



Physics-informed probabilistic modelling of ship arrival time and fuel consumption under wave-induced uncertainty

Downloaded from: <https://research.chalmers.se>, 2026-10-13 10:03 UTC

Citation for the original published paper (version of record):

Zhang, C., Fu, S., Zhang, M. et al (2027). Physics-informed probabilistic modelling of ship arrival time and fuel consumption under wave-induced uncertainty. Reliability Engineering and System Safety, 278. <http://dx.doi.org/10.1016/j.ress.2026.113472>

N.B. When citing this work, cite the original published paper.



Physics-informed probabilistic modelling of ship arrival time and fuel consumption under wave-induced uncertainty

Chi Zhang^a, Shanshan Fu^b, Mingyang Zhang^c, Wengang Mao^{a,*}, Tsoulakos Nikolaos^d

^a Department of Mechanical Engineering, Chalmers University of Technology, Gothenburg, Sweden

^b College of Transport & Communications, Shanghai Maritime University, Shanghai, China

^c State Key Laboratory of Submarine Geoscience, Shanghai Jiao Tong University, Shanghai, China

^d Laskaridis Shipping Company, Kifissia, Athens, Greece

ARTICLE INFO

Keywords:

Physics-informed
Gaussian process regression
Energy efficiency
Arrival time prediction
Bayesian inference

ABSTRACT

Reliable prediction of fuel consumption and estimated time of arrival (ETA) is essential for maritime operational decision support. However, added resistance due to waves is often treated as a deterministic correction, limiting the quantification of schedule delay risk and fuel consumption uncertainty under variable sea states. This study develops a physics-informed heteroscedastic Gaussian process regression (Phi-HGPR) framework for probabilistic prediction of added wave resistance and its propagation to ETA and fuel consumption uncertainty. A semi-empirical model is embedded as the Gaussian process mean function, while a non-parametric residual process and an input-dependent noise model capture model discrepancy, experimental scatter, and data sparsity in oblique wave conditions. A hybrid resampling strategy mitigates severe heading imbalance in 3081 towing tank records covering 52 ships and 288 experimental cases. The model is evaluated on five representative hulls and compared with a semi-empirical baseline and standard Gaussian process regression. Across oblique sea, it reduces prediction error relative to standard Gaussian process regression while preserving physically plausible curve shapes. The learned transfer function is embedded in a Bayesian sequential updating scheme using full-scale North Atlantic voyage data. Uncertainty in added wave resistance prediction is propagated through a fixed ship performance model, while observed leg-wise sailing times are used to update the posterior weights of candidate power profiles. Posterior distributions of ETA and fuel consumption are updated as observations are assimilated, enabling schedule reliability evaluation and fuel-exceedance risk assessment.

1. Introduction

As the most efficient mode of transport, shipping carries about 90 % of the world's traded goods. Global shipping has grown by approximately 20% over the past decade and continues to rise, thereby heightening the pressure on the shipping industry to reduce its carbon emissions [1]. In response, the International Maritime Organization (IMO) set a target of net-zero greenhouse-gas emissions by 2050 at the 80th session of its Marine Environment Protection Committee (MEPC)

[2]. Ship energy performance models, which are used to predict ship speed, power demand, fuel consumption and emissions under given operating and environmental conditions, are fundamental to improving energy efficiency and meeting regulatory targets [3–6]. A ship is subject not only to calm water resistance but also to external impacts such as waves and wind [7]. Wave-induced resistance is recognized as the dominant environmental load in the open sea. Under severe sea states, this additional resistance can cause speed losses and directly degrade propulsive performance, influencing the selection and operating speed

Abbreviations: ANN, Artificial Neural Network; CFD, Computational Fluid Dynamics; CII, Carbon Intensity Indicator; CTH, Semi-empirical added-resistance model adopted from Lang and Mao (2021); DWT, Deadweight tonnage; EEXI, Energy Efficiency Existing Ship Index; ETA, Estimated Time of Arrival; GPR, Gaussian Process Regression; HGPR, Heteroscedastic Gaussian Process Regression; IMO, International Maritime Organization; ISO, International Organization for Standardization; ITTC, International Towing Tank Conference; MAE, Mean Absolute Error; MCR, Maximum Continuous Rating; PCA, Principal Component Analysis; Phi-HGPR, Physics-Informed Heteroscedastic Gaussian Process Regression; PI, Predictive Interval; PICP, Prediction Interval Coverage Probability; PINN, Physics-Informed Neural Network; RBF, Radial Basis Function; RMSE, Root Mean Square Error; RQ, Rational Quadratic; RRMSE, Relative Root Mean Square Error; SFOC, Specific Fuel Oil Consumption; SMOTE, Synthetic Minority Over-sampling Technique.

* Corresponding author.

E-mail address: wengang.mao@chalmers.se (W. Mao).

<https://doi.org/10.1016/j.ress.2026.113472>

Received 13 May 2026; Received in revised form 31 August 2026; Accepted 6 September 2026

Available online 7 September 2026

0951-8320/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

of the main engine, fuel consumption, and consequently emission levels [8]. The problem has regained urgency under the IMO short-term greenhouse gas measures, which entered into force on 1 January 2023: the Energy Efficiency Existing Ship Index (EEXI) and the Carbon Intensity Indicator (CII) explicitly link technical and operational efficiency to regulatory compliance and ratings [2]. Accurate prediction of speed loss and power increase due to added wave resistance is therefore integral to design, retrofitting, voyage planning, and regulatory reporting [9].

Empirical and semi-empirical methods are developed to meet the demand for rapid estimation of added resistance in waves, providing standardized predictions within calibrated applicability ranges and therefore serving as widely used tools for sea-trial correction and operational assessments. Liu and Papanikolaou [10] improved practical head sea formulas by adding correction terms tied to ship particulars and by calibrating against multi-type datasets to enhance robustness across hull families. Lang and Mao [11] proposed a semi-empirical model for predicting added wave resistance and speed loss in head sea by introducing a significant-wave-height correction factor to capture high wave impacts associated with nonlinear ship motions. Kim et al. [12] proposed a new semi-empirical model that combines the predictions of these two models using hyperbolic tangent weighting functions based on the wavelength-to-ship-length ratio and wave heading.

In addition, recent studies have applied machine learning methods to the prediction of added resistance in waves. Cepowski [13,14] trained artificial neural networks using only basic design parameters, first providing a closed-form ANN surrogate for preliminary design and later an ensemble of five ANNs that improves reliability over a single network when only model test data are available. Duan et al. [15] proposed a deep feedforward neural network for fast prediction of added resistance in heading waves. Zhang et al. [16] applied GPR (Gaussian process regression) for added wave resistance prediction in head sea.

However, both types of methods have their own limitations. For example, empirical and semi-empirical formulas are often insufficient for capturing complex nonlinear phenomena, and their updating and further refinement are not convenient. In contrast, machine learning models usually lack interpretability and are accurate only for interpolation within the range of the training data; once the prediction point moves far beyond the observed data domain, reliable extrapolation becomes difficult [17]. Since most towing-tank experiments for added wave resistance estimation are conducted under head sea conditions, the extrapolation capability of the model is particularly important.

Physics-informed machine learning can be implemented in different ways. One approach is to combine physics-based and data-driven predictions through weighted fusion [18]. For example, Fan et al. [7] combined the outputs of a physical ship model with neural network and random forest models for ship fuel-consumption prediction. A second approach is to use a data-driven model as a surrogate for a key component of a physics-based framework. For example, Kanazawa et al. [19] used neural networks to compensate for the single-step-ahead velocity prediction error of a physics-based ship dynamic model during docking operations. A third approach is to embed physical knowledge as prior information in a data-driven model. For instance, Deo et al. [20] incorporated a physics-informed mean function into a sparse variational Gaussian process for underwater radiated noise prediction.

