Data-driven prediction of added-wave resistance on ships in oblique waves

A comparison between tree-based ensemble methods and artificial neural networks

Mittendorf, Malte; Nielsen, Ulrik D.; Bingham, Harry B.

Published in:
Applied Ocean Research

Link to article, DOI:
10.1016/j.apor.2021.102964

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Applied Ocean Research 118 (2022) 102964

Contents lists available at ScienceDirect

Applied Ocean Research

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

Data-driven prediction of added-wave resistance on ships in oblique waves—A comparison between tree-based ensemble methods and artificial neural networks

Malte Mittendorf [a,*], Ulrik D. Nielsen [a,b], Harry B. Bingham [a]

[a] DTU Mechanical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[b] NTNU Centre for Autonomous Marine Operations and Systems, 7052 Trondheim, Norway

ARTICLE INFO

Keywords:
Ship hydrodynamics
Added-wave resistance
Machine learning
Tree-based ensemble methods
Artificial neural networks

ABSTRACT

The present paper implements machine learning methods for the prediction of the added-wave resistance of ships in head to beam wave conditions. The study is focused on non-linear regression algorithms namely Random Forests, Extreme Gradient Boosting Machines and Multilayer Perceptrons. The employed dataset is derived from results of three different potential flow methods covering a wide range of operational conditions and 18 hull forms in total. The rational data preprocessing makes up the core part of the paper having its focal point on practical application. Moreover, a rigorous hyperparameter study based on Bayesian optimization is conducted, and the validation of the final models for three case studies against numerical and experimental data as well as two established prediction techniques shows satisfactory generalization in case of the neural network. The tree-based ensemble methods, on the other hand, are not able to generalize sufficiently from the given parameter discretization of the underlying dataset.

1. Introduction

The maritime industry is changing: The increased environmental awareness of society and policymakers has lead to the introduction of several rules and regulations incentivizing the enhancement of energy efficiency of ships, e.g. the mandatory compliance of new and existing ship designs with the Energy Efficiency Design Index (EEDI) baseline, International Maritime Organization (2011). Broadly speaking, the EEDI is the ratio of transport work and CO^2-emissions, i.e. it sets the environmental pollution in relation to the socioeconomic benefit, see Kristensen (2010) for greater detail. The legislation of the EEDI framework enforced a trend of reducing installed engine power and slow steaming. The latter (slow steaming) increases the added-wave resistance fraction of the total resistance which in combination with the former makes the risk of losing maneuverability in adverse weather conditions pivotal in ship design and operation. Moreover, the EEDI-equation includes the weather factor f_w incorporating the effect of the added resistance or, rather, the consequent involuntary speed loss in a seaway due to wind and waves. However, the reliable determination of the added-wave resistance and hence of the factor f_w is problematic both when using model test data and results of numerical calculations as has been shown by Shigunov et al. (2018). For this reason, the present work provides a machine learning (ML) approach for the prediction of the added resistance in waves and compares the model estimates to traditional prediction methods. It is stressed that the ML task itself is considered as a supervised regression approach for an efficient estimate of the added-wave resistance without any detailed hull shape information.

1.1. Literature review

In early ship design stages as well as in the assessment of a vessels' operational performance an efficient prognosis of the added-wave resistance is essential; hence, it is compelling to pursue an empirical regression approach in this respect. Empirical formulas were and are still widely applied in naval architecture, but machine learning regression techniques are also well established in ship design – e.g. Bertram and Mesbahi (2000) – and ship hydrodynamics, especially as surrogate models that approximate results of computationally expensive simulation methods by regression models. For instance, Sclavounos and Ma (2018) estimate ship roll motions and their decay by a Support Vector Machine (SVM), which is a non-linear, instance-based ML-algorithm. Moreover, Mittendorf and Papanikolaou (2021) present a global surrogate-based multi-objective hull form optimization using a genetic algorithm applied to Multivariate Adaptive Regression Splines

* Corresponding author.
E-mail address: mamit@mek.dtu.dk (M. Mittendorf).

https://doi.org/10.1016/j.apor.2021.102964
Received 28 June 2021; Received in revised form 24 September 2021; Accepted 10 November 2021
Available online 9 December 2021

(MARS) providing the calm water resistance of the case study. Oberhagemann et al. (2019) provide an approach for structural hull monitoring including extreme waves by combining Computational Fluid Dynamics (CFD) calculations with spectral methods and a tree-based ensemble approach by a Random Forest (RF). In contrast to the above-mentioned publications, artificial neural networks or rather multilayer perceptrons (MLP) are predominantly used in state-of-the-art publications for the prediction of added-wave resistance on ships. Cepowski (2020), for instance, trains an MLP on benchmark model test data in head wave conditions and extracts an empirical formula from the resulting model parameters and activations. Moreover, Herradon de Grado and Bertram (2016) train a MLP based on CFD results of the added-wave resistance of one container vessel for headings ranging from head to following waves. In addition, they consider the added-wind resistance by training another MLP on data derived from an empirical formula.

1.2. Motivation and objective

As can be inferred from the presented literature, the individual approaches are either limited to one case ship or to one wave direction. Therefore it is the aim of this paper to provide a more widely applicable surrogate model; both in terms of geometric variation of the hull and with regards to the relative wave direction. The latter is of high relevance as model test data suggests that the added-wave resistance, R_{AW}, will be most prominent in head to bow oblique waves, Valanto and Hong (2015). Another novelty will be the focus on the data preprocessing (feature engineering) and hyperparameter study as these steps are highly influential on the model performance and are not considered in the above-mentioned publications. Moreover, it is envisioned to investigate a variety of machine learning models, and not MLP exclusively, since these are characterized by numerous inherent caveats, such as no scale invariance, general opaqueness and sensitivity to multicollinearity,[1] cf. Hastie et al. (2009). In contrast, non-parametric tree-based ensemble models, such as random forests and (extreme) gradient boosting methods are not affected by the mentioned drawbacks. Hence, this contribution is focused on a comparison between MLP and tree-based ensemble methods applied to the prediction of the added-wave resistance of ships based on a database generated using three different numerical calculation methods in head and bow oblique waves, i.e. in head to beam seas.

Finally, the goal of this contribution is to illuminate the following research question: Do tree-based ensemble methods yield higher accuracy and transparency compared to multilayer perceptrons for the prediction of the added-wave resistance on ships? In this respect, the developed statistical models will be compared to the semi-empirical formula of Liu and Papanikolaou (2020), model test data and RANS (Reynolds averaged Navier Stokes) CFD results for several validation cases. It is stressed that this study is limited to long-crested and regular waves; however, the extension to irregular waves is trivial given a suitable wave energy density spectrum.

1.3. Composition

In Section 2, a brief sketch of the theory behind the added-wave resistance and the applied machine learning models will be provided. Section 3 introduces the underlying database and contains the applied methodology, i.e. data preprocessing, model selection and the overall training configuration. The main findings and results are displayed in Section 4. Moreover, the final models will be applied to three case studies and the models' performances are discussed. Finally, in Section 5, conclusions are drawn and several suggestions for future work are presented.

[1] Multicollinearity expresses when two or more predictors are highly related to each other, i.e. do not provide unique or independent information to the regression approach and thus negatively impact model performance.

2. Theoretical background

2.1. Added-wave resistance

The added-wave resistance R_{AW} is the unsteady longitudinal force a ship experiences apart from the calm-water and wind resistances in a realistic seaway. The so-called drift force is of second order and constitutes three components: Wave diffraction, radiation and their interaction. Following Faltinsen (2005), the added resistance in waves results from the ship's ability of generating unsteady waves. A ship in resonant heave and pitch conditions generates the most pronounced waves per unit of wave amplitude due to the large relative vertical motions. For more comprehensive theoretical considerations regarding the added resistance in waves reference is made to the literature, e.g. Strøm-Tejsen et al. (1973).

The added-wave resistance is expressed in non-dimensional form in Eq. (1).

$$C_{AW} = \frac{R_{AW} L_{pp}}{\rho g \zeta^2 B_m^2} \tag{1}$$

In view of Eq. (1), the added-wave resistance is dependent on the ship's length and the beam denoted as L_{pp} and B_m respectively, as well as the wave amplitude ζ, the gravity acceleration is g, and the seawater density ρ. In addition, the ship's draft T_D, the longitudinal radius of gyration k_{yy}, the bow shape as well as the fullness of the hull, i.e. the block coefficient C_B, have a pronounced impact on the magnitude of R_{AW}. Furthermore, operational conditions, such as loading conditions (i.e. GM_T), Froude number $Fn = \frac{U}{\sqrt{gL_{pp}}}$, relative wave direction β, intrinsic wave frequency ω_0 and wave steepness H/λ are also of significant importance.

There are two main approaches towards the calculation of the added-wave resistance using potential flow theory: (1) The far-field methods and (2) the near-field methods. The far-field methods are based on energy considerations for the reflected and radiated waves, the momentum flux at infinity and the work performed by the floating structure. These calculation methods are split further into two sub-methods for the determination of added resistance in waves: Maruo (1957) and Salvesen (1978) proposed approaches based on momentum-conservation, whereas Gerritsma and Beukelman (1971) formulated a far-field method based on the radiated energy method. Both approaches were implemented using the Salvesen, Tuck and Faltinsen (STF) strip theory developed by Salvesen et al. (1970) and are still widely used today. Conversely, the near-field methods determine the steady second order force by integration of the hydrodynamic, steady pressure acting on the wetted surface of the ship's hull. This pressure is be calculated exactly from the first order potential and its derivatives (cf. Boese (1970)). Faltinsen et al. (1980) proposed a near-field approach by an asymptotic formula for the added resistance in short waves which is based on slender body theory. The herein employed dataset consists of results of three different methods in their respective preferred definition range: (1) A Green's Function Method is used for the calculation of the drift force in non-forward speed conditions, (2) results of Salvesen's formulation are employed for the determination of the added-wave resistance in large to medium wavelengths and (3) Faltinsen's asymptotic formula is applied in the region of high wave frequencies. Since details about the dataset are given in Section 3, it is sufficient to mention that higher wave frequencies have the most practical relevance since ships sail the majority of the time in low to moderate sea states, (Faltinsen, 2005). Liu and Papanikolaou (2017) show that relatively short waves $\lambda/L_{pp} < 0.5$ contribute significantly to the added resistance of large ships in seafaring conditions, but are often neglected.

Ultimately, the inherent non-linear characteristics of the added-wave resistance lead to major complexities for theoretical calculations. In fact, the SHOPERA (energy efficient Safe SHip OPERAtion) benchmark study revealed remarkable deviations of numerical, theoretical simulation results and model test data, cf. Shigunov et al. (2018).

Even though RANS methods are, in fact, able to reproduce non-linear characteristics such as wave breaking, the accuracy of these methods is not superior in the above mentioned SHOPERA benchmark study. CFD simulations require a very fine spatial and temporal discretization in order to prevent numerical diffusion of the wave elevation in the free surface region, resulting in a high computational cost. The majority of studies pertaining the added resistance in waves is focused on head wave conditions (e.g. Sigmund and el Moctar (2018) or Sadat-Hosseini et al. (2013)), but recent studies such as Uharek (2019) or Kim et al. (2021) consider relative wave headings $\beta \in [180, 0]$ deg. discretized by 30 deg. increments.

2.2. Machine learning

Mitchell (1997) defines a machine learning task as the following: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E". In the herein presented work, the task T is considered as a multivariate regression problem, i.e. the learning algorithm is asked to approximate the function $f : \mathbb{R}_n \rightarrow \mathbb{R}$. The supervised machine learning algorithm experiences E, a dataset $D\{(\mathbf{x_i}, y_i)\}(|D| = n, \mathbf{x_i} \in \mathbb{R}^d, y_i \in \mathbb{R})$, where a feature vector $\mathbf{x_i}$ is associated with a known scalar target y_i and the feature matrix is of shape $n \times d$, where n corresponds to the samplesize and d reflects the dimensions of the feature space. Moreover, the approximation $\hat{y_i}$ is obtained by training the model, i.e. minimizing the performance measure P which is specific to the task T. In the following, the fundamentals of the individual learning algorithms are provided. The models are selected such that they are able to handle high-dimensional, non-linear and continuous data.

