/L = 0.4

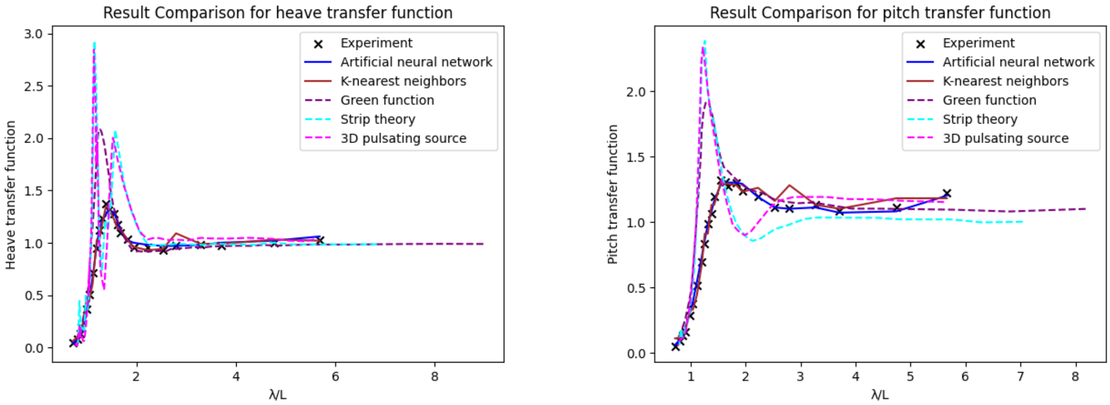


Figure 26: Comparison with numerical approach for NPL 4b at Fn = 0.53 and S/L = 0.4

According to the obtained results, the developed approach demonstrates the potential of machine-learning models to reproduce motion trends within the available experimental dataset and to support comparative assessment of the considered benchmark cases. However, due to the limited range of hull-form and operating parameters represented in the dataset, the present approach should be regarded as a constrained interpolation framework rather than a general design-stage prediction tool.

4. Conclusions

There has been a growing interest in incorporating data science into both the design and operation phases of shipping in recent years. The shipping industry has witnessed a substantial growth in the use of machine learning and statistical methods for ship motion prediction during the early design phase, owing to advancements in computer processing capacity and data accessibility. This facilitates the enhancement of the adaptability, analytical, and predictive capacities of data-driven models founded on machine learning through the extraction of latent information from gathered datasets. Such insights might unveil the impact of diverse parameters that influence the motion of ships.

This research presents a comparative analysis of different machine learning methods with respect to their predictive accuracy in modeling catamaran motion. The algorithms are trained and tested using a dataset derived from experimental findings obtained from NPL catamarans in long-crested head seas. A range of data analysis and data preparation approaches have been employed to enhance the performance of algorithms. Hyperparameter adjustment was conducted to mitigate the issues of overfitting and underfitting in the models and the models were assessed using K-fold cross-validation and test-set prediction. This research also offers comparative insight into several machine-learning models and their respective strengths and drawbacks. In addition, a numerical approach based on the 3D Green function was developed for the NPL 4b catamaran and compared with other numerical approaches. For the benchmark cases considered in this study, some of the machine-learning models showed closer agreement with the experimental data than the selected numerical results. However, this observation should be interpreted only within the present dataset and should not be generalized beyond the tested cases. The following are the summary of the main findings:

- All the regression models assessed exhibited varying levels of in-domain estimation capability for the motion transfer functions, although polynomial regression performed poorly in comparison with the other models.
- K-nearest neighbor, regression tree, random forest, and artificial neural network models showed good in-domain interpolation performance within the available dataset, indicating that these methods can capture local trends in the experimental data.
- The dataset utilized in this study was obtained exclusively from the experimental outcomes of NPL catamarans in long-crested head seas. Consequently, the block coefficient and breadth-to-draft ratio remained constant across the whole dataset, and only two demi-hull spacing ratios were included. As a result, the developed models cannot capture the influence of these parameters beyond the studied range and should therefore be interpreted within this restricted experimental domain.

- A significant limitation of the employed machine-learning models is their limited capacity for extrapolation. This implies that the present models should be viewed primarily as interpolation tools within the available NPL dataset.
- Since heading angle is not included as an input variable, the proposed methodology is only suitable for long-crested head-sea conditions and does not address more general seaway conditions involving sway, roll, and yaw responses.
- Although the artificial neural network and K-nearest neighbor models showed strong in-domain interpolation performance, these results should be interpreted cautiously because the limited dataset structure may increase the risk of overfitting, overly local interpolation, or overestimation of performance under random train/test and K-fold validation.