Within the data-driven component of such physics-informed machine learning approaches, Gaussian Process Regression is one of the more suitable choices, because it has a clear mathematical interpretation and can provide credible intervals for its predictions [21]. Ko and Nam [22] proposed a physics-informed GPR model for ship maneuverability prediction, in which speed-dependent hydrodynamic coefficients are embedded in the GPR mean function. Li et al. [23] proposed a physics-constrained GPR model for wave-energy-converter arrays by embedding constraints into the kernel design to predict hydrodynamic interactions. Jiang et al. [21] proposed a physics-informed Gaussian process model, where physical knowledge, historical degradation data,

and limited current observations are combined to construct nonstationary priors for probabilistic degradation prognostics and fatigue-damage prediction.

However, modelling added wave resistance at arbitrary wave headings using a physics-informed GPR still has several challenges. First, the collected towing-tank experimental data for added wave resistance are imbalanced, as most experiments are conducted under head sea conditions. Second, the standard GPR formulation is derived under the assumption that the output noise variance is constant and independent of the input. However, the uncertainty and noise associated with added wave resistance are strongly input-dependent and, from a physical perspective, are governed by factors such as wave heading, wavelength, Froude number, and hull form. Therefore, standard GPR is insufficient for characterizing the relevant uncertainty in added wave resistance under arbitrary wave headings.

Motivated by these limitations, this study develops a Physics-Informed Heteroscedastic Gaussian Process Regression (Phi-HGPR) model for predicting added wave resistance at arbitrary wave headings. The overall workflow is shown in Fig. 1. The model is trained using towing-tank experimental data collected from the open literature, covering multiple hull forms, wave headings, and Froude numbers. To improve extrapolation in sparsely sampled oblique wave conditions, a semi-empirical model is embedded as the prior. In addition, a heteroscedastic noise model is introduced to represent input-dependent uncertainty in added wave resistance. To mitigate the imbalance of the experimental database, Synthetic Minority Over-sampling Technique (SMOTE) is applied to oversample sparse regions, while random under-sampling is used in dense regions. Finally, we integrate the proposed model with full-scale onboard performance-monitoring data from a chemical tanker and, within a Bayesian-inference framework, perform real-time probabilistic estimation of the vessel's estimated time of arrival (ETA) and fuel consumption. To the best of our knowledge, the use of physics-informed machine learning for modelling added wave resistance at arbitrary wave headings, the application of hybrid resampling to imbalanced towing-tank experimental data, and the analysis of how uncertain added wave resistance affects ETA and fuel-consumption estimation have not yet been discussed in the existing literature.

The remainder of the paper is organized as follows. Section 2 introduces the theoretical foundations of Phi-HGPR. Section 3 describes the training dataset and the SMOTE procedure. Section 4 presents added wave resistance predictions for various hull types using the proposed model. Section 5 demonstrates the application to real-time estimation of a vessel's ETA and fuel consumption.

2. Methodology of the proposed Phi-HGPR model

In this study, the Physics-Informed Heteroscedastic Gaussian Process Regression (Phi-HGPR) model is developed to predict added wave resistance across a range of ship types and Froude numbers under arbitrary wave headings. Section 2.1 outlines the fundamental formulation of standard Gaussian Process Regression (GPR). Section 2.2 introduces the incorporation of physics-based knowledge into GPR, focusing on the modelling of the mean function. Section 2.3 addresses the modelling of the variance function within the GPR framework, thereby forming the heteroscedastic GPR (HGPR) approach.

Changes in a ship's added resistance in waves are governed by several factors, including hull form, operating condition, and wave properties. For the nondimensional transfer function of mean resistance in regular waves, the main influencing variables are wave frequency, wave heading, and ship speed. Therefore, the performance of different approaches for predicting C_{aw} is typically evaluated by classifying the predictions by Froude number, encounter angle and wavelength ratio for various ship types. The corresponding input and output variables are summarized in Table 1. L_E denotes the entrance length. It is defined as the horizontal distance between the forepeak and the point along the waterline where the local breadth reaches 99% of the ship's maximum

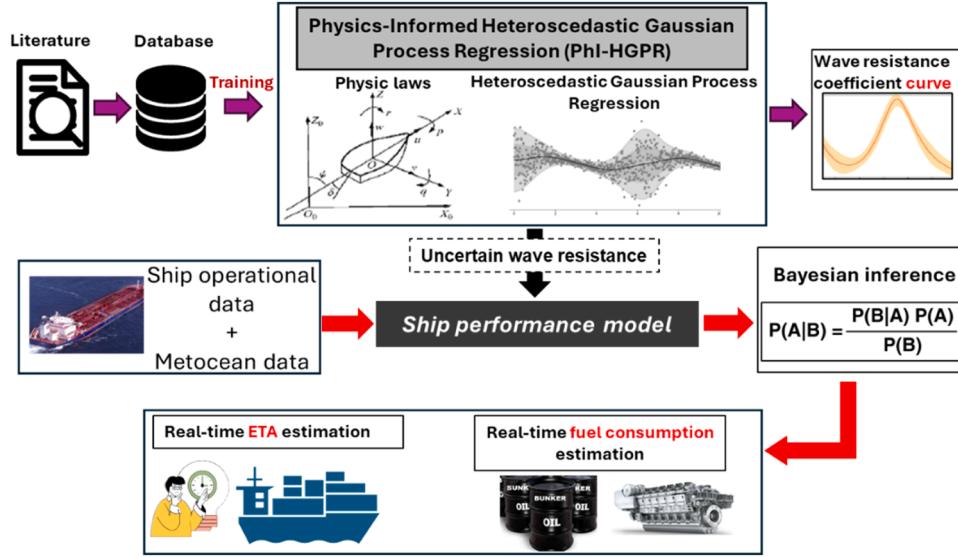


Fig. 1. Workflow of this paper.

Table 1
Input and output variables for C_{aw} prediction.

	Description	Attributes
Input	Wavelength-to-ship length ratio	λ/L
	Encounter angle	β
	Froude number	F_n
	Block coefficient	C_B
	Ship length-to-beam ratio	L/B
	Beam-to-draft ratio	B/T
	Non-dimensional radius of gyration in yaw	K_{yy}/L
	Effective ship length ratio	L_E/L
Output	Added wave resistance coefficient	C_{aw}

breadth. This quantity is regarded as an important variable for wave-resistance prediction in the semi-empirical formulations proposed by Lang and Mao [11]. In this study, the encounter angle β is defined as the relative angle between the incident wave direction and the ship heading, where $\beta = 0^\circ$ corresponds to head sea, $\beta = 90^\circ$ to beam sea, and $\beta = 180^\circ$ to following sea.

2.1. Standard Gaussian process regression

Gaussian Process Regression (GPR) is a Bayesian non-parametric regression framework that represents an unknown input–output relationship as a stochastic process [24–26]. For each training sample with input vector $\mathbf{x}_i$, the observation model is written as:

$$C_{aw}^i = f(\mathbf{x}_i) + \varepsilon_i \quad (1)$$

where $\varepsilon_i \sim \mathcal{N}(0, \sigma_n^2)$, and σ_n^2 denotes the noise variance. The latent function $f(\mathbf{x})$ is assigned a Gaussian process prior $f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$, where $m(\mathbf{x})$ is the mean function, usually set to zero for convenience, and $k(\mathbf{x}, \mathbf{x}') = \text{cov}(f(\mathbf{x}), f(\mathbf{x}'))$ is the covariance (kernel) function. The input vectors are collected in the set:

$$\mathbf{X} = \{\mathbf{x}_i, i = 1, \dots, n\} \quad (2)$$

The corresponding latent function values at the training inputs are defined as:

$$\mathbf{f} = [f_1, f_2, f_3, \dots, f_n]^T = [f(\mathbf{x}_1), f(\mathbf{x}_2), f(\mathbf{x}_3), \dots, f(\mathbf{x}_n)]^T \quad (3)$$

where $\mathbf{f} \sim \mathcal{N}(m(\mathbf{X}), K(\mathbf{X}, \mathbf{X}))$. All the values of C_{aw} in the data set can be defined as:

$$C_{aw} = \{C_{aw}^i, i = 1, \dots, n\} \quad (4)$$

which is jointly Gaussian distributed,

$$C_{aw} \sim \mathcal{N}(m(\mathbf{X}), K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}) \quad (5)$$

where $K(\mathbf{X}, \mathbf{X})$ is the kernel matrix of training inputs and $\mathbf{I}$ is an identity matrix. For a new query input $\mathbf{x}_*$, we assume that the associated latent value f_* and the training outputs C_{aw} share a joint Gaussian distribution:

$$\begin{bmatrix} C_{aw} \\ f_* \end{bmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m(\mathbf{X}) \\ \mu(\mathbf{x}_*) \end{pmatrix}, \begin{pmatrix} K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I} & K(\mathbf{X}, \mathbf{x}_*) \\ K(\mathbf{x}_*, \mathbf{X}) & k(\mathbf{x}_*, \mathbf{x}_*) \end{pmatrix}\right) \quad (6)$$