2.2.1. Artificial neural networks

Neural networks are widely applied as information processing systems and common learning algorithms which are inspired by the governing principles of mammalian brains and their dendrites. These models are directed acyclic graphs and are used as universal function estimators, since the combination of affine functions, i.e. matrix operations, and non-linearities introduced via activation functions, lead to the *Universal Approximation Theorem*, cf. Goodfellow et al. (2016). The most widespread artificial neural network is the multilayer perceptron (MLP), and in its basic configuration it comprises three layers: An input layer, a hidden layer and an output layer; hence, it can be considered as a composite function.

$$\hat{y_i} = f_L[W_L, f_{L-1}(W_{L-1}, \dots, f_1\{W_1, \mathbf{x_i}\})] \tag{2}$$

In Eq. (2), f_i corresponds to the non-linearity or activation function applied in layer L on the weight matrix W_i. It is noted that the usual bias term b is omitted. The parameters or weight matrices W_i are initialized randomly and optimized during training. In general, the training process of neural networks is formulated as a gradient-based optimization problem: The loss function P is obtained by backpropagation and minimized by, e.g. a stochastic gradient descent algorithm. Recently, deep architectures achieved outstanding results in a variety of different disciplines ranging from image recognition to natural language processing. In the herein presented work, relatively shallow MLP architectures are applied for the non-linear regression approach. Further and more elaborate details of neural networks can be found in Goodfellow et al. (2016) and Hastie et al. (2009).

2.2.2. Tree-based ensemble methods

Ensemble methods combine predictions from multiple so-called *weak* machine learning algorithms or base learners for more accurate predictions. The building block of the herein employed ensemble models is the non-parametric Decision Tree. This greedy[2] learning algorithm infers decision rules through features in a supervised environment by splitting the feature space into cuboid regions maximizing the information gain and minimizing impurity in a recursive and progressive fashion. Impurity – or variance in case of regression – refers to the homogeneity of the labels at each node and the information gain is the difference between the parent node impurity and the weighted sum of the child node impurities, i.e. the variance reduction. In this study, the CART (Classification and Regression Tree) implementation of Breiman et al. (1984) is utilized and it constructs binary trees using the feature and threshold that yields the largest information gain at each node. In a regression context it is vital to note that predictions of decision trees are not smooth, but *constant and piecewise approximations*. For this reason, TEM (Tree-Based Ensemble Methods) are known – like most other ML methods – for their insufficient extrapolation capabilities. For more detailed information about regression trees, cf. Hastie et al. (2009) or Bishop (2006).

In this study, two different ensemble methods are considered: (1) In a *bagging* (bootstrap aggregating) approach, a multitude of weak learners are trained in *parallel* on random subsets of the entire training data with replacement. (2) *Boosting*, on the other hand, is a sequential approach and each individual, base learner is influenced by the imperfections of the previous model. Due to the *sequential* manner it allows the definition and minimization of an arbitrary loss function. However, boosting does obviously not reach the computational efficiency of bagging. In the following, additional details and comments about the considered ensemble methods (Random Forests and Extreme Gradient Boosting Trees) are given.

Random forests

Random Forest (RF) Regressors were proposed by Breiman (2001) and grow decision trees in a probabilistic scheme. The algorithm employs decorrelated decision trees in its CART implementation as weak, underfitting learners circumventing the pitfall of overfitting. The output is calculated as the average prediction of the individual decision trees.

$$\hat{y_i} = \frac{1}{M} \sum_{m=1}^{M} \phi_m(\mathbf{x_i}) \tag{3}$$

As can be inferred in Eq. (3), M corresponds to the number of base learners and ϕ_m is considered as the output of the decision tree trained on a bootstrap of the data with a random subsample of the features, i.e. random variable selection at each node. These two sources of randomness decrease the variance of the random forest estimator. In fact, the weak base learners typically show high variance, but the enforced randomness yields decision trees with decoupled or rather decorrelated prediction errors. Hence, by averaging (cf. Eq. (3)) the individual predictions, some errors cancel each other. The variance reduction leads to increased model generalization and robustness to outliers, but might in turn slightly increase the bias.

Extreme gradient boosting trees

Extreme Gradient Boosting Regressors (XGBR) are based on the boosting meta-algorithm, which yields an ensemble model using decision trees as sequential base learners. XGBRs construct an additive, stage-wise model and it based on the optimization of arbitrary, differentiable and regularized loss functions. Gradient descent algorithms are utilized for the creation of new trees based on the residual errors of the previous one, i.e. with the addition of each tree, the model becomes more expressive. In contrast to random forests, each decision

[2] Greedy algorithms build a solution to an optimization problem in a heuristic and recursive multi-step manner by making locally optimal choices at each level.

tree contains weights w_j assigned to each leaf j and its value as well as the tree structure itself are determined by minimizing the objective function. The loss function $\mathcal{L}$ comprises of either the $L1$ or $L2$ error and a regularization term penalizing model complexity, cf. Eq. (4).

$$\mathcal{L} = \sum_i l(y_i, \hat{y}_i) + \sum_k \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{4}$$

It is noted that l is a differentiable loss function, T corresponds to the number of leaf nodes, whereas γ and λ are used as hyperparameters for regularization. The hyperparameter λ decreases the risk of overfitting, whereas γ penalizes each tree k for growing additional leaves. The optimal w_j is obtained by using a second-order Taylor expansion of the formulated loss function (cf. Eq. (4)). The determination of the optimal tree structure is performed by a greedy method enabling the selection of the best split candidate for adding new leaves. Lastly, the sequential, additive prediction process is evident in the following equation due to the iteration index denoted as t.

$$\hat{y}_i^{(t)} = \sum_{m=1}^{M} \phi_m(\mathbf{x_i}) = \hat{y}_{i,t-1} + \phi_t(\mathbf{x_i}) \tag{5}$$

The scores w_j and respective tree structure are summarized in ϕ_t in Eq. (5). Only the fundamentals of XGBR were conveyed herein, detailed information, such as split finding algorithms, weighted quantile sketch and system design are presented more elaborately in Chen and Guestrin (2016). Lastly, the XGBR algorithm is versatile and works well on structured data; however, its use is problematic for sparse and dispersed data. In comparison to the random forest, the XGBR algorithm has more adjustable hyperparameters requiring more computational effort during the parameter study. Ultimately, this algorithm does not only reduce variance – similarly to the random forest – due to the ensemble approach, but in most cases also the bias due to its sequential approach as every decision tree is affected by the imperfections or errors of the previous one.

3. Data and methods

3.1. Dataset

The underlying data of the DTU Design Tool, cf. Nielsen (2015) and Martinsen (2016), was used for the training setup of the machine learning models. The DTU design tool aggregates results of three different calculation methods for the determination of the added-wave resistance R_{AW} in irregular waves and employs linear interpolation on a set of quadratic transfer functions. As mentioned in Section 1, the momentum conservation method by Salvesen (1978) was used for low to medium wave frequencies, and the pressure integration method by an asymptotic formula proposed by Faltinsen et al. (1980) was utilized for the high wave frequencies. The required quantities, such as added mass and damping matrices as well as first order potential were computed by the strip theory solver I-Ship Petersen (2000), which is based on the STF method (Salvesen et al., 1970). Lastly, the calculations for zero-forward speed were carried out using the Greens Function panel method WAMIT (Lee and Newman, 2013). The main particulars of the target ship and the parameters of a parametric wave energy spectrum are then specified as an input to the MATLAB tool which performs the interpolations.

The program constitutes two sets of hull variants: One slender, relatively fast container vessel type ($C_B = 0.58$) and a blunt, relatively slow bulk carrier or tanker type ($C_B = 0.84$). It is stressed that the results for the two parent hulls are considered as separate datasets in the presented work. In addition, both datasets are identical considering their shape and derivation by numerical tools. One database is set up for nine scaling variants totaling 18 different hulls and their non-dimensional main particulars are presented in Table 1. The ratios are selected for minimizing the epistemic uncertainty, i.e. the entire design space is taken into account. The hull #9 may appear to be unrealistic, however, the ratios of the main particulars of contemporary cruise ships closely match the selected values.

Table 1
The scaling variants of the parent hulls from Martinsen (2016).

	L_{pp}/B_m	B_m/T_D	L_{pp}/T_D
#1	4.0	2.0	8.0
#2	4.0	3.5	14.0
#3	4.0	5.0	20.0
#4	6.0	2.0	12.0
#5	6.0	3.5	21.0
#6	6.0	5.0	30.0
#7	8.0	2.0	16.0
#8	8.0	3.5	28.0
#9	8.0	5.0	40.0

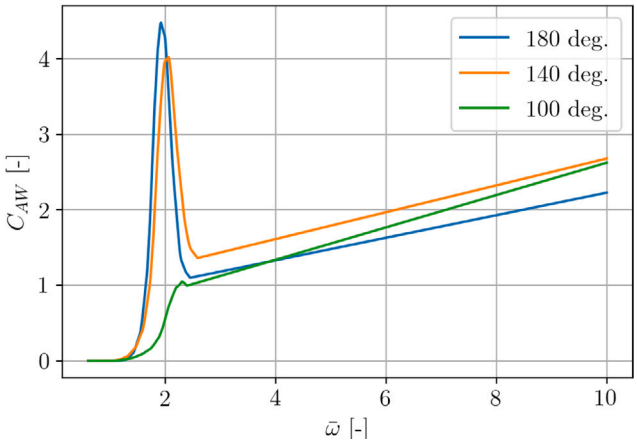


Fig. 1. Sample plot of hull #1 ($C_B = 0.84$) at $Fn = 0.1$ of the non-dimensional added-wave resistance C_{AW} and for various relative wave headings β.

On a sidenote, all nine blunt-type scaling variants share the same longitudinal radius of gyration $k_{yy}/L_{pp} = 0.25$, block coefficient $C_B = 0.84$ and waterplane coefficient $C_{WP} = 0.92$. The same holds for the slender container ship hulls with the following characteristics: $k_{yy}/L_{pp} = 0.25$, $C_B = 0.58$ and $C_{WP} = 0.79$. It is noteworthy that the scaling combinations are chosen to reflect the variety of the current world fleet, Nielsen (2015). The database consists of a large set of quadratic transfer functions given for the non-dimensional wave frequency $\bar{\omega} = \omega_0/\sqrt{g/L_{pp}}$ discretized into 941 values and $\bar{\omega} \in [0.6, 10.0]$.

In view of Fig. 1, it becomes clear that assembling the database using three different calculation method leads to discontinuities in the slope of the curves and somewhat arbitrary transitions posing a challenge to the present regression approach. Especially, since the intersection points of the individual quadratic transfer functions between Faltinsen's formula and Salvesen's method lead to inconsistencies, as in some cases the two curves do not intersect at all. In these rare cases, the Salvesen estimate is used instead for C_{AW} in short waves, (Martinsen, 2016).

Six different ship speeds have been taken into account considering Froude numbers in the interval $Fn \in [0.0, 0.25]$ with increments of 0.05. The wave direction $\beta \in [90, 180]°$ is discretized into 10 directions covering head to beam seas. It is stressed that the added resistance for beam to following seas is assumed to be zero. Moreover, the slenderness ratio $L/\nabla^{1/3}$, where ∇ corresponds to the ship's volume displacement, and the following ratios L_{pp}/B_m, B_m/T_D and L_{pp}/T_D are included in the dataset. Ultimately, each database has a sample size of $n = 508,140$ and one exemplary and properly non-dimensionalized sample in a ML setup has the following shape $\mathbf{x_i} = [L_{pp}/B_m, B_m/T_D, L_{pp}/T_D, L_{pp}/\nabla^{1/3}, \beta, Fn, \bar{\omega}]$ and $y_i = C_{AW}$.