Overall, the present work should be interpreted as a methodological case study on the application and comparison of machine-learning models for interpolating heave and pitch motion characteristics within a well-known but limited experimental catamaran dataset. The study demonstrates that useful in-domain interpolation can be achieved from sparse experimental data within a constrained parameter space; however, it does not establish a general preliminary design tool for catamarans. Future work may investigate hybrid and physics-informed learning frameworks to incorporate hydrodynamic knowledge into the prediction process, reduce dependence on sparse experimental data, and potentially improve robustness outside the presently studied parameter range.

CRediT Author Statement

Argha Saha: Conceptualization, Methodology, Software, Validation, Writing-Original draft preparation; **Safia Alam Sumaiya:** Data curation, Writing-Original draft preparation, Visualization, Investigation**; N. M. Golam Zakaria**: Supervision, Writing-Reviewing and Editing.

References

Ameen, M., Phillips, A. B., Turnock, S. R. and Wilson, P. A. (2022): Ship speed prediction based on machine learning for efficient shipping operation, Ocean Engineering, Vol. 245, Article 110449. https://doi.org/10.1016/j.oceaneng.2021.110449

Bhattacharya, R. (1978): Dynamics of marine vehicles. ISBN-13: 9780471072065

Caramoy, S. and Aryawan, W. D. (2025): The application of artificial neural network–MISO (multi-input single-output) in catamaran ship roll motion prediction, IOP Conference Series: Earth and Environmental Science, Vol. 1461, No. 1, Article 012038. https://doi.org/10.1088/1755-1315/1461/1/012038

Cepowski, T. (2020): The prediction of ship added resistance at the preliminary design stage by the use of an artificial neural network, Ocean Engineering, Vol. 195, Article 106657. https://doi.org/10.1016/j.oceaneng.2019.106657

Cepowski, T. (2023): The use of a set of artificial neural networks to predict added resistance in head waves at the parametric ship design stage, Ocean Engineering, Vol. 281, Article 114744. https://doi.org/10.1016/j.oceaneng.2023.114744

Chollet, F. et al. (2015): Keras, GitHub, available at: https://github.com/fchollet/keras

D'Agostino, D., Serani, A., Stern, F. et al. (2022): Time-series forecasting for ships maneuvering in waves via recurrent-type neural networks, Journal of Ocean Engineering and Marine Energy, Vol. 8, pp. 479–487. https://doi.org/10.1007/s40722-022-00255-w

Danışman, D. B. (2014): Reduction of demi-hull wave interference resistance in fast displacement catamarans utilizing an optimized centerbulb concept, Ocean Engineering, Vol. 91, pp. 227–234. https://doi.org/10.1016/j.oceaneng.2014.09.018

Dogrul, A., Kahramanoglu, E. and Çakıcı, F. (2021): Numerical prediction of interference factor in motions and added resistance for Delft catamaran 372, Ocean Engineering, Vol. 223, Article 108687. https://doi.org/10.1016/j.oceaneng.2021.108687

Goodfellow, I., Bengio, Y. and Courville, A. (2016): Deep learning, Vol. 1, MIT Press.

Gourlay, T. (2008): Sinkage and trim of a fast displacement catamaran in shallow water, Journal of Ship Research, Vol. 52, No. 3, pp. 175–183.

Guedes Soares, C., Fonseca, N. and Pascoal, R. (2005): Experimental and numerical study of the motions of a turret-moored FPSO in waves, Journal of Offshore Mechanics and Arctic Engineering, Vol. 127, August 2005. https://doi.org/10.1115/1.1951774

Han, J., Kitazawa, D. and Murai, M. (2023): Numerical simulation of motion responses of a cabin-suspended catamaran in head waves, Proceedings of the ASME 2023 42nd International Conference on Ocean, Offshore and Arctic Engineering (OMAE2023), Paper No. OMAE2023-103284, Melbourne, Australia, 11–16 June 2023. https://doi.org/10.1115/OMAE2023-103284

Hassan, H. M., Elsakka, M. M., Refaat, A., Zhang, H., Yin, Z., Ahmed, A., Nashaat, H., Rizk, R. and Moustafa, M. M. (2025): Identifying the optimal gap separation of National Physics Laboratory (NPL) catamaran model using machine learning algorithms, Journal of Ocean Engineering and Marine Energy, Vol. 11, pp. 543–560. https://doi.org/10.1007/s40722-025-00386-w

Hawkins, D. M. (1980): Identification of outliers, Vol. 11, Springer.

Inoue, Y. and Makino, Y. (1989): The influence of forward speed upon three-dimensional hydrodynamic forces, Journal of the Society of Naval Architects of Japan, Vol. 1989, pp. 207–216. https://doi.org/10.2534/jjasnaoe1968.1989.166_207

Islam, M. M. (2021): Numerical analysis of dynamic response and wave-load on multihull ship advancing in waves, Postgraduate dissertation (thesis), Naval Architecture and Marine Engineering (NAME), BUET Institutional Repository.