Within the Bayesian formulation, the central result of GPR is the posterior predictive distribution of f_* conditional on the observed data, which remains Gaussian:

$$f_* | \mathbf{X}, C_{aw}, \mathbf{x}_* \sim \mathcal{N}(\bar{f}_*, \Sigma_{f_*}) \quad (7)$$

Its mean and variance can be calculated as:

$$\bar{f}_* = K(\mathbf{x}_*, \mathbf{X}) [K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} C_{aw} \quad (8)$$

$$\sigma(f_*) = k(\mathbf{x}_*, \mathbf{x}_*) - K(\mathbf{x}_*, \mathbf{X}) [K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} K(\mathbf{X}, \mathbf{x}_*) \quad (9)$$

where $\bar{f}_*$ is the predictive mean, representing the estimated value of the added resistance coefficient for the input $\mathbf{x}_*$, while the predictive variance $\sigma(f_*)$ quantifies the corresponding uncertainty. These expressions enable us to compute, for any specified input (a specific ship, speed, wave heading, and wavelength/frequency), both the mean and uncertainty (variance) of the added wave resistance coefficient.

2.2. Proposed method for including physical information

Although GPR is a powerful tool for regression and has shown success in predicting added wave resistance in head sea, its performance is highly dependent on the distribution and coverage of the training data. Therefore, when the data are sparse or extrapolation beyond the observed domain is required, such as in wave headings other than head sea, the model's generalization capability on unseen data becomes constrained. In practical engineering applications, employing a purely black-box model to predict added wave resistance requires sufficient data. Achieving comprehensive coverage of data needs extensive towing tank experiments, which are both costly and time-consuming; for

extreme sea states or rare operating conditions, obtaining sufficient data may be impossible. However, by incorporating physics-based models, one can enhance extrapolation capability, thereby reducing the need for exhaustive training data coverage [27]. Physics-informed learning provides a general framework for incorporating physical knowledge into data-driven models, either through neural networks such as Physics-Informed Neural Networks (PINNs) or through kernel-based regression models [28]. Since GPR has shown promising performance in head sea added wave resistance prediction, this study adopts the kernel-based route, where the prior physical knowledge is embedded as the Gaussian process prior for reliable estimation of added wave resistance for arbitrary wave headings.

Semi-empirical formulas are widely used physics-based tools for predicting added resistance in waves. Researchers have proposed several well-known versions, including the classic formula by Fujii and Takahashi [29], the STAwave-1 and STAwave-2 models, and the ITTC guidelines based on large model-test data sets [30,31]. Liu and Papanikolaou [10] first proposed a statistical method that combines elements of the formulations by Faltinsen [32] and Jinkine and Ferdinande [33]. In parallel, Lang and Mao [11] developed a head sea added wave resistance model based on the approaches of Tsujimoto et al. [34] and Jinkine and Ferdinande [33]. Their formulation follows the general structure of Liu and Papanikolaou's [10] equations and was further tuned using additional experimental datasets. This model was later generalized to arbitrary wave headings by introducing a correction factor depending on the encounter frequency. For the full expressions and practical guidance on their use, the reader is referred to the original publications.

The mean function $m(\cdot)$ in a standard GPR model is usually assumed to be zero. Replacing the zero mean with a prescribed physical function guides the posterior prediction toward the prior mean, especially in regions with limited data support. In this paper, we use the semi-empirical formulas developed by Lang and Mao [35] (we will hereafter refer to this semi-empirical model as the CTH model for short) as the mean, i.e., we set $m(\mathbf{x}) = f_{CTH}(\mathbf{x})$. The added wave resistance coefficient can be modelled as:

$$C_{aw}(\mathbf{x}) \sim f_{CTH}(\mathbf{x}) + \mathcal{GP}(0, k(\mathbf{x}, \mathbf{x}')) \quad (10)$$

Eq. (10) indicates that the GP term serves to adjust the residuals. By substituting the semi-empirical formula $f_{CTH}(\cdot)$ into Eq. (8), we obtain the predicted C_{aw} for the current input vector $\mathbf{x}_*$:

$$\bar{f}_*(\mathbf{x}_*) = f_{CTH}(\mathbf{x}_*) + K(\mathbf{x}_*, \mathbf{X}) [K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} (C_{aw} - f_{CTH}(\mathbf{X})) \quad (11)$$

which shows that when the semi-empirical model provides an accurate physical description of added wave resistance, the influence of the error term $(C_{aw} - f_{CTH}(\mathbf{X}))$ is marginal and the prediction is driven mainly by $f_{CTH}(\cdot)$. Furthermore, if the query test point $\mathbf{x}_*$ is far from the training set $\mathbf{X}$, the corresponding kernel value approaches zero (an intrinsic property of common kernels such as the squared exponential). Under such circumstances, the predicted $C_{aw}(\mathbf{x}_*)$ is governed essentially by the $f_{CTH}(\mathbf{x}_*)$ term. This is an important property, because a finite (and possibly small) training dataset cannot span the entire state space. When the input conditions to be predicted lie outside the sampled domain (i.e., where the trained GPR model does not generalize well), the prediction will revert to the semi-empirical model.

However, the fixed CTH mean function may also introduce systematic prior bias. Since its parameters are not recalibrated using the present dataset, errors in the baseline prediction can only be corrected by the residual GP. This correction is effective where sufficient local training data are available, but becomes weaker in data-sparse regions, where the final prediction may inherit part of the bias of the CTH prior. Therefore, the fixed physical prior provides useful regularization but may still reduce model flexibility when the baseline prediction is inaccurate.

2.3. Heteroscedastic Gaussian process regression

In Gaussian Process Regression, the predicted mean value for a test point $\mathbf{x}^*$ can be represented as a weighted combination of the training outputs; the more similar a training sample is to $\mathbf{x}^*$, the greater its contribution to this weighted average. As shown in Eq. (9), the variance equals the prior uncertainty minus the portion that has already been “explained away” by the observed data, thus quantifying the residual uncertainty of the model at the given input. The posterior variance reflects epistemic uncertainty due to limited data, while the input-dependent noise models aleatoric uncertainty [36]. When the training data become denser and more strongly correlated, the additional explanatory power further compresses the model variance. The standard treatment for noise in GPR assumes homoscedasticity, i.e., a zero-mean Gaussian distribution with constant variance, which often fails in real-world applications. Because the noise is not modelled as input-dependent, the model tends to be over-confident in high-noise regions and to underestimate uncertainty in low-noise regions. To address this problem, a heteroscedastic Gaussian Process Regression (HGPR) is introduced, in which a second GP is employed to model the noise. This GP is placed on the logarithm of the input-dependent noise variance, following the standard heteroscedastic GP formulation [37,38], which ensures that the noise variance remains non-negative after exponentiation:

$$g(\mathbf{X}) = \log \sigma^2(\mathbf{X}) \quad (12)$$

$$g(\mathbf{X}) \sim \mathcal{GP}(m_g(\mathbf{X}), k_g(\mathbf{X}, \mathbf{X}')) \quad (13)$$

where $m_g(\mathbf{X})$ is the mean function of the variance and $k_g(\mathbf{X}, \mathbf{X}')$ is its covariance function. The kernels of both GPs use a composite kernel comprising K_{RBF} , K_{RQ} , and K_{Matern} ($\nu=3/2$), where:

$$k_{RBF}(\mathbf{x}, \mathbf{x}') = \exp \left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\ell_{RBF}^2} \right) \quad (14)$$

$$k_{RQ}(\mathbf{x}, \mathbf{x}') = \left(1 + \frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\alpha\ell_{RQ}^2} \right)^{-\alpha} \quad (15)$$

$$k_{Matern}(\mathbf{x}, \mathbf{x}') = \left(1 + \frac{\sqrt{3}\|\mathbf{x} - \mathbf{x}'\|}{\ell_{Matern}} \right) \exp \left(-\frac{\sqrt{3}\|\mathbf{x} - \mathbf{x}'\|}{\ell_{Matern}} \right) \quad (16)$$

$$k(\mathbf{x}, \mathbf{x}') = C[k_{RBF}(\mathbf{x}, \mathbf{x}') + k_{RQ}(\mathbf{x}, \mathbf{x}')] + k_{Matern}(\mathbf{x}, \mathbf{x}') \quad (17)$$

where ℓ is the length scale and C is a scaling constant. The parameter α in Eq. (15) governs the weighting among different length scales. The total predictive variance for a test point $\mathbf{x}_*$ can be computed as:

$$\sigma_T = \sigma_*^2(\mathbf{x}_*) + \sigma^2(\mathbf{x}_*) \quad (18)$$

where $\sigma_*^2(\mathbf{x}_*)$ denotes the predictive variance yielded by the mean function of the physics-informed GPR, whereas $\sigma^2(\mathbf{x}_*)$ denotes the input-dependent noise variance. Thus, heteroscedastic Gaussian process regression (HGPR) remedies the shortcomings of standard GPR in variance estimation by modelling the input-dependent noise variance, making it better suited to real-world data.

3. Data for modelling and validation

In this study, a database of publicly available towing-tank measurements is developed to train a data-driven model for estimating added wave resistance. The dataset contains 3081 records spanning 52 ships and 288 distinct experimental cases. The variables include both operating and hull-form descriptors, $[\lambda/L, F_r, C_B, B/T, L/B, K_{yy}/L, L_E/L, \beta]$, whereas the response variable is the non-dimensional coefficient of wave added resistance C_{aw} . Fig. 2 shows the distribution of these

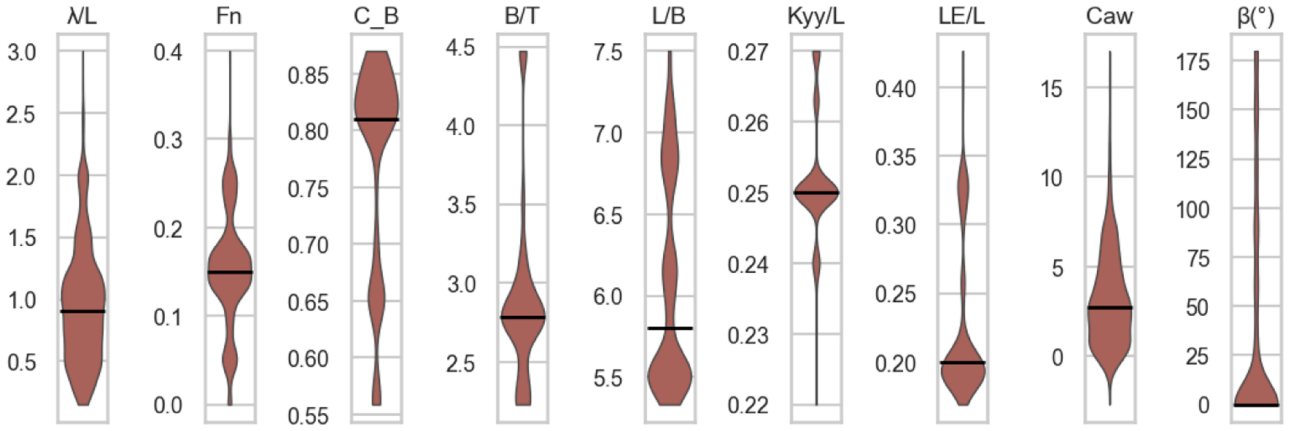


Fig. 2. The distribution of the database according to various ship parameters.

parameters for the ships used in the study, where the dataset spans both short- and long-wave regimes, with the short-wave range more densely sampled. The Froude numbers are concentrated at low values ($\approx 0.10 - 0.15$).

Fig. 3 shows the number of samples under each wave heading angle, with 1836 samples under head sea and 1245 samples under other wave angles. The sample sizes across encounter angles are highly imbalanced (e.g., 0° accounts for the majority). One cause of this data imbalance is that added resistance in waves reaches its maximum near head sea conditions, leading researchers to focus more on head sea experiments. Moreover, experiments in oblique sea are more complex than head sea tests, which involve predominantly longitudinal motions, placing higher demands on experimental apparatus and facilities and leading to greater uncertainty in the results.

Training a model directly on such data causes the optimization to be dominated by the global loss, so the model “prefers” dense regions and achieves lower errors there. Thus, overall performance may look acceptable, but predictions in critical data-poor regimes are biased. To address this issue, we use a hybrid resampling scheme: over-sample sparse operating conditions and under-sample overrepresented ones. This keeps the dataset size in check, balances learning across regimes, and improves the estimation of heteroscedastic uncertainty. This study adopts a hybrid resampling strategy, combining over-sampling and under-sampling, to bring the sample sizes at each encounter angle (0° , 30° , 60° , 90° , 120° , 150° , 180°) close to a predefined target. Specifically,

synthetic minority over-sampling (SMOTE) is applied to the six non-zero angles to expand their sample counts to the set threshold [39]. SMOTE generates new samples in feature space by linear interpolation:

$$s' = s_0 + \zeta(s_j - s_0), \quad \zeta \sim \mathcal{U}(0, 1) \quad (19)$$

where s_0 is a sample from the sparse subset and s_j is the nearest neighbour sample of s_0 . Random under-sampling is applied to the dataset of 0° (head sea) by uniformly selecting, without replacement, a subset of the required size from the original sample pool.

To examine the impact of SMOTE, we applied principal component analysis (PCA) for dimensionality reduction and visualization. Fig. 4 presents the two-dimensional PCA embedding of the features, contrasting the data distributions before and after SMOTE resampling. Prior to resampling, the samples at non-zero encounter angles appear as thin, discontinuous slanted bands with large voids and isolated points. After applying resampling, the representation across angles becomes more balanced; each color-coded vertical band is denser and more continuous, and gaps on the plane are substantially reduced.

4. Results of the constructed Phi-HGPR model

To validate the proposed Phi-HGPR model, a baseline GPR model is also built based on the resampled dataset for comparative analysis. Since Zhang et al. [16] have already demonstrated the application of GPR for predicting head sea wave resistance, the present study focuses on other

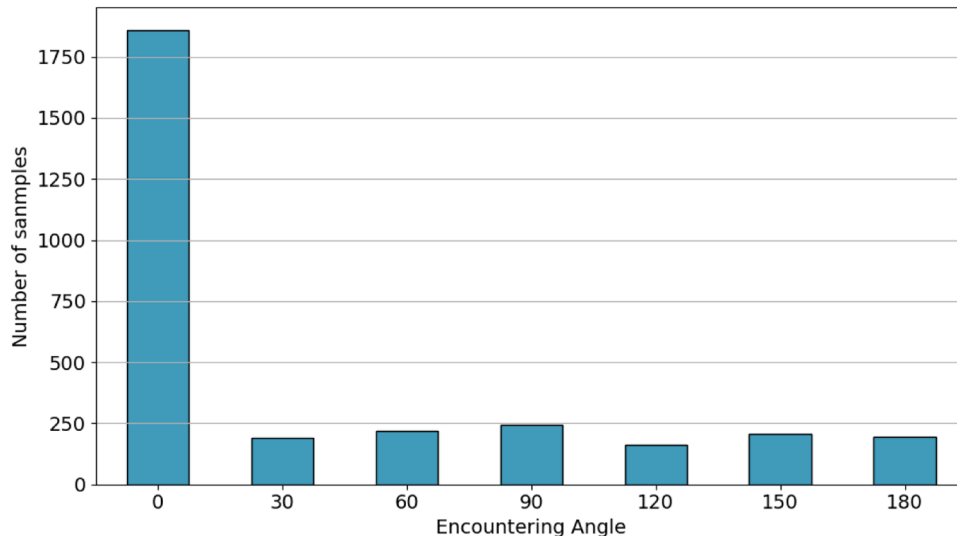


Fig. 3. The distribution of encounter angles in the database.