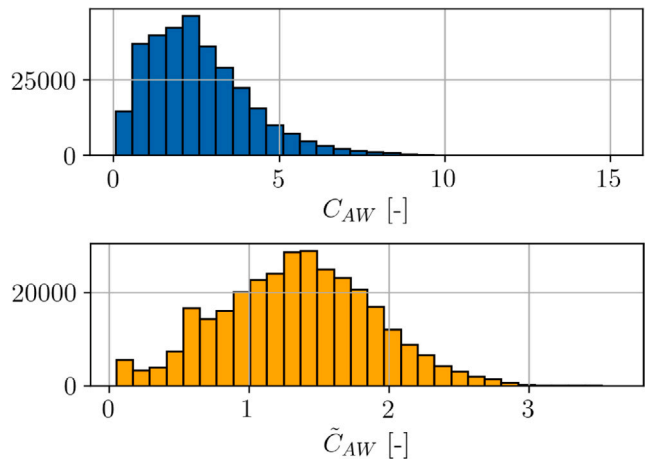


Fig. 2. The application of the Power Transform to the target distribution in black leads to the transformed one in orange. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Methodology

The applied methodology of the presented work consists of three parts: The initial data preprocessing and the model selection part in which the hyperparameter optimization is conducted. The final model assessment is carried out with respect to accuracy, generalization capability, and transparency. From a statistical point of view, the dataset has three distinct characteristics: (1) The target distribution is highly skewed, (2) the behavior of the target is non-linear and (3) the dataset is - taking both databases into account - with a samplesize of $n > 10^6$ relatively rich. Hence, the learning algorithms were selected such that they are capable of handling non-linear and high-dimensional, tabular data. It is noted, that the data preprocessing and model selection part is only carried out for the $C_B = 0.84$ dataset and the findings are adopted for the second database ($C_B = 0.58$).

A transformation of the dependent variable y_i is required as the original target distribution is biased and unbalanced towards lower C_{AW} values (cf. Fig. 2). The Gaussianity or normality of the target distribution is a presupposition of linear regression techniques in statistical modeling and has also a beneficial effect on the performance of non-linear machine learning methods. Logarithmic transforms are monotonic and do not change the locations of optima and also lead to far more stable numerical algorithms. Thus, prior to the training process, the target C_{AW} was transformed according to the Yeo and Johnson (2000) power transformation to $\tilde{C}_{AW}$. The Yeo-Johnson transformation was chosen over the Box–Cox transformation, since it does not impose any restrictions on y_i. The Yeo-Johnson transformation maximizes the marginal likelihood function by the modification of the hyperparameter λ and is defined in the following Eq. (6).

$$y_i^{(\lambda)} = \begin{cases} ((y_i+1)^{\lambda}-1)/\lambda & \text{if } \lambda \neq 0, y \geq 0 \\ \log(y_i+1) & \text{if } \lambda = 0, y \geq 0 \\ -[(-y_i+1)^{(2-\lambda)}-1]/(2-\lambda) & \text{if } \lambda \neq 2, y < 0 \\ -\log(-y_i+1) & \text{if } \lambda = 2, y < 0 \end{cases} \tag{6}$$

It is stressed that the used transformation is more versatile than the well-known Box-Cox transformation due to the treatment of negative values. After the first calculation, λ was set to $\lambda = 0.1814$ in order to maintain reproducibility. Such a low λ value indicates that the optimal transformation is quite close to the logarithm. The effect of the Yeo–Johnson transformation on the specific target distribution can clearly be seen in Fig. 2.

Fig. 2 shows that the initial target distribution (black) has been transformed into a symmetric and asymptotic distribution (orange) which resembles a standard normal distribution. It is stressed that the transformation does not impact values equal to zero and thus these are disregarded in Fig. 2 for the sake of clarity. Neural networks are not scale invariant, hence a transformation is necessary for their application. On the other hand, tree-based ensemble methods are scale invariant, but their performance is negatively affected by dispersed data. For this reason and in order to maintain a consistent training methodology across all models, transformed values were fed into both model types. Due to the evenly discretized dataset only normalization, i.e. scaling according to the extreme values, is applied to the feature space.

Furthermore, the polynomial feature expansion up to second order was applied, in order to capture the non-linear behavior of the target by extending the initial set of predictors. It is common practice to generate additional feature incorporating non-linear interactions between the independent variables $\mathbf{x_i}$. Hence, the dimensions of the feature space is increased to $d = 35$ due to the linear combination and squaring of all features. In general, the polynomial expansion is a well-established approach when applying linear regression to non-linear data, and it does also have a positive impact on the performance of non-linear regression algorithms.

Feature importance measures are examined for the examination of the models' transparency and the effect of the feature expansion/elimination. The TEM provide an intrinsic estimate by the mean decrease impurity (MDI) method which was proposed by Breiman (2001). MDI determines each feature importance as the decrease in node impurity weighted by the probability of reaching that node. The probability can be obtained by dividing the number of samples reaching this node by the total samplesize. Averaging the estimates of the predictive ability over several decision trees leads to a variance decrease of the importance measure and thus it can be used for feature selection. Since neural networks are in general considered as opaque and do not provide a built-in feature importance estimate, the permutation feature importance method is applied. Permutation or mean decrease accuracy (MDA) feature importance is an iterative, model-agnostic algorithm and indicates the impact of one particular feature on the model score by disregarding it during training. The method was proposed by Breiman (2001) for random forests and it is considered trivial that the computational effort scales with the dimensions of the feature space. However, the estimates are not affected by multicollinearity as in the case of the MDI approach. Ultimately, it is stressed that both methods are not directly comparable, but complement each other.

The increase in dimensions of the feature space leads to sparsity and machine learning algorithms are negatively affected by the so-called *curse of dimensionality* (for mathematical details cf. Hastie et al. (2009)). Hence, the number of features will be reduced to the most important subset in three iterations: (1) Initially, the most unrelated features are dropped based on a chosen threshold in view of the individual feature importances. (2) The tree-based methods 'vote' via recursive feature elimination by their cross validation score on the discard of features, noticing that Recursive Feature Elimination (RFE) was proposed by Guyon et al. (2002). It is a selection method that fits a model and removes the weakest feature until the loss function converges. (3) The previous method is only based on the TEM and thus on the MDI feature importance estimate which is affected by collinear features, which will inevitably result from the applied feature expansion technique. The performance of MLP are negatively influenced by the occurrence of multicollinearity. For this problem, the iterative variance inflation factor (VIF) method is used for eliminating highly correlated predictors. The VIF is defined in Eq. (7).

$$VIF = \frac{1}{1-R_i^2} \text{ with } R^2 = 1 - \frac{\Sigma(y_i-\hat{y}_i)^2}{\Sigma(y_i-\bar{y}_i)^2} \tag{7}$$

The VIF approach was proposed by Lin et al. (2011) and is a measure for the increase in variance of ordinary least squares (OLS) parameter estimates if an additional variable i is added to the linear

regression model. In Eq. (7), R^2 corresponds to the coefficient of determination and the overbar indicates the mean. The VIF is a measure of multicollinearity and the threshold for the VIF of each feature is set to $VIF < 10$.

The dataset does not require any further preprocessing such as filtering or imputation as the data shows only minor variance and no missing values. Prior to the training phase, the modified dataset is resampled and split into a training and a test set, where the test set makes up 20% of the initial data. The hyperparameter study as well as the model selection is only carried out on the training set and the considered metrics are the root mean squared error (RMSE) (cf. Eq. (8)) and the mean absolute error (MAE) which is defined in Eq. (9). It is stressed that both metrics have the same unit and smaller values indicate more accurate predictions.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{8}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{9}$$

The RMSE puts more weight on outliers and is sensitive to variance, whereas the MAE expresses the magnitude of the error regardless of the direction and thus is more robust to noise.

The *No Free Lunch Theorem* (cf. Goodfellow et al. (2016)) states that one cannot assume, that one particular architecture or parameter combination performs better on a variety of tasks; hence, it is of paramount importance to conduct a hyperparameter optimization. Instead of a uninformed method, such as exhaustive grid search, an informed and thus more elegant optimization approach by Bayesian optimization was utilized. Bayesian optimization is a sequential, semi-supervised method employing Bayes theorem for adaptively approximating the objective function and obtaining the optimum without considering any gradient information, but a surrogate of the function of interest by a Gaussian process. The algorithm requires an exploration as well as an exploitation phase and initially a prior is forced onto the loss landscape. After the function evaluation, or update, the posterior is used for adjusting the acquisition function. The next candidate is queried based on the value of the acquisition function. Bayesian optimization is advantageous for objective function that are expensive to evaluate, like the model score during hyperparameter optimization. The herein applied optimization procedure was set up as a minimization problem of the RMSE of a holdout set (20% of training data) with ten initial samples (exploration), 20 exploitative iterations maximum and the model type was a Gaussian Process with Expected Improvement as acquisition function. The overall objective was to find the optimal model capacity, i.e. the tradeoff between variance and bias. For more elaborate details about global Bayesian optimization cf. Mockus (1989).

All calculations were performed on a virtual machine on a server equipped with two Intel Xeon processors (2.2 GHz), 13 GB RAM and an NVidia Tesla K80 GPU. The utilized machine learning frameworks were scikit-learn (Pedregosa et al., 2011), XGBoost (Chen and Guestrin, 2016), GPyOpt (The GPyOpt authors, 2016) for the hyperparameter tuning and statsmodels (Seabold and Perktold, 2020) for the VIF procedure. It is stressed that both scikit-learn and XGBoost employ an optimized version of the CART algorithm for the TEM.

4. Results and discussion

4.1. Feature selection

In the present study, the *Pearson* correlation coefficient is neither suited for data exploration nor for feature selection, as it is only capable of expressing linear interdependencies. To pinpoint the *non-linear* driving factors of R_{AW}, according to the chosen models, the features are presented according to their relevance obtained from the already described mean decrease impurity (MDI) and mean decrease accuracy (MDA) method. The threshold for the *sequential* feature selection is set to 2.5% in the first iteration and is given in Fig. 3 alongside the individual feature influences on the model performance. It is noted that all values of one model sum up to one and that the results are computed using the libraries scikit-learn (Pedregosa et al., 2011) and XGBoost (Chen and Guestrin, 2016).

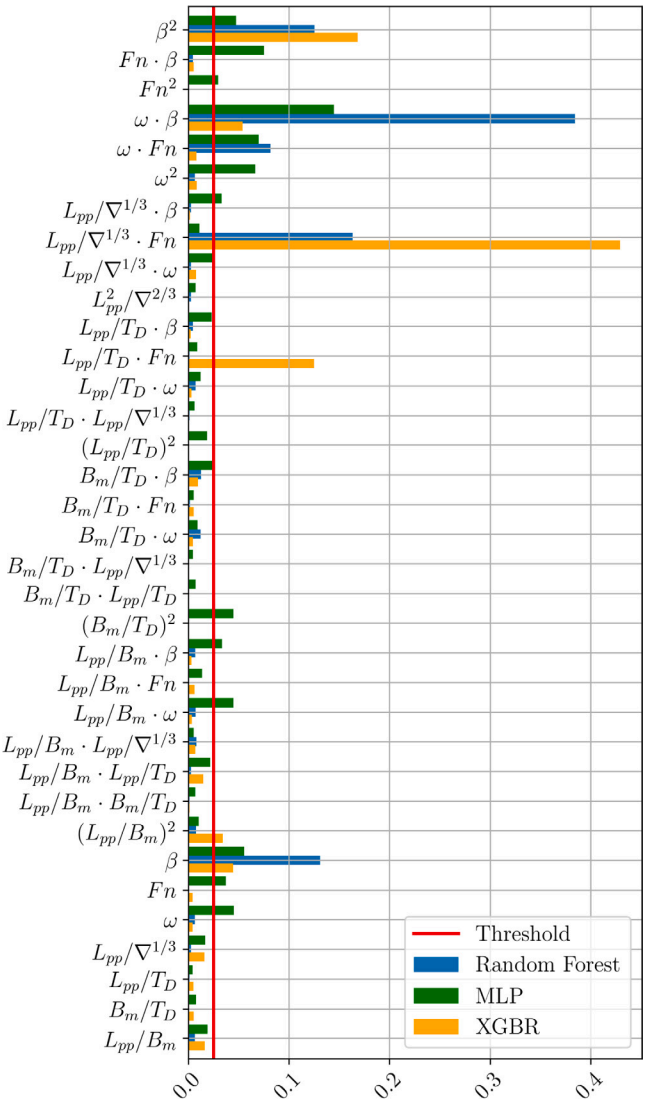


Fig. 3. Individual feature importance estimates for the expanded dataset including a threshold (red) for feature selection. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In view of Fig. 3 it is appreciated that the models' estimates agree mostly on the relevance of the predictors. Moreover, exclusively investigating the initial set of features (i.e. the original database) it is clear that the operational conditions, such as β, Fn and $\bar{\omega}$ are most influential for the models' predictions. Furthermore, Fig. 3 shows that the polynomial expansion leads on one hand to highly expressive variables as the linear combinations (e.g. $\bar{\omega}\,\beta$), but also the squared versions (e.g. β^2) are of great relevance for the models. Hence, the overall tradeoff between model accuracy and transparency is underlined as it can be seen that the machine learning models choose unphysical, highly abstract features over physical and human-readable ones. But on the other hand, the feature expansion also leads to multicollinearity. An indication for that are the discrepancies between the MDA and MDI estimates (cf. Fig. 3), where the latter