Lee, J.-H., Kim, Y., Kim, B.-S. and Gerhardt, F. (2021): Comparative study on analysis methods for added resistance of four ships in head and oblique waves, Ocean Engineering, Vol. 236, Article 109552. https://doi.org/10.1016/j.oceaneng.2021.109552

Li, G., Kawan, B., Wang, H. and Zhang, H. (2017): Neural-network-based modelling and analysis for time-series prediction of ship motion, Ship Technology Research, Vol. 64, No. 1, pp. 30–39. https://doi.org/10.1080/09377255.2017.1309786

Liu, L., Yang, Y. and Peng, T. (2022): Machine learning prediction of 6-DOF motions of KVLCC2 ship based on RC model, Journal of Ocean Engineering and Science, ISSN 2468-0133. https://doi.org/10.1016/j.joes.2022.08.004

Liu, S., Xu, R., & Papanikolaou, A. (2021): Prediction of the motion of a ship in regular head waves using artificial neural networks. In ISOPE International Ocean and Polar Engineering Conference (pp. ISOPE-I). ISOPE.

Martić, I., Degiuli, N., & Grlj, C. G. (2023): Prediction of Added Resistance of Container Ships in Regular Head Waves Using an Artificial Neural Network. Journal of Marine Science and Engineering, 11(7), 1293. https://doi.org/10.3390/jmse11071293

Mittendorf, M., & Papanikolaou, A. D. (2021): Hydrodynamic hull form optimization of fast catamarans using surrogate models. Ship Technology Research, 68(1), 14–26. https://doi.org/10.1080/09377255.2020.1802165

Molland, A., Wellicome, J. and Couser, P. (1994): Resistance experiments on high-speed displacement catamaran forms: Variation of length-displacement ratio and breadth-draught ratio, Ship Science Report No. 71, Department of Ship Science, University of Southampton.

Najafi, A., Nowruzi, H. and Ghassemi, H. (2018): Performance prediction of hydrofoil-supported catamarans using experiment and ANNs, Applied Ocean Research, Vol. 74, pp. 66–84. https://doi.org/10.1016/j.apor.2018.02.017

Nazemian, A., Boulougouris, E. and Aung, M. Z. (2024): Utilizing machine learning tools for calm-water resistance prediction and design optimization of a fast catamaran ferry, Journal of Marine Science and Engineering, Vol. 12, No. 2, Article 216. https://doi.org/10.3390/jmse12020216

Nguyen, T. T., Phan, T. L. and Le, T. H. (2025): Numerical study on the motion and wave loads of an uncrewed surface vehicle catamaran at low speed in sea waves, Journal of Applied Fluid Mechanics, Vol. 18, No. 9. https://doi.org/10.47176/jafm.18.9.3330

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

Qian, P., Yi, H. and Li, Y. (2015): Numerical and experimental studies on hydrodynamic performance of a small-waterplane-area-twin-hull (SWATH) vehicle with inclined struts, Ocean Engineering, Vol. 96, pp. 181–191. https://doi.org/10.1016/j.oceaneng.2014.12.039

Santoso, B., Romadhoni, Custer, J. and Ariyany, Z. (2026): Optimization of electric propulsion systems for ferry vessels: A case study in riverine operations, Journal of Naval Architecture and Marine Engineering, Vol. 23, No. 1, pp. 75–92. https://doi.org/10.3329/jname.v23i1.81886

Tian, X. and Song, Y. (2023): Machine learning for short-term prediction of ship motion combined with wave input, Applied Sciences, Vol. 13, No. 9, Article 5298. https://doi.org/10.3390/app13095298

van 't Veer, AP. (1998). Experimental results of motions, hydrodynamic coefficients and wave loads on the 372 catamaran model. WbMT.

Wellicome, J. F., Tamarel, P., Molland, A. F. and Couser, P. R. (1995): Experimental measurements of the seakeeping characteristics of fast displacement catamarans in long-crested head seas, Ship Science Report No. 89, University of Southampton.

Yang, Y., Zhang, Z., Zhao, J., Zhang, B., Zhang, L., Hu, Q. and Sun, J. (2024): Research on ship resistance prediction using machine learning with different samples, Journal of Marine Science and Engineering, Vol. 12, No. 4, Article 556. https://doi.org/10.3390/jmse12040556

Zhang, Z., Jiang, B., Wang, L., Wei, S., Li, T., Zhang, G. and Zong, Z. (2024): Hydrodynamic analysis of different formation configurations of catamaran in regular head waves, Journal of Marine Science and Engineering, Vol. 12, No. 9, Article 1577. https://doi.org/10.3390/jmse12091577