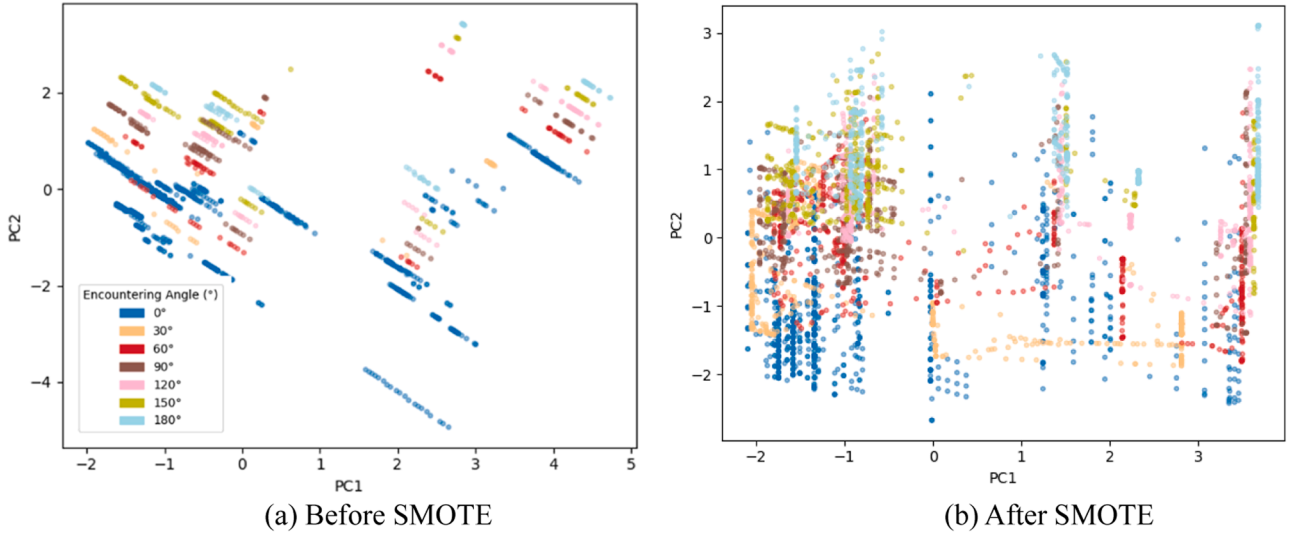


Fig. 4. Comparison of data distribution before and after resampling.

wave headings. Model performance is evaluated using the root mean square error (RMSE), relative root mean square error (RRMSE) and prediction interval coverage probability (PICP),

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (p_i - a_i)^2}{n}} \quad (20)$$

$$\text{RRMSE} = \frac{\text{RMSE}}{\sqrt{\frac{\sum_{i=1}^n a_i^2}{n}}} \times 100\% \quad (21)$$

$$\text{PICP} = \frac{1}{n} \sum_{i=1}^n I_i \times 100\%, \quad I_i = \begin{cases} 1, & L_i \leq a_i \leq U_i, \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

where p_i is the predicted value; a_i denotes the actual value; n is the number of samples; L_i and U_i are the lower and upper bounds of the prediction interval. Added wave resistance experimental data have a special structure, in which observations from different experiments may be treated as independent, while observations within the same experiment are highly correlated, revealing a distinct trend of C_{aw} changing with λ/L . Therefore, complete experimental series under non-head sea conditions are randomly selected from different ship types, such as KVLCC, S175, and HSVA, to form the test set, which accounts for approximately 10% of the total dataset. In addition, all DTC samples are excluded from the training set and reserved as an out-of-sample test case.

Before model training, all variables were scaled using min-max normalization [40]. During training, the mean GP and the variance GP construct their corresponding covariance matrices from the training samples, using the same form of composite kernel (see Eq. (17)). An initial mean GP is first trained and then used to predict the training samples themselves, yielding an error for each training sample. Next, in the standardized input space, the squared errors of neighbouring samples are locally averaged to obtain a local noise-variance estimate for each training point. The logarithm of this local variance is then used as the training target of the second GP. The variance GP is subsequently trained to learn how the noise intensity varies under different input conditions, and the trained variance GP is used to re-estimate the input-dependent noise variance at the training samples. Finally, these sample-wise noise variances obtained from the variance GP are incorporated into the covariance matrix of the mean GP, and the final mean GP is retrained. After training, according to Eq. (18), the Phi-HGPR model estimates a Gaussian distribution characterizing the C_{aw}

uncertainty with a mean and a predictive variance at each λ/L point.

4.1. Case analysis

We choose several typical hull types with their predicted mean values and predictive intervals for detailed case analysis [41].

(1) S175 container ship (4.8% of the total dataset)

As illustrated in Fig. 5, the three models show pronounced differences in their predictions of added wave resistance for the S175 container ship. The proposed Phi-HGPR demonstrates consistently strong performance across most wave headings. By learning the underlying physical trends of the added wave resistance coefficient C_{aw} from a semi-empirical model and introducing a systematic downward adjustment, it produces predictions that are both more accurate and physically plausible. In contrast, the basic GPR model fails to reproduce the overall shape of the response curve, yielding nearly constant predictions over the entire range of headings; nevertheless, at a wave heading of 90° , it happens to achieve a relatively low prediction error. Except for the 150° case, the predictive intervals (PIs) generated by Phi-HGPR successfully encompass most of the experimental data points, indicating reliable uncertainty quantification. By comparison, basic GPR produces excessively wide PIs, which limits their practical usefulness for engineering applications.

(2) HSVA cruise vessel (7.1% of the total dataset)

As shown in Fig. 6, for the HSVA case the basic GPR model again fails to fully reproduce the overall shape of the response curve. Nevertheless, its performance has notably improved compared with that observed for the S175 hull. With increasing encounter angle, the GPR predictions exhibit a clear decreasing trend and achieve the highest accuracy among the compared models at certain headings, most notably at 60° , where the predictions closely match the experimental measurements. In addition, the associated predictive intervals are substantially narrower than those obtained for S175, indicating improved uncertainty estimation. This improvement is likely attributable to the presence of hull forms in the training dataset that are similar to the HSVA model. A closer inspection of the datasets reveals that the KCS hull ($L_{pp} \approx 230$ m, $B \approx 32.2$ m, $T \approx 10.8$ m, $C_B \approx 0.651$) is comparable to the HSVA hull in terms of principal dimensions, and that KCS-related data are well represented in the training set. The Phi-HGPR model continues to effectively correct the semi-empirical CTH baseline and yields a further, albeit modest, reduction in prediction error. Importantly, cases in which the Phi-HGPR predictions lie above the CTH curve demonstrate that the model does not simply apply a uniform downward correction; instead, it adaptively

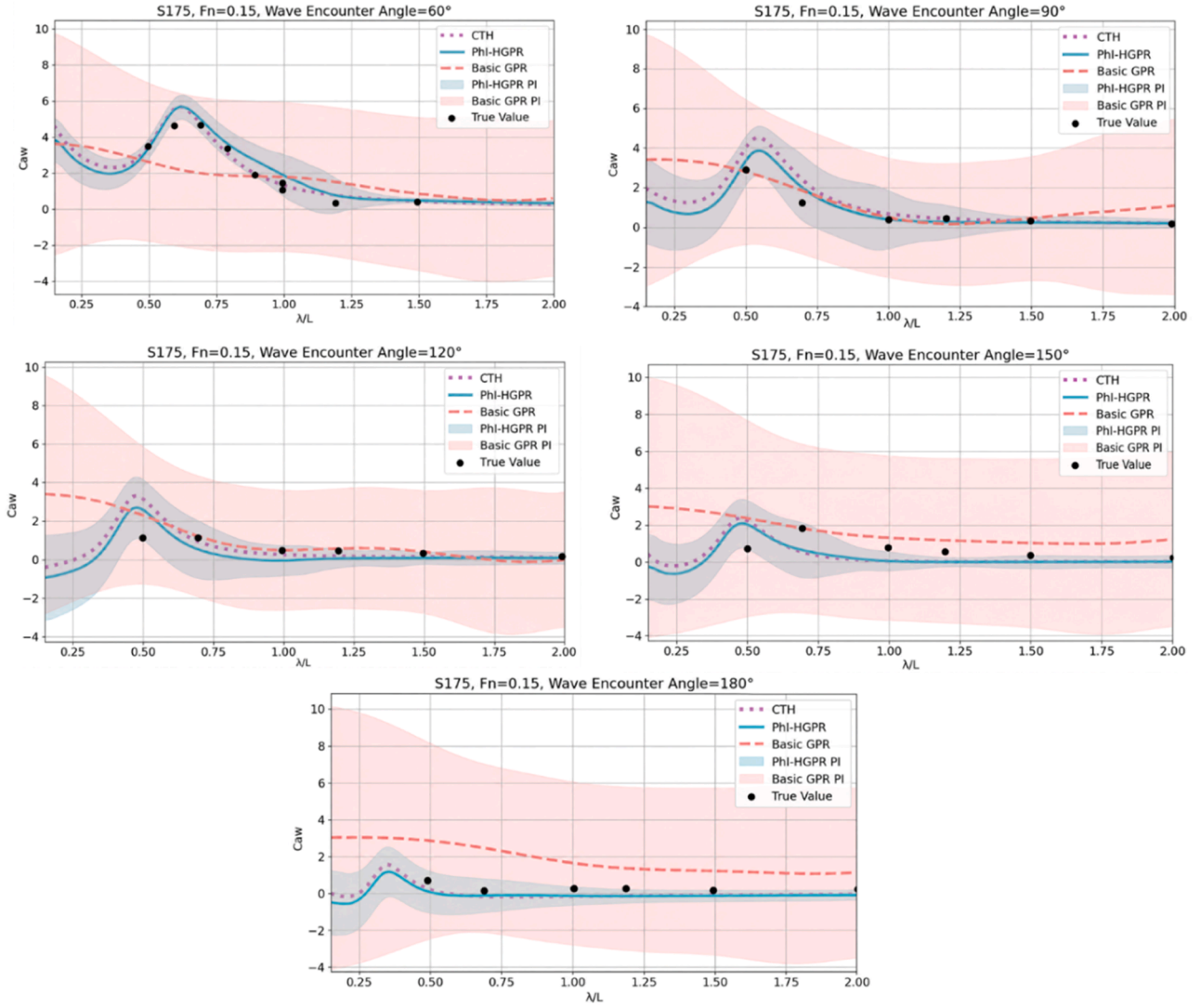


Fig. 5. Results of added wave resistance prediction for the S175 container ship.

modifies the baseline response based on data-driven learning. Moreover, the predictive intervals produced by Phi-HGPR remain consistently narrower and more physically reasonable than those of the basic GPR, reinforcing its advantage in terms of both predictive accuracy and reliable uncertainty quantification.

(3) DTC container ship (1.2% of the total dataset)

All DTC samples are excluded from the training dataset. As shown in Fig. 7, for DTC, the basic GPR still fails to capture the expected behavior; as with S175, the predicted curve does not exhibit a decrease with increasing angle, which is inconsistent with the expected physical trend. Phi-HGPR outperforms the other two models while preserving the physical shape of the resistance curve, and its PIs are more physically plausible than those of the basic GPR. Phi-HGPR performs well mainly because the semi-empirical CTH formula already gives good results for this hull form. The predictive intervals produced by Phi-HGPR generally cover the measured data. However, for wave angles greater than 120° and $\lambda/L > 0.5$, the predicted added resistance coefficient shows a brief increase that is inconsistent with the trend given by the CTH model. This behavior indicates that the model is sensitive to the available data. In contrast, the very wide predictive intervals of the plain GPR model are probably caused by the scarcity of training data from hull types with principal dimensions similar to those of DTC.

(4) KVLCC2 tanker (16.6% of the total dataset)

For this case study hull, the test Froude number is relatively low. As

shown in Fig. 8, at encounter angles of 60° and 90° , the basic GPR curve shows only limited physically meaningful behavior. In contrast, Phi-HGPR introduces physical information by combining a semi-empirical baseline with a data-driven correction, and it can preserve the correct physical trend in most cases. At an encounter angle of 90° , the basic GPR achieves the lowest RMSE, which indicates that the GPR model can fit the available data points very well. In addition, the semi-empirical formula alone can capture the main physical trend reasonably well, but it is less flexible than Phi-HGPR. When the encounter angle increases to 150° (close to following sea), the basic GPR prediction changes little over the entire wavelength range. By comparison, Phi-HGPR can be interpreted as correcting the semi-empirical baseline by shifting the curve downward, which allows it to match the trend observed in the measurements more effectively and achieve the lowest RMSE for this condition. Moreover, the Phi-HGPR predictive interval is narrower and more physically plausible, whereas the basic GPR gives unrealistically wide predictive intervals in both the short-wavelength and long-wavelength regions.

(5) 170k bulk carrier (7.8% of the total dataset)

Fig. 9 illustrates that the basic GPR model predicts the peak values of the added resistance coefficient with forward shifts, the magnitude of which varies across different wave headings. These phase shifts are most likely induced by confounding influences from data points associated with other, geometrically similar hull forms in the training set. Except

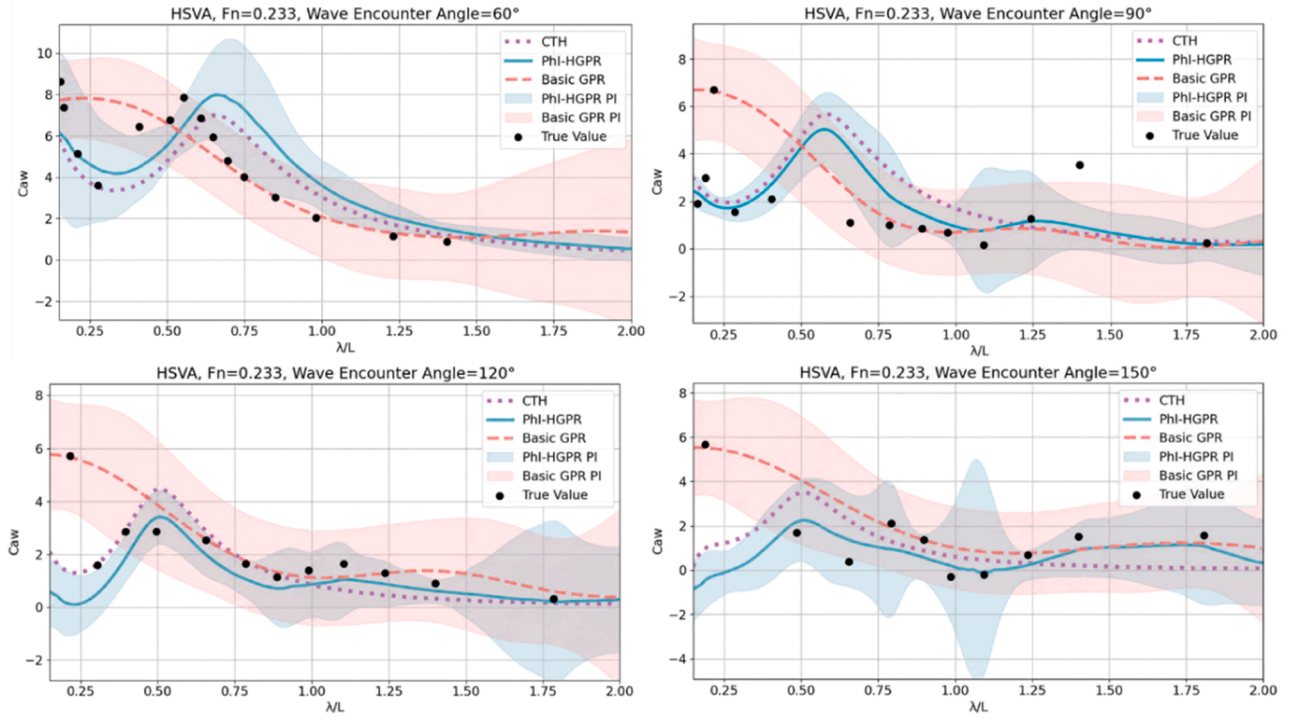


Fig. 6. Results of added wave resistance prediction for the HSVA cruise ship.

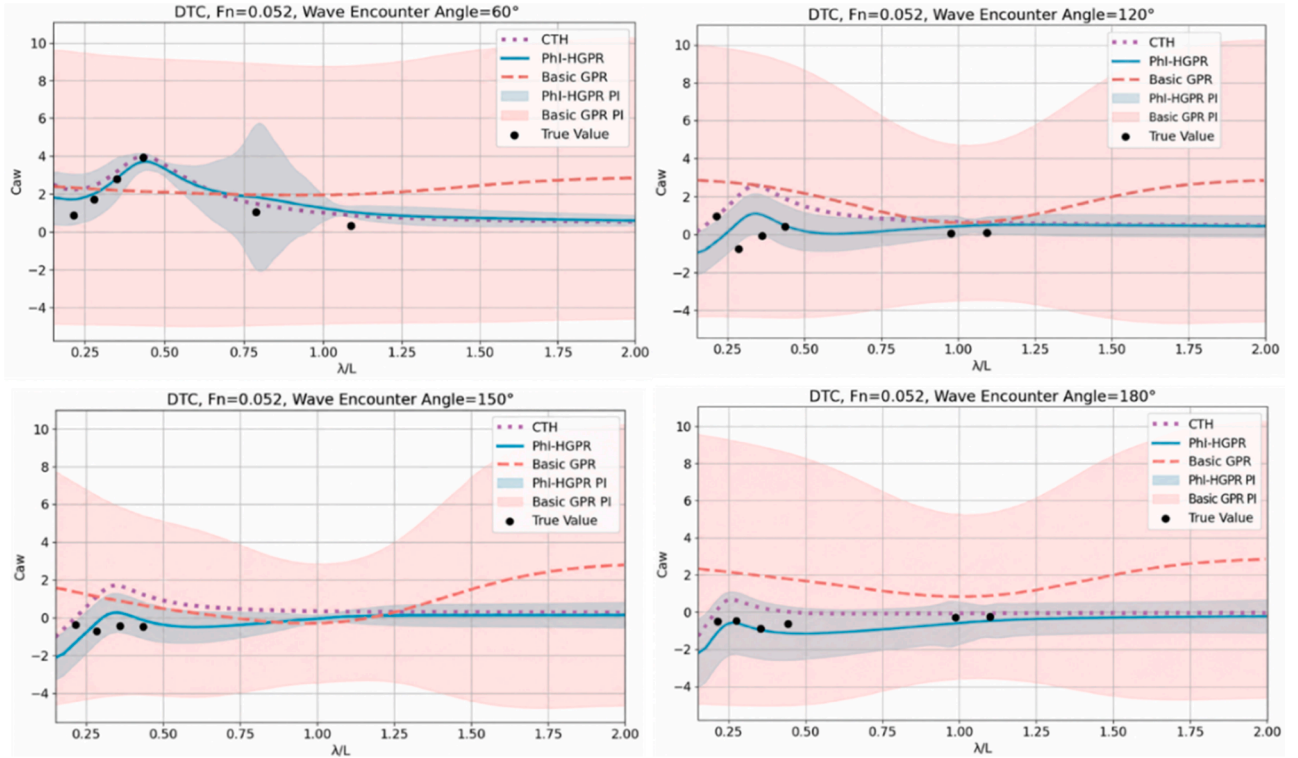


Fig. 7. Results of added wave resistance prediction for the DTC container ship.

for the 60° case, the Phi-HGPR model delivers the most accurate predictions while consistently preserving the expected physical behavior of the response. Although the experimental data for this hull form were not used to calibrate the parameters of the CTH model, the CTH baseline performs well in terms of prediction accuracy and maintains a physically reasonable curve shape. By inheriting these physical characteristics and

augmenting them with data-driven corrections, Phi-HGPR achieves superior overall performance compared with the standalone CTH model. In this case, both GPR-based approaches produce predictive intervals of reasonable width, which indicates adequate uncertainty quantification.

Table 2 and Fig. 10 present the RMSE, RRMSE and PICP of C_{aw} predictions for three methods (CTH, Phi-HGPR, and GPR) over 21

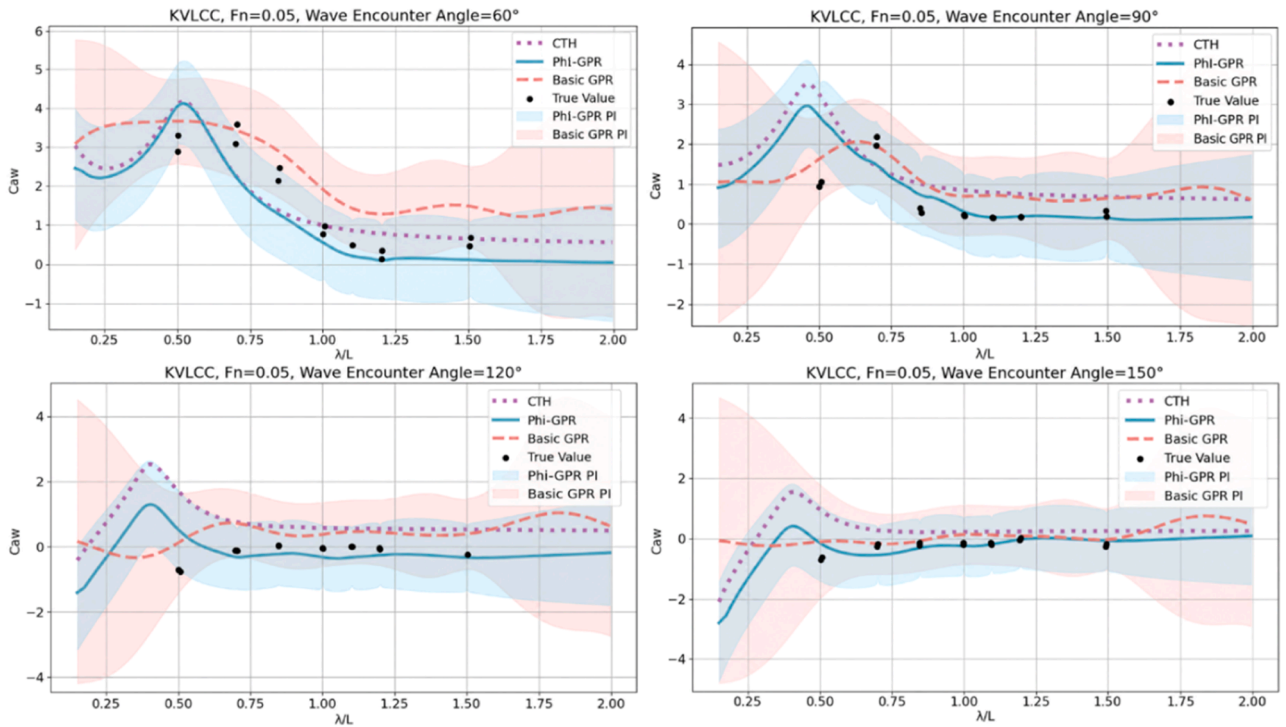


Fig. 8. Results of added wave resistance prediction for the KVLCC2 tanker.

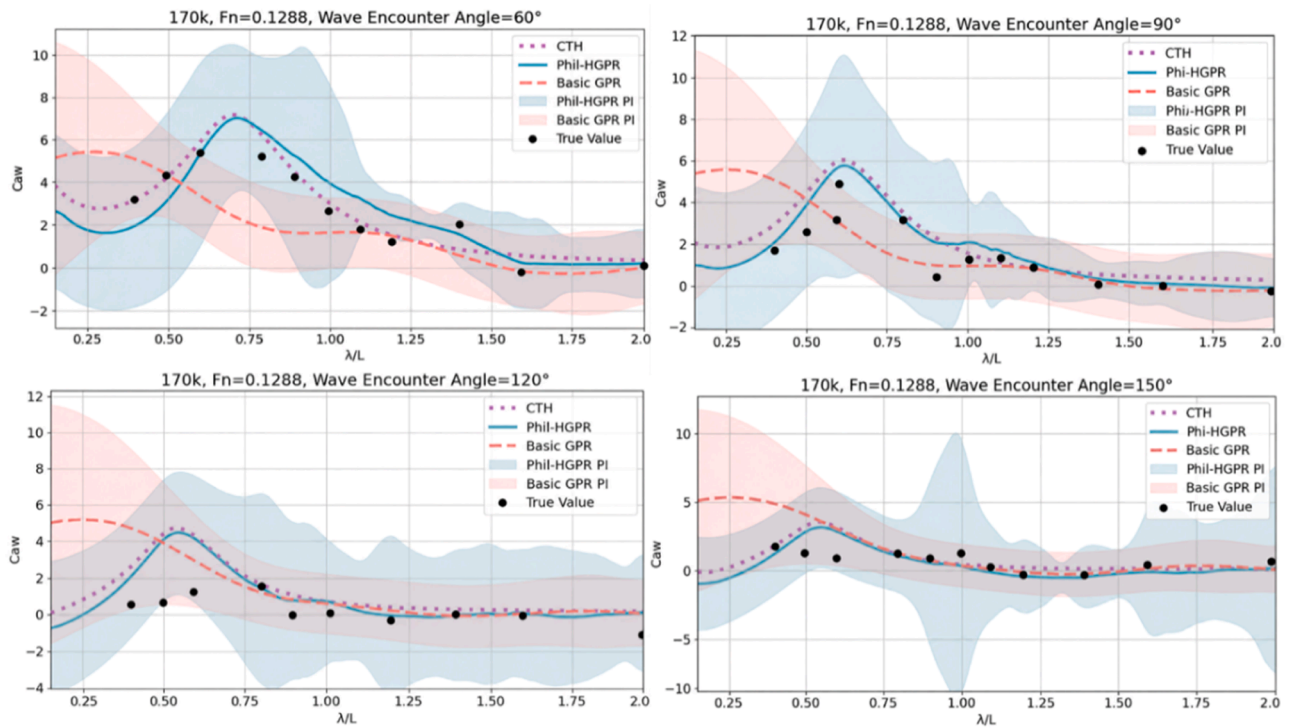


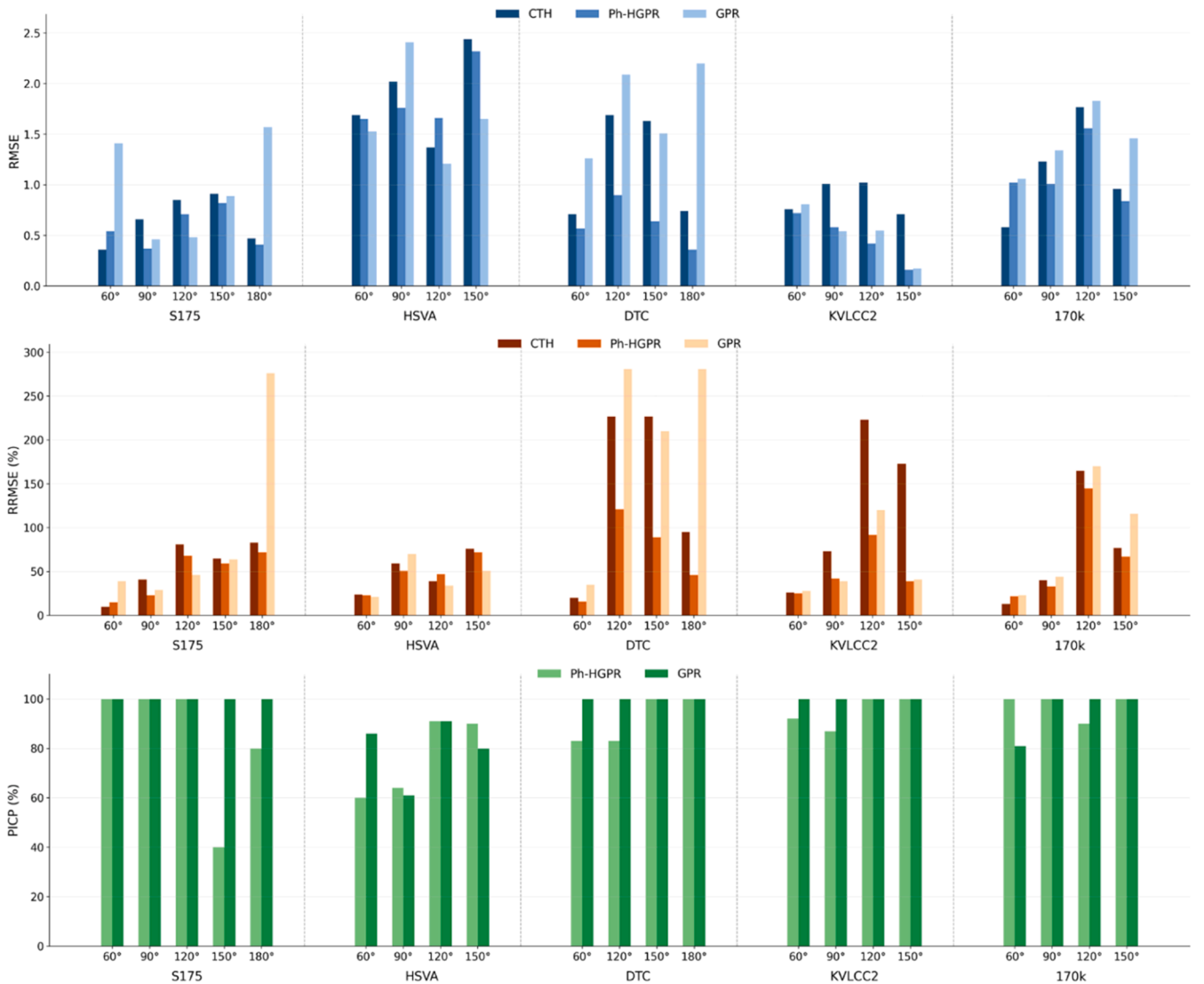
Fig. 9. Results of added wave resistance prediction for the 170k bulk carrier.

combinations of hull type, encounter angle, and Froude number. Across the 21 tested conditions, Phi-HGPR achieves the lowest error in 14 cases; when grouped by hull type, it performs best for DTC, KVLCC2, 170k, and S175. The only systematic exception is HSVA, where the basic GPR performs best on average and wins in three of the four angles. Apart from two isolated cases (S175-60° and 170k-60°) where CTH is superior, the hybrid approach provides the most accurate and robust C_{aw}

predictions across the tested conditions. The RRMSE for added wave resistance prediction at arbitrary wave headings generally increases as the encounter angle changes from near head sea conditions to following sea conditions. For all hull types at 60°, Phi-HGPR achieves very low RRMSE values, and it also gives acceptable results in most cases at 90°. When the wave angle is between 120° and 180°, the RRMSE values become much larger. For the CTH and GPR models, extremely large

Table 2RMSE, RRMSE and PICP of C_{aw} prediction using different methods.

Ship type	β (°)	F_n	CTH			PhI-HGPR			GPR		
			RMSE	RRMSE	PICP	RMSE	RRMSE