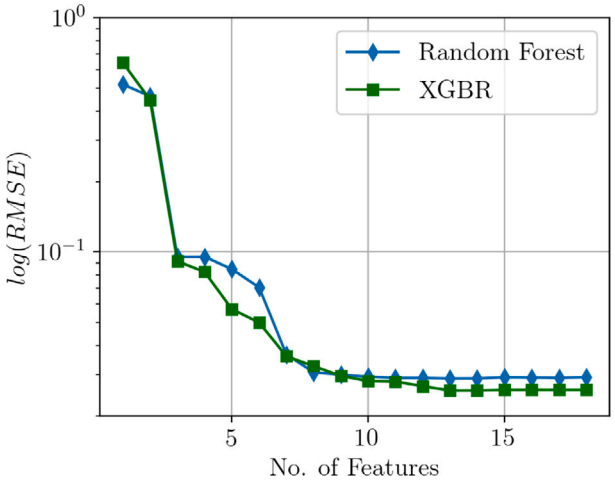


Fig. 4. Convergence study of the $RMSE$ of the validation set applying Recursive Feature Elimination (RFE) for the tree-based ensemble methods.

is affected by multicollinearity. These differences are visible for instance for $L_{pp}/\nabla^{1/3}\, Fn$ or $\bar{\omega}\,\beta$. Ultimately, the MDI estimates show general agreement with the MDA feature importance of the multilayer perceptron (MLP), underlining the sufficient transparency of the tree-based ensemble methods (TEM) at much lower computational cost. The positive effect of the feature expansion was proven and the dimensions of the feature space were reduced according to the chosen threshold of 2.5% from $d = 35$ to $d = 18$. Considering Fig. 3, the feature vector $\mathbf{x_i}$ has the following shape after the first iteration: $\mathbf{x_i} = [\beta^2, Fn\,\beta, Fn^2, \bar{\omega}\,\beta, \bar{\omega}\,Fn, \bar{\omega}^2, L_{pp}/\nabla^{1/3}\,\beta, L_{pp}/\nabla^{1/3}\,Fn, L_{pp}/\nabla^{1/3}\,\bar{\omega}, L_{pp}/T_D Fn, B_m/T_D\,\beta, (B_m/T_D)^2, L_{pp}/B_m\,\beta, L_{pp}/B_m\,\bar{\omega}, (L_{pp}/B_m)^2, \beta, Fn, \bar{\omega}]$. It is stressed that the described first iteration facilitates the second iteration and saves computational effort considerably.

The second feature selection iteration employs recursive feature elimination (RFE), i.e. it is based on the individual model MDI estimates of the TEM. For this purpose, the scikit-learn (Pedregosa et al., 2011) function including cross validation was employed. The convergence of $RMSE$ of both models applied on a cross validation set is presented in Fig. 4 for the increasing number of variable dimensions d.

Considering Fig. 4, it shows that there is good overall agreement between random forest (RF) and extreme gradient boosting regressor (XGBR) in their convergence behavior and the sensitivity to the individual features. According to both models $d = 14$ is optimal and Fn, Fn^2, $Fn\,\beta$ as well as $\bar{\omega}$ were excluded from the feature space based on the ranking that results from the feature elimination procedure. It is interesting that operational features, such as Fn and β, are disregarded in this step, even though they appear as highly relevant in Fig. 3.

Subsequently, the application of the variance inflation factor (VIF) procedure is carried out, in order to eliminate highly correlated and thus redundant features. The feature with the highest VIF is dropped and the procedure is conducted again with the updated set of features until all features exhibit $VIF < 10$. The respective VIF values for all predictors are presented for all 6 necessary iterations in Table 2. It is noted that the statsmodels library (Seabold and Perktold, 2020) was used for computing the variance inflation factor.

Lastly, the final feature vector has the following shape: $\mathbf{x_i} = [\beta^2, \bar{\omega}\,\beta, \bar{\omega}\,Fn, \bar{\omega}^2, L_{pp}/\nabla^{1/3}\,Fn, L_{pp}/T_D\,Fn, (B_m/T_D)^2, L_{pp}/B_m\,\bar{\omega}, (L_{pp}/B_m)^2]$. As only features resulting from the polynomial expansion are part of the final feature space the relevance of this step in the data preprocessing step is emphasized. Carrying out the feature selection methodology, the expanded feature space is reduced from $d = 35$ to $d = 9$ (cf. Table 2) positively affecting computational efficiency. The possible beneficial effect on model performance of the presented feature selection is studied in the following part.

Table 2
Results of VIF obtained after iterative multicollinearity analysis.

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
β^2	19.16	18.49	11.83	11.5	7.12	2.04
$\bar{\omega}\,\beta$	7.09	6.98	6.97	6.66	6.65	6.64
$\bar{\omega}\,Fn$	1.35	1.35	1.35	1.35	1.35	1.31
$\bar{\omega}^2$	3.4	3.37	3.36	3.22	3.22	3.05
$L_{pp}/\nabla^{1/3}\,Fn$	8.81	8.20	8.2	8.06	8.04	5.4
$L_{pp}/T\,Fn$	6.36	5.69	5.69	5.57	5.55	3.53
$(B_m/T_D)^2$	7.26	6.33	6.28	6.01	0.85	0.51
$L_{pp}/B_m\,\bar{\omega}$	21.22	19.34	19.32	5.17	5.17	5.14
$(L_{pp}/B_m)^2$	9.69	9.61	2.12	2.06	2.07	1.85
β	11.42	11.12	11.05	11.05	10.97	elim.
$B_m/T_D\,\beta$	50.47	15.42	15.33	14.98	elim.	elim.
$L_{pp}/\nabla^{1/3}\,\bar{\omega}$	27.6	23.98	23.98	elim.	elim.	elim.
$L_{pp}/B_m\,\beta$	137.36	25.24	elim.	elim.	elim.	elim.
$L_{pp}/\nabla^{1/3}\,\beta$	222.38	elim.	elim.	elim.	elim.	elim.

4.2. Model selection

In this section, the final model characteristics are determined using Bayesian optimization. Moreover, the individual performances of the methods will be assessed in use of the test set, but initially the baseline architectures of the considered models are presented. It is stressed that the naming convention of the employed software packages are used herein.

The baseline random forest is characterized by 100 estimators, whereas the hyperparameter `min_samples_split`, which represents the minimum number of samples required to split an internal node, is set to 2. Increasing this parameter, the tree becomes more constrained as it has to consider more samples at each node. The parameter `min_samples_leafs` is defined as 4 and it is similar to `min_samples_splits`; however, the latter describes the minimum number of samples at the leafs. The maximum depth of the tree `max_depth` is equal to 10 indicating how deep the tree can be built. The deeper the tree, the more splits and thus the more information can be captured. Furthermore, the number of features `max_features` to be considered while searching for the best split is specified as d - the number of features. In addition, either the $L1$ or $L2$ error at the leafs is minimized.

The initial XGBR model is characterized by 100 `n_estimators`, a linear loss function and a `learning_rate` of 0.1. Moreover, `max_depth` is set to 10 and the regularization term $\gamma = 0.0$. The parameter `colsample_by_tree` denotes the subsample ratio of features for each split at each level and is thus comparable to `max_features`. Also, `min_child_weight` defines the minimum sum of weights w_j of all samples required in a child node, i.e. related to `min_child_leaf` and capable to prevent overfitting.

The baseline MLP architecture comprises one hidden layer with 100 neurons including $L2$ regularization as well as a ReLU (Rectifier Linear Unit, $x = max(0, x)$) activation following input and hidden layer. The output neuron is activated by a linear function and the Adam (Adaptive Moment Estimation) optimizer (cf. Kingma and Ba (2015)) minimizes the $L2$ error w.r.t. the learning rate.

Throughout the parameter study, the `hidden_layer_sizes`, the `activation_function` as well as the $L2$ penalty or regularization term α are optimized for the MLP. All mentioned hyperparameters of RF and XGBR are variables during optimization. The baseline and final model hyperparameters as well as their definition range are provided in Table 3. The default values are adopted from the chosen software packages scikit-learn (Pedregosa et al., 2011) and XGBoost (Chen and Guestrin, 2016), whereas the ranges are located around them.

The Bayesian optimization was fed with 10 initial explorative samples and ran for 20 iterations during the exploitation optimizing the $RMSE$ of a cross validation set. In view of Table 3, it is evident that the resulting parameters mostly correspond to either the upper parameter bound or the baseline value. For instance, `n_estimators`

Table 3
Hyperparameters of the three examined models including baseline and bounding values.

	Hyperparameter	Baseline	Range	Result
RF	`n_estimators`	100	[50, 150]	150
	`max_depth`	10	[1, 30]	30
	`min_sample_leafs`	4	[1, 5]	1
	`min_sample_split`	2	[2, 10]	2
	`max_features`	d	$\{d, \sqrt{d}, \log_2(d)\}$	d
	`criterion`	MSE	$\{MSE, MAE\}$	MSE
XGBR	`n_estimators`	100	[50, 150]	150
	`learning_rate`	0.1	[0.05, 0.1]	0.1
	`max_depth`	10	[1, 30]	20
	`min_child_weight`	1	[1, 10]	1
	γ	0	[0.05, 5]	0.1
	`colsample_by_tree`	1	[0.5, 0.8]	0.8
MLP	`hidden_layer_sizes`	(100)	{(100), (50, 50, 50), (50, 100, 50), (100, 100)}	(50, 100, 50)
	`activation`	ReLU	{ReLU, sigmoid, tanh}	ReLU
	α	10^{-4}	$[10^{-4}, 0.05]$	10^{-4}

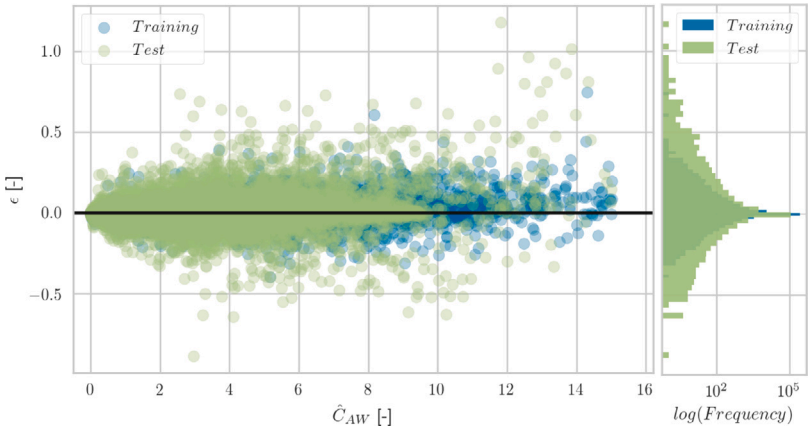


Fig. 5. Residuals plot for the RF on the $C_B = 0.84$ database. Note that the histogram (right-side plot) of the residuals presents the occurrence frequency in logarithmic scale.

in case of the TEM and `hidden_layer_sizes` reach their upper constraint after the optimization for the largest model complexity. On first sight this fact may increase the risk of overfitting (insufficient model capacity), but the cross validation score is taken into account in the parameter study in order to prevent overfitting. The possible occurrence of overfitting and the overall model performance will be assessed in the following.

In Figs. 5–7 the residuals ϵ of the models' predictions are presented for both the training and test datasets, accompanied by the histogram of the residuals in logarithmic scale. It is stressed that only the results for the $C_B = 0.84$ dataset are presented herein and the corresponding results for the $C_B = 0.58$ dataset are placed in the appendix. It is evident that all three models are biased towards underpredicting the non-dimensional added resistance C_{AW} as the histograms are shifted towards positive ϵ values. Furthermore, considerable heteroskedasticity is noticeable for all three models, i.e. an increase in variance with increasing magnitude of the target variable. This phenomenon is also observed in the residual analysis of the semi-empirical formula of Liu and Papanikolaou (2020) and results from the non-linear behavior of the added-wave resistance. However, the heteroskedasticity is most pronounced for the MLP and decreases for the TEM. Moreover, we see the largest bias and variance for the MLP. The RF on the other hand, exhibits robust estimates with little noise; however, it is conspicuous that the test predictions are more biased towards underpredicting the target than the training predictions. The discrepancies in the residual distributions are not as pronounced for the XGBR and the MLP. Whether this is an indication of overfitting in case of the random forest will be assessed in the following analysis of the metrics of the test predictions. It is noted that a minor discrepancy between training and test performance is expected as it is unseen data.

The models are assessed based the metrics $RMSE$ and MAE using the out-of-sample (test) dataset in Table 4. It is stressed that (a) corresponds to the application of the Yeo–Johnson-Transformation and the described data preprocessing method. In addition, the performance on a dataset (b) without the described preprocessing methodology, i.e. only normalizing the target according to the extreme values and the use of the initial set of features is presented. Hence, the effect of the preprocessing methodology is examined. In addition, the performance of the chosen baseline models is given in order to underline the increase in accuracy resulting from the conducted parameter study. It is noted that Table 4 presents the model performance for both datasets and that the metrics for the in-sample set for both datasets are located in the appendix of this contribution.

The observations from Table 4 are two fold: (1) The performance increase resulting from the hyperparameter optimization and (2) the effect of the data preprocessing described in the previous section. Regarding the first point, the MLP model experiences the largest performance gain of 72.7%, whereas the RF and XGBR achieve an increase of 40.05% and 11.25%, respectively. It is stressed that the presented percentage increase is an average for both metrics and databases. Regarding the second point, it shows that the described data preprocessing method leads to superior model accuracy in comparison to the dataset without the applied methodology (b). The performance of the XGBR is affected the most with 170.5% higher accuracy relating (a) to (b). The MLP as well as the RF yield a performance increase of 35.51% and 36.87%, respectively. For the sake of clarity the results of Table 4 are presented in Fig. 8 in histograms.

Ultimately, the RF achieves the highest accuracy according to both metrics (bold in Table 4). Whether this is an indication of the superior generalization capability will be assessed in the model validation

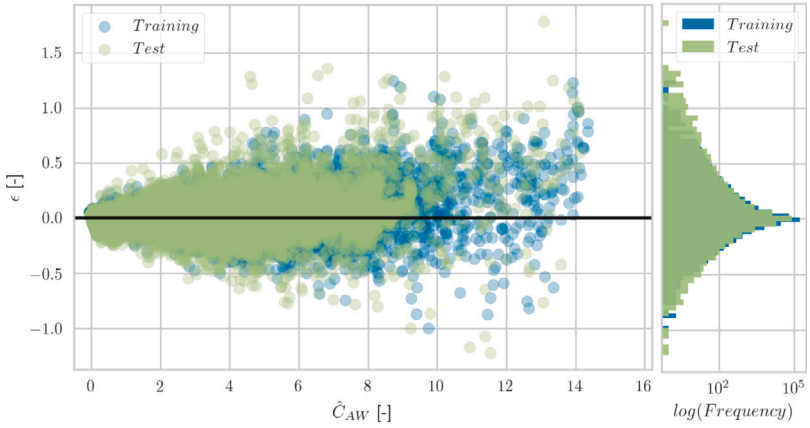


Fig. 6. Residuals plot for the XGBR on the $C_B = 0.84$ database. Note that the histogram (right-side plot) of the residuals presents the occurrence frequency in logarithmic scale.

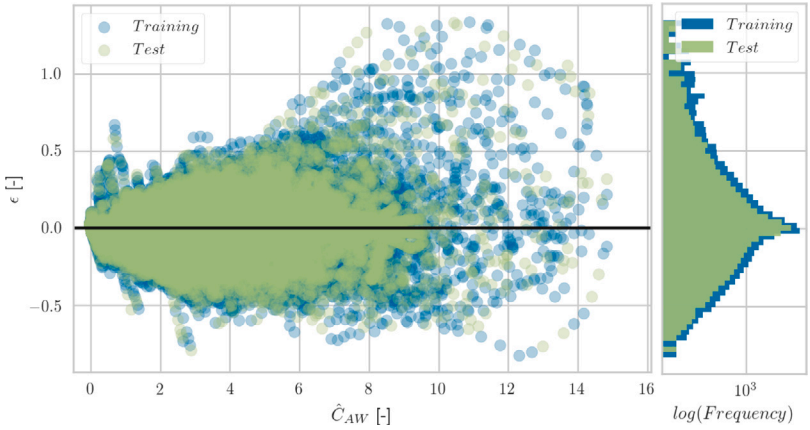


Fig. 7. Residuals plot for the MLP on the $C_B = 0.84$ database. Note that the histogram (right-side plot) of the residuals presents the occurrence frequency in logarithmic scale.

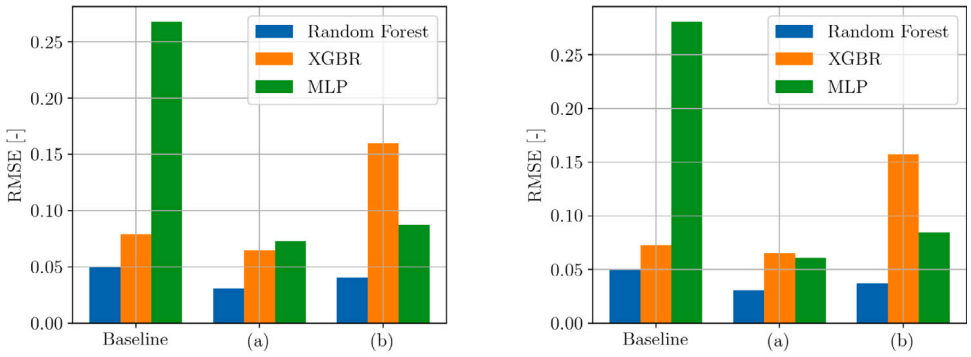
Table 4
Metrics of predictions using the test dataset for both datasets.

		$C_B = 0.84$			$C_B = 0.58$		
		RF	XGBR	MLP	RF	XGBR	MLP
Baseline	$RMSE$	0.0502	0.0789	0.2680	0.0497	0.0727	0.2808
	MAE	0.0130	0.0343	0.1304	0.0130	0.0328	0.1392
(a)	$RMSE$	**0.0306**	0.0645	0.0730	0.0310	0.0647	0.0611
		−38.98%	−18.27%	−72.74%	−37.61%	−11.04%	−78.25%
	MAE	**0.0075**	0.0310	0.0393	0.0076	0.0308	0.0416
		−41.83%	−9.65%	−69.84%	−41.78%	−6.07%	−70.11%
(b)	$RMSE$	0.0407	0.1597	0.0878	0.0369	0.1573	0.0848
		−18.85%	102.57%	−67.23%	−25.70%	116.27%	−69.80%
	MAE	0.0101	0.0921	0.0578	0.0122	0.0906	0.0566
		−22.39%	168.60%	−55.65%	−5.69%	175.96%	−59.38%

section. However, it proves that the RF did not overfit on the in-sample data despite the discrepancies of the residual histograms found in Fig. 5. In addition, it can be observed that the RF shows the smallest variance in model performance, cf. Fig. 8. Hence, it is concluded that the RF is in this case not only the most accurate regressor, but also the most robust one. The MLP on the other hand shows a considerable increase in performance after the parameter study and also the data preprocessing method affects the results noticeably. For the XGBR, the parameter study has only a small influence on performance and the poor performance on the (b) dataset is conspicuous. It is assumed that the dispersed target distribution (cf. upper plot in Fig. 2) is the reason for this performance drop. Outliers have a negative impact on boosting as each new tree builds on previous tree residuals. Thus boosting ensemble methods have a disproportionate bias towards outliers as they have much larger residuals than other samples. Still, the insufficient performance on (b) is higher than expected. Lastly, one can conclude that it is justifiable to only optimize the models using one dataset ($C_B = 0.84$) as both the performance increase due to the hyperparameter optimization and the data preprocessing are similar.

4.3. Model validation

In the following, the models are validated on three case studies against publicly available model test and computational fluid dynamics results — RANS in particular. Additionally, the estimates of the DTU Design Tool and a semi-empirical formula proposed by Liu and Papanikolaou (2020) are considered. The following hull forms were selected as benchmarks from the literature: The full-type KVLCC2 (KRISO Very Large Crude oil Carrier) is investigated in head waves.



(a) Results for the $C_B = 0.84$ database. (b) Results for the $C_B = 0.58$ database.

Fig. 8. Histogram of the $RMSE$ of the model predictions applied on the test set.

Table 5
Main particulars and operational conditions of the three case studies adopted from Sadat-Hosseini et al. (2013), Fujii and Takahashi (1975) and Kim et al. (2021).

	Unit	KVLCC2	S175	LNG-C
L_{pp}	[m]	320.0	175.0	290.0
B_m	[m]	58.0	25.4	45.0
T_D	[m]	20.8	9.5	11.5
C_B	[–]	0.808	0.561	0.79
k_{yy}/L_{pp}	[–]	0.25	0.24	0.25
Fn	[–]	0.142	0.15	0.19
β	[deg.]	180	150	120

On the other hand, the S175 container ship Fujii and Takahashi (1975) and the LNG carrier of Kim et al. (2021) are considered in bow oblique waves. The main particulars of the ships as well as their forward speed and relative wave direction are provided in Table 5.

As can be seen in Table 5, one slender and two full-type vessels were selected for model validation. All geometries remain within the definition range of the database (cf. Table 1), but do not exactly match the scaling ratios of the initial hull forms; thus, the generalization capability of the data-driven models can be assessed in a proper fashion. Besides the agreement in the magnitude of C_{AW}, the adequate representation of the resonant conditions is of high interest as the mean added resistance $\bar{R}_{AW}$ in irregular waves is a function of the overlap between the individual peaks of the prevalent wave energy density spectrum and the quadratic transfer function.

Initially, the blunt KVLCC2 tanker is considered in head waves $\beta = 180°$ at $Fn = 0.142$. In Fig. 9, the estimated quadratic transfer functions of C_{AW} are presented vs. the relative wavelength λ/L_{pp}. The experimental fluid dynamics (EFD) samples of Lee et al. (2013) and Sadat-Hosseini et al. (2013) as well as the RANS results of the latter publication are considered for validation.

In view of Fig. 9, it is stated that all estimates of the ML models are in accordance with the DTU Design Tool which is expected as the underlying database is the same. Moreover, the models show quantitative agreement with the CFD as well as EFD samples in long waves and qualitative agreement with their trend in the short wave regime. However, both the resonance frequency and the corresponding magnitude are not very well captured. Lastly, the semi-empirical formula provides the most appropriate estimate in comparison to the model test and CFD results. On a sidenote, the two separate experimental datasets show minor variability or rather uncertainty - especially near resonance. In addition, the RANS results differ notably from the experimental data in long and short waves.

The slender S175 container ship is examined in bow oblique waves ($\beta = 150°$) at $Fn = 0.15$. In Fig. 10, the data-driven models are validated against the RANS results of Uharek (2019) and the experimental data of Fujii and Takahashi (1975).

As can be inferred from Fig. 10, none of the ML methods captures the resonance region correctly as the individual peaks deviate in both wavelength and in magnitude. Furthermore, the ML models show a large variation in their accuracy. The MLP exhibits good agreement in long and short waves and it is evident that the tree-based ensemble methods (TEM) provide lower C_{AW} estimates in the proximity of resonance compared to the corresponding MLP values. The DTU Design Tool and the semi-empirical formula by Liu and Papanikolaou (2020) are in general in agreement with the numerical and experimental results. Once again, the discrepancy between RANS and EFD data is noted and the TEM are not able to reflect the resonant conditions correctly.

The individual results of the twinscrew LNG carrier, which is examined in oblique waves $\beta = 150°$ and moderate forward speed conditions ($Fn = 0.19$), are provided in Fig. 11. The benchmark data is established by numerical and experimental data adopted from Kim et al. (2021).

The immediate observation from Fig. 11 is the discrepancy between RANS and EFD data. In addition, the MLP, the DTU Design Tool as well as the semi-empirical formula by Liu and Papanikolaou (2020) show matching estimates — in resonant conditions in particular. However, the results of the TEM deviate both quantitatively and qualitatively in this regime. The data of the tree-based ensemble methods shows minor irregularities due to the fact that TEM only provide piecewise approximations. The satisfactory agreement of the ML models and DTU Design Tool is evident in short waves, where the semi-empirical formula of Liu and Papanikolaou (2020) yields higher values.

4.4. General discussion

The herein examined data-driven models were able to represent the frequency coherent character of the added-wave resistance. The validation shows sufficient performance for the MLP and highlights its applicability as a design aid and surrogate model. The TEM on the other hand show deviant results and high bias in resonant conditions when applied on unseen data. In long and even short waves, the agreement is considered as reasonable in most cases. The MLP shows overall sufficient resemblance to the two established prediction methods, i.e. the DTU Design Tool (Nielsen, 2015; Martinsen, 2016) and the semi-empirical formula by Liu and Papanikolaou (2020). As the added-wave resistance is generally considered as a non-viscous phenomenon, the application of potential theory methods is appropriate. However, considering further perspectives and the application to real data, all of the underlying methods are subject to the limitations of potential theory, such as the assumption of linearity including only small perturbations, small amplitudes, moderate wave steepness and the absence of viscous effects. Salvesen's method disregards the hull geometry above the waterline and includes only the Neumann–Kelvin linearized interaction between the steady and the unsteady flows. In

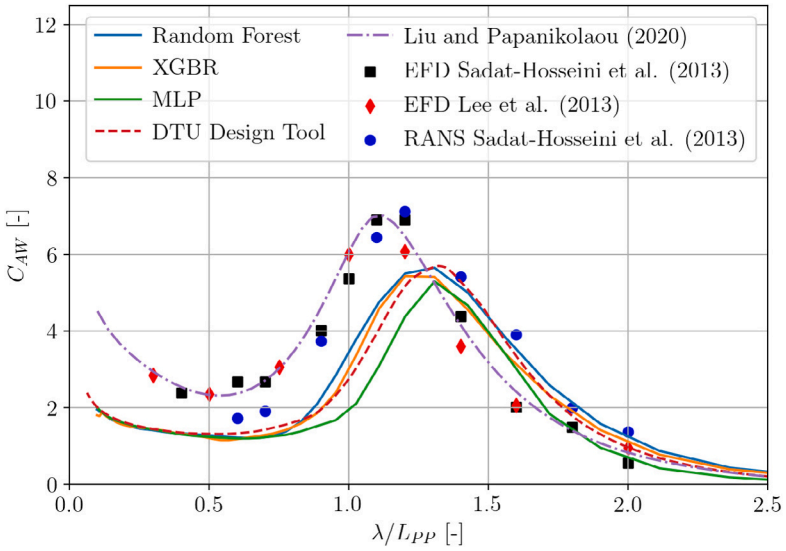


Fig. 9. Non-dimensional added-resistance C_{AW} of the KVLCC2 tanker for $\beta = 180°$ and $Fn = 0.142$.

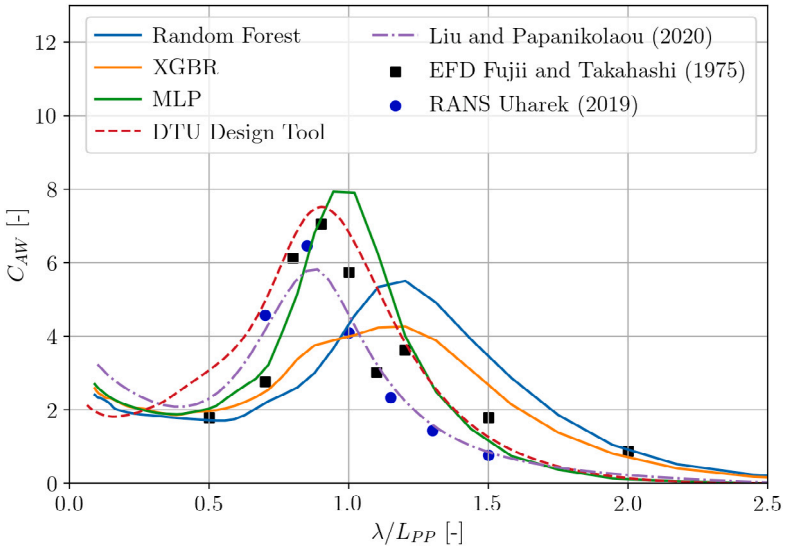


Fig. 10. Non-dimensional added-resistance C_{AW} of the S175 containership for $\beta = 150°$ and $Fn = 0.15$.

addition it is based on strip theory and inherits the corresponding assumptions and approximations. Similarly, Faltinsen's formula relies on the assumption that the ship has vertical walls, i.e. no flare is considered. Ultimately, the TEM match the abilities of the MLP in numerous characteristics, for instance capturing non-linear interactions between the predictors or handling the transition from Salvesen's to Faltinsen's estimate in the underlying database, but essentially the fundamental tradeoff in machine learning between transparency and accuracy, or rather in generalization capability, manifests itself in the present study through the transparent TEM showing inferior results compared to the opaque MLP. The MLP shows higher model generalization due to its multilayered architecture, but as the number of hidden layers increases the estimates and the mathematics behind them get more complex and opaque. The TEM on the other hand, have a clear and transparent structure and provide not only a decision path for every prediction, but also a direct estimate of each feature's predictive power. Still, the ensemble methods are affected by the fundamental limitations of decision trees and the CART implementation. On one hand, the CART algorithm is known for its transparency and simplicity as it grows binary tree graphs. In turn, this comes at the cost of instabilities and inaccuracies as soon as the estimator is applied on unseen data. This instability is due to the hierarchical structure because the effect of an imperfection in the top split is propagated down to all child nodes. The MARS algorithm (cf. (Mittendorf and Papanikolaou, 2021)) has a mathematical resemblance to the CART algorithm and alleviates its lack of smoothness, but is not as scalable and hence not considered in the present work. Lastly, considerable variance in the estimates of the added-wave resistance was noticeable, not only for data-driven models, but also in the numerical results and even in the experimental data uncertainty was present, cf. Fig. 9. This reminds us that both the experimental and numerical determination of added-wave resistance, in particular near resonance and in oblique waves, is challenging. It could be even argued that a deterministic prediction is not *yet* suited for practical applications. Park et al. (2015) provide several sources of uncertainty regarding experiments pertaining to the added-wave resistance. This motivates dedicated studies for the uncertainty quantification in the direction of predicting the added-wave resistance using both numerical and empirical methods. From the literature it is evident that the uncertainty of added resistance is, indeed, well-known, but *not* well characterized.

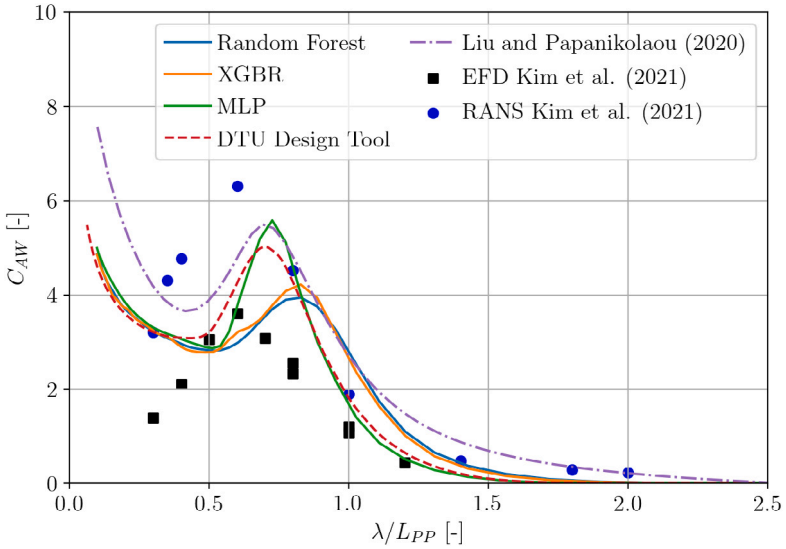


Fig. 11. Non-dimensional added-resistance C_{AW} of the LNG carrier for $\beta = 120°$ and $Fn = 0.19$.

5. Conclusions

The prediction of the second order steady wave forces on ships is a classical hydrodynamics problem and has been addressed in this contribution using different non-linear machine learning techniques. This study proposed a data-driven methodology for the estimation of the non-dimensional added-wave resistance C_{AW} based on results obtained from three different numerical calculation methods based on potential flow theory. The multivariate regression approach was focused on the comparison of tree-based ensemble methods and artificial neural networks in the context of prediction accuracy, robustness and transparency. Moreover, the data preprocessing and hyperparameter optimization were key features of the suggested methodology. As a result, all applied methods were able to produce a non-linear and consistent mapping from abstract features to the target C_{AW} with sufficient accuracy on in- and out-of-sample data. The application to real world case studies revealed insufficient results of the tree-based ensemble methods compared to the neural network. In comparison to well established design aids, the study proves that the multilayer perceptron is able to generate C_{AW} estimates of similar quality. Lastly, the present work highlights that the reliable and accurate prediction of the added-wave resistance is still a challenging task for all examined methods and will remain a competitive and demanding research topic in the coming future.

Addressing the research question from the paper's introduction, it appears that tree-based ensemble methods show comparable or even higher precision than the multilayer perceptron on in-sample data; however, the final validation study revealed that tree-based ensemble methods are rigid in their application and become unstable applied on new data differing from the initial discretization or distribution. This is a result of the inherent drawbacks of the Classification and Regression Trees (CART) algorithm as its simplicity comes at the cost of instability and inaccuracies as the weak learners tend to become overly complex. In fact, this leads to decreased generalization capability and overfitting. Moreover, tree-based ensemble methods show high bias in the sense that they underestimate the magnitude of the added-wave resistance under resonant conditions for most cases. As already noted, tree-based ensemble methods do not perform sufficiently well on sparse data, therefore the minimization of epistemic uncertainty is pivotal for the improvement of the model performance. Thus, the generation of an updated database with additional parameter combinations and parameter variation is an important aspect for future work. In addition, the consideration of different longitudinal radii of gyration k_{yy} and block coefficients C_B values is required for improving the prediction of the added-wave resistance and the corresponding resonance frequency. Lastly, the calibration on experimental data of the models using transfer learning is a promising aspect of future work.

The implicit feature processing capabilities of deep neural networks make their application appealing as the herein labor intensive preprocessing is not required using these advanced architectures. Considering deep learning methods and their ability handling high-dimensional tensors, it is possible to utilize the hull shape itself as input as has been proposed in the study by Taniguchi and Ichinose (2020). They employ a Generative Adversarial Network for the generation of novel hull shapes and a convolutional neural network as a discriminator trained on CFD results of the wake profile. Such sophisticated layer types and architectures have immense potential for future work regarding the development of initial design aids and also hull shape optimization approaches.

CRediT authorship contribution statement

Malte Mittendorf: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Validation, Visualization, Writing – original draft & editing. **Ulrik D. Nielsen:** Writing – review & editing, Supervision, Project administration, Resources, Funding acquisition. **Harry B. Bingham:** Writing – review & editing, Supervision, Resources, Project administration.

Declaration of competing interest

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

Acknowledgments

The financial support from The Danish Maritime Fund (Projekt 2019-043), Orients Fond and Department of Mechanical Engineering, Denmark, Technical University of Denmark is highly appreciated. The second author has received funding by the Research Council of Norway through the Centres of Excellence scheme, project number 223254 AMOS.

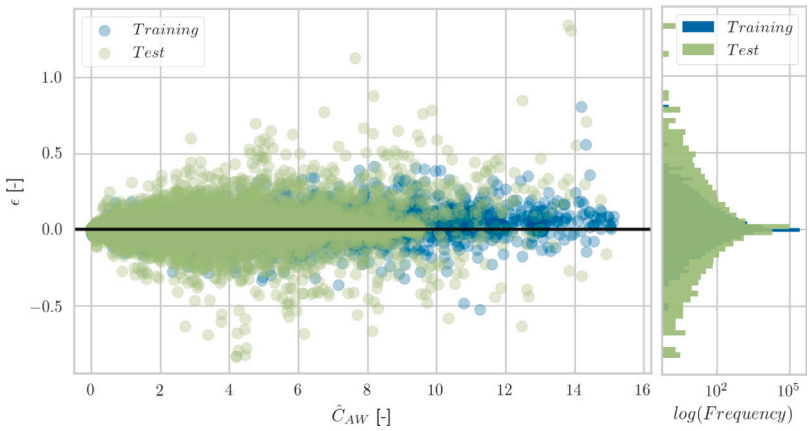


Fig. 12. Residuals plot for the RF on the $C_B = 0.58$ database.

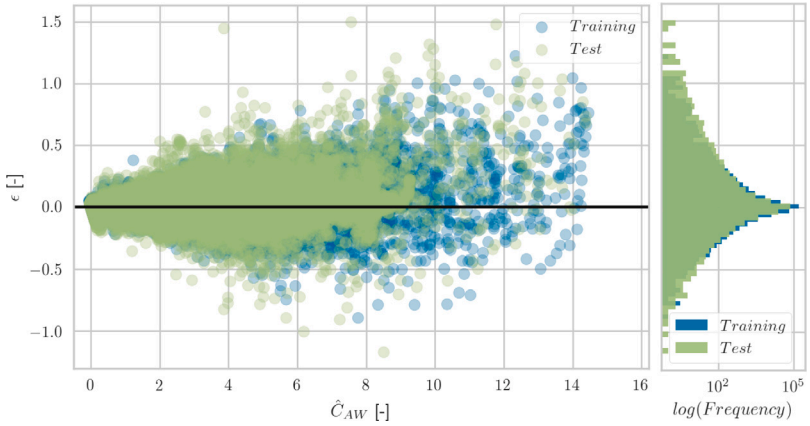


Fig. 13. Residuals plot for the XGBR on the $C_B = 0.58$ database.

Table 6
Metrics for the prediction of the train dataset for both databases.

		$C_B = 0.84$			$C_B = 0.58$		
		RF	XGBR	MLP	RF	XGBR	MLP
Baseline	$RMSE$	0.0322	0.0634	0.2674	0.0318	0.0596	0.2844
	MAE	0.0083	0.0295	0.1299	0.0082	0.0285	0.1393
(a)	$RMSE$	0.0121	0.0492	0.0726	0.0121	0.0484	0.0605
	MAE	0.0029	0.0258	0.0392	0.0029	0.0256	0.0414
(b)	$RMSE$	0.0154	0.1439	0.0877	0.0150	0.1423	0.0840
	MAE	0.0043	0.0870	0.0577	0.0054	0.0859	0.0563

Table 7
Main particulars of the additional four case studies taken from Valanto and Hong (2015), Shigunov et al. (2018), Yasukawa et al. (2019) and Park et al. (2019).

	Unit	HSVA	DTC	S-Cb84	SNU
L_{pp}	[m]	220.32	355.0	178.0	323.0
B_m	[m]	32.04	51.0	32.26	60.0
T_m	[m]	7.2	14.5	11.57	21.0
C_B	[-]	0.654	0.661	0.84	0.83
k_{yy}/L_{pp}	[-]	0.263	0.27	0.25	0.258
Fn	[-]	0.233	0.052	0.0	0.137
β	[deg.]	120	180	150	150

Appendix

A.1. Performance on training dataset

See Table 6.

A.2. Residuals plots for the $C_B = 0.58$ database

See Figs. 12–14.

A.3. Additional validation cases

In addition to the three validation cases in Section 4, the models are validated for four other case studies. The ships are selected such that they cover different operational conditions and hull shapes. The main particulars of the hulls are presented in Table 7 and it is evident that two slender and two full-type ships are chosen.

Initially, the slender HSVA (Hamburg Ship Model Basin) cruise ship is considered in oblique waves $\beta = 120°$ and relatively high forward speed $Fn = 0.233$. In Fig. 15, the transfer function of C_{AW} is presented for the normalized wavelength λ/L_{pp} alongside the experimental samples of Valanto and Hong (Valanto and Hong, 2015).

As can be inferred from Fig. 15, all three ML models are consistent with both model test results and the semi-empirical formula. However, it is conspicuous that none of the methods capture the resonance region correctly as the individual peaks deviate in both wave length and in magnitude. It is evident that the TEM provide lower C_{AW} estimates in the proximity of resonance compared to the MLP. Lastly, the MLP provides the most accurate ML prediction considering both model test data and the formula of Liu and Papanikolaou (2020). One possible explanation for the deviations in case of the HSVA cruise ship could be that the k_{yy}/L_{pp} is kept constant in the database at 0.25 and the

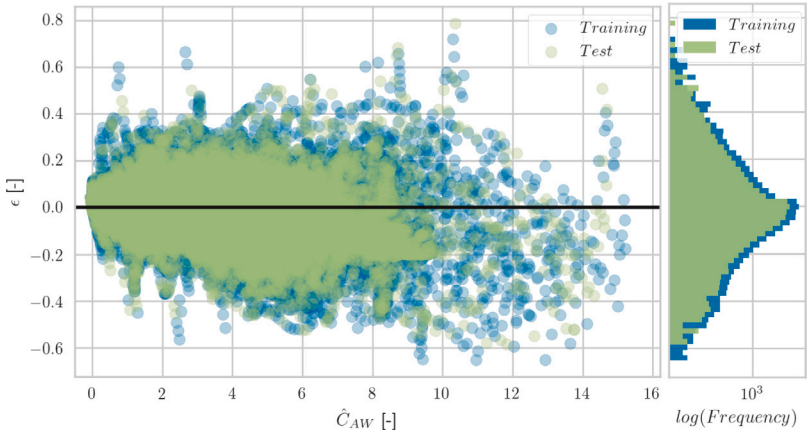


Fig. 14. Residuals plot for the MLP on the $C_B = 0.58$ database.

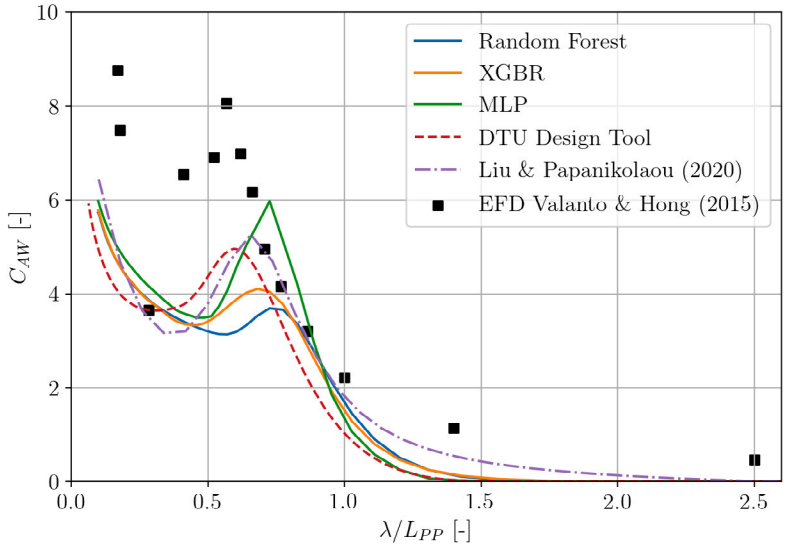


Fig. 15. C_{AW} for the HSVA cruise ship for $\beta = 120°$ and $Fn = 0.233$.

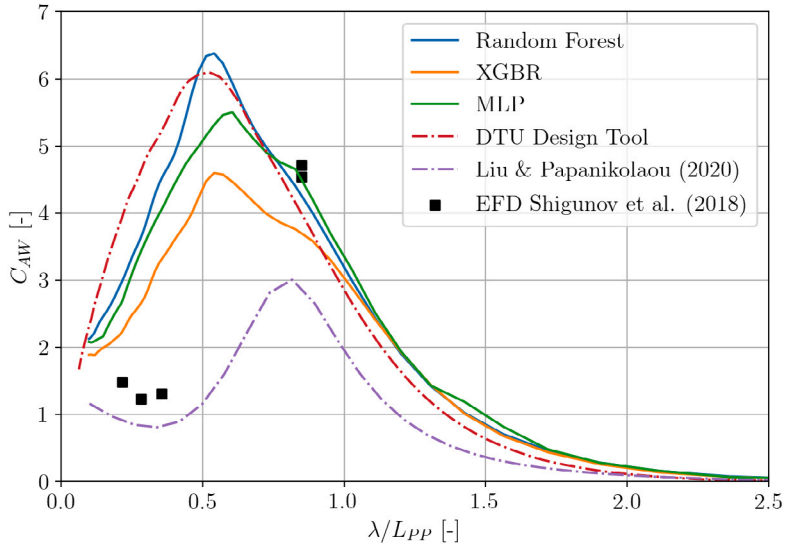


Fig. 16. C_{AW} for the DTC container ship for $\beta = 180°$ and $Fn = 0.052$.

cruise ship has a $k_{yy}/_{Lpp} = 0.263$, (Valanto and Hong, 2015). In addition, the loading condition (i.e. GM) is neither variable in the semi-empirical formula of Liu and Papanikolaou (2020) nor in the used dataset. The weight distribution in general is a considerable source of

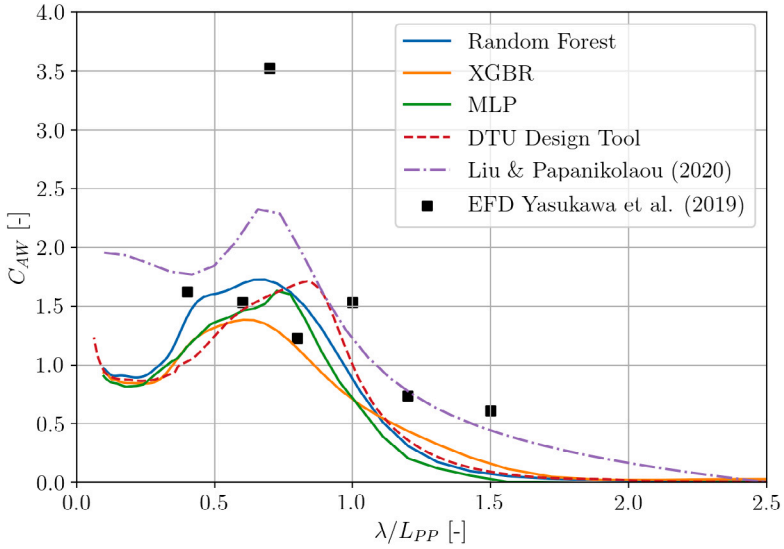


Fig. 17. C_{AW} for the S-Cb84 bulk carrier for $\beta = 150°$ and $Fn = 0.0$.

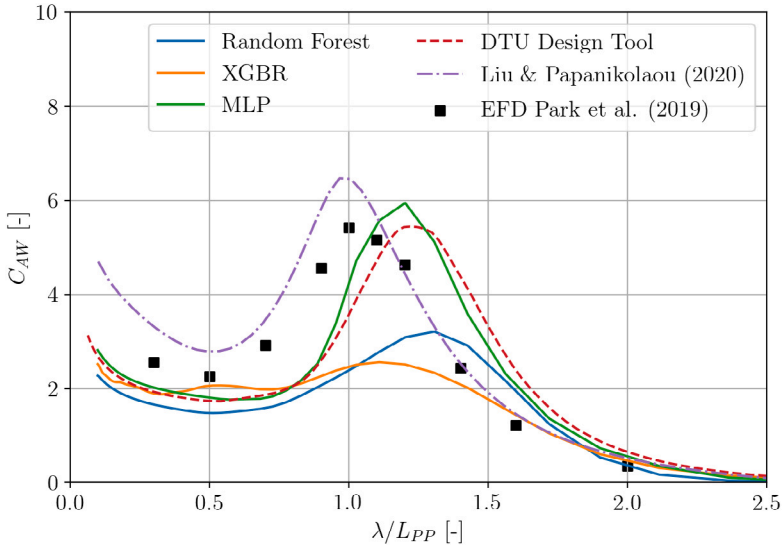


Fig. 18. C_{AW} for the SNU tanker for $\beta = 150°$ and $Fn = 0.137$.

uncertainty during model tests as Park et al. (2015) mention. Moreover, the occurrence of breaking waves as reported by Valanto and Hong (2015), the relatively high forward speed and potential wall reflections for oblique waves could be additional driving factors of experimental uncertainty.

The slim-shaped Duisburg Test Case (DTC) was studied extensively for its manoeuvering characteristics in waves during the SHOPERA EU project. The case with head waves ($\beta = 180°$) and relatively low forward speed with $Fn = 0.052$ was examined. The results for this case are presented in Fig. 16 with the samples from Shigunov et al. (2018) as benchmark.

In view of Fig. 16, it is assumed that the ML models provide reasonable estimates in the regime of long waves due to the asymptotic behavior, but in short wavelengths drastic deviations are noticeable in comparison to the benchmark results. Conversely, the semi-empirical formula captures the added-wave resistance in short waves sufficiently, but underestimates the magnitude for the resonant frequencies. Also, the estimates of the TEM – the XGBR in particular – show considerable variance due the fact that TEM only yield piecewise approximations.

In the following, a case study without forward speed ($Fn = 0.0$) is investigated with the full-type S-Cb84 bulk carrier in oblique waves $\beta = 150°$ and the results are presented in Fig. 17 together with the model test data of Yasukawa et al. (2019).

As can be inferred in Fig. 17, the estimates of the ML models are qualitatively in accordance with the benchmark data and the Liu Papanikolaou formula. Moreover, the semi-empirical method shows superior accuracy in long waves, whereas the ML methods exhibit good agreement in shorter waves.

The last test case is the SNU (Seoul National University) tanker, which is characterized by high fullness, in oblique waves $\beta = 150°$ and medium forward speed conditions ($Fn = 0.137$). The experimental data of Park et al. (2019) is considered for validation.

The ML models show a large variance in their accuracy, viz. Fig. 18. In particular, the results of the TEM deviate dramatically both quantitatively and qualitatively. The predictions of the MLP match the magnitude of the experimental data, but the peak frequencies differ slightly. Once again, the data of the TEM shows irregularities in short waves.

References

Bertram, V., Mesbahi, E., 2000. Applications of neural nets in ship design. In: Jahrbuch der Schiffbautechnischen Gesellschaft. Springer.

Bishop, C., 2006. Pattern Recognition and Machine Learning. Springer Verlag, New York, http://dx.doi.org/10.1117/1.2819119.

Boese, P., 1970. Eine Einfache Methode Zur Berechnung Der Widerstandserhoehung eines Schiffes im Seegang, Vol. 258. Schriftenreihe Schiffbau, pp. 1–9. http://dx.doi.org/10.15480/882.648.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. http://dx.doi.org/10.1023/A:1010933404324.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and regression trees. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, ISBN: 9780412048418.

Cepowski, T., 2020. The prediction of ship added resistance at the preliminary design stage by the use of an artificial neural network. Ocean Eng. 195, http://dx.doi.org/10.1016/j.oceaneng.2019.106657.

Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, http://dx.doi.org/10.1145/2939672.2939785.

Faltinsen, O.M., 2005. Hydrodynamics of High-Speed Marine Vehicles. Cambridge University Press, http://dx.doi.org/10.1017/CBO9780511546068.

Faltinsen, O.M., Minsaas, K.J., Liapis, N., Skjordal, S.O., 1980. Prediction of resistance and propulsion of a ship in a seaway. In: Inui, T. (Ed.), Proceeding of 13th Symposium on Naval Hydrodynamics. The Shipbuilding Research Association of Japan, Tokyo (Japan), pp. 505–529.

Fujii, H., Takahashi, T., 1975. Experimental study on the resistance increase of a large full ship in regular oblique waves. J. Soc. Naval Archit. Japan 1975 (137), 132–137.

Gerritsma, J., Beukelman, W., 1971. Analysis of the Resistance Increase in Waves of a Fast Cargo Ship.. Technical Report 334, Shipbuilding laboratory, Technological University Delft, Delft, The Netherlands.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. The MIT Press, ISBN: 0262035618.

Herradon de Grado, E., Bertram, V., 2016. Predicting added resistance in wind and waves employing artificial neural nets. In: In Proc. of the 1st Hull Performance & Insight Conference, HullPIC'16. Pavone.

Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46 (1), 389–422. http://dx.doi.org/10.1023/A:1012487302797.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. Springer Science+Business Media, New York, http://dx.doi.org/10.1007/978-0-387-84858-7.

International Maritime Organization, 2011. Amendments To the Annex of the Protocol of 1997 to Amend the International Convention for the Prevention of Pollution from Ships, 1973, As Modified By the Protocol of 1978 Relating Thereto, Vol. 203 no. 62. Resolution MEPC.

Kim, T., Yoo, S., Kim, H.J., 2021. Estimation of added resistance of an LNG carrier in oblique waves. Ocean Eng. 231, 109068. http://dx.doi.org/10.1016/j.oceaneng.2021.109068.

Kingma, D.P., Ba, J., 2015. Adam: a method for stochastic optimization. In: Proc. of 3rd International Conference on Learning Representations, ICLR '15, San Diego.

Kristensen, H.O.H., 2010. Model for environmental assessment of container ship transport. Soc. Nav. Archit. Mar. Eng. Trans. 118, 122–139. http://dx.doi.org/10.5957/SMC-2010-T41.

Lee, C.H., Newman, J.N., 2013. Wamit user manual. http://www.wamit.com/manual.html version 7.0.

Lee, J.H., Seo, M.G., Park, D.M., Yang, K.K., Kim, K.H., Kim, Y., 2013. Study on the effects of hull form on added resistance. In: Proc. of the 12th International Symposium on Practical Design of Ships and Other Floating Structures, Changwon, Korea, pp. 329–337.

Lin, D., Foster, D.P., Ungar, L.H., 2011. VIF regression: A fast regression algorithm for large data. J. Amer. Statist. Assoc. 106 (493), 232–247. http://dx.doi.org/10.1198/jasa.2011.tm10113.

Liu, S., Papanikolaou, A., 2017. On the prediction of the added resistance of large ships in representative seaways. Ships Offshore Struct. 12 (5), 690–696. http://dx.doi.org/10.1080/17445302.2016.1200452.

Liu, S., Papanikolaou, A., 2020. Regression analysis of experimental data for added resistance in waves of arbitrary heading and development of a semi-empirical formula. Ocean Eng. 206, http://dx.doi.org/10.1016/j.oceaneng.2020.107357.

Martinsen, M.A., 2016. A design tool for estimating wave added resistance of container ships (Master's Thesis). Technical University of Denmark.

Maruo, H., 1957. The excess resistance of a ship in rough seas. Int. Shipbuild. Prog. 4 (35), 337–345. http://dx.doi.org/10.3233/ISP-1957-43501.

Mitchell, T.M., 1997. Machine Learning. McGraw-Hill, New York.

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

Mockus, J., 1989. Bayesian Approach to Global Optimization. Kluwer Academic Publishers, Dordrecht.

Nielsen, C.S., 2015. A ship design tool for estimating added resistance in waves (Master's Thesis). Technical University of Denmark.

Oberhagemann, J., Kaufmann, J., Ervik, A., Gramstad, O., Helmers, J., Sireta, F.X., 2019. Wave Load and Response Predictions Combining HOSM, CFD and Machine Learning. In: Proc. of the 38th Int. Conf. on Ocean, Offshore and Arctic Engineering OMAE 2019, Glasgow, Scotland, http://dx.doi.org/10.1115/OMAE2019-95352.

Park, D.-M., Lee, J.-H., Jung, Y.-W., Lee, J., Kim, Y., Gerhardt, F., 2019. Experimental and numerical studies on added resistance of ship in oblique sea conditions. Ocean Eng. 186, 106070. http://dx.doi.org/10.1016/j.oceaneng.2019.05.052.

Park, D.-M., Lee, J., Kim, Y., 2015. Uncertainty analysis for added resistance experiment of KVLCC2 ship. Ocean Eng. 95, 143–156. http://dx.doi.org/10.1016/j.oceaneng.2014.12.007.

Pedregosa, F., et al., 2011. Scikit-learn: Machine learning in python. Mach. Learn. Res. 12, 2825–2830.

Petersen, T., 2000. Wave Load Prediction-a Design Tool. Department of Naval Architecture and Ocean Engineering, Technical University of Denmark.

Sadat-Hosseini, H., Wu, P., Carrica, P., Kim, H., Toda, Y., Stern, F., 2013. CFD verification and validation of added resistance and motions of KVLCC2 with fixed and free surge in short and long head waves. Ocean Eng. 59, 240–273. http://dx.doi.org/10.1016/j.oceaneng.2012.12.016.

Salvesen, N., 1978. Added resistance of ships in waves. J. Hydronaut. 12 (1), 21–34. http://dx.doi.org/10.2514/3.63110.

Salvesen, N., Tuck, E., Faltinsen, O., 1970. Ship motions and sea loads. In: Annual Meeting of the Society of Naval Architects and Marine Engineers, SNAME, New York.

Sclavounos, P.D., Ma, Y., 2018. Artificial Intelligence Machine Learning in Marine Hydrodynamics. In: Proc. 37th Int. Conf. on Ocean, Offshore & Arctic Engineering OMAE 2018, Madrid, Spain.

Seabold, S., Perktold, J., 2010. statsmodels: Econometric and statistical modeling with python. In: Proc. of 9th Python in Science Conference, http://dx.doi.org/10.25080/Majora-92bf1922-011.

Shigunov, V., el Moctar, O., Papanikolaou, A., Potthoff, R., Liu, S., 2018. International benchmark study on numerical simulation methods for prediction of maneuverability of ships in waves. Ocean Eng. 165, 365–385. http://dx.doi.org/10.1016/j.oceaneng.2018.07.031.

Sigmund, S., el Moctar, O., 2018. Numerical and experimental investigation of added resistance of different ship types in short and long waves. Ocean Eng. 147, 51–67. http://dx.doi.org/10.1016/j.oceaneng.2017.10.010.

Strøm-Tejsen, J., Yeh, H.Y.H., Moran, D.D., 1973. Added resistance in waves. Soc. Nav. Archit. Mar. Eng. Trans. 81, 109–143.

Taniguchi, T., Ichinose, Y., 2020. Hull form design support tool based on machine learning. In: Proc. of the 19th Conference on Computer and IT Applications in the Maritime Industries, COMPIT'20, Pontignano.

The GPyOpt authors, 2016. GPyOpt: A Bayesian optimization framework in python. http://github.com/SheffieldML/GPyOpt.

Uharek, S.A., 2019. Numerical Prediction of Ship Manoeuvring Performance in Waves (Doctoral Thesis). Technical University of Berlin.

Valanto, P., Hong, Y., 2015. Experimental investigation on ship wave added resistance in regular head, oblique, beam, and following waves. In: Proc. of the 25th International Society of Offshore and Polar Engineers.

Yasukawa, H., Hirata, N., Matsumoto, A., Kuroiwa, R., Mizokami, S., 2019. Evaluations of wave-induced steady forces and turning motion of a full hull ship in waves. J. Mar. Sci. Technol 24 (1), 1–15. http://dx.doi.org/10.1007/s00773-018-0537-3.

Yeo, I.-K., Johnson, R.A., 2000. A new family of power transformations to improve normality or symmetry. Biometrika 87 (4), 954–959. http://dx.doi.org/10.1093/biomet/87.4.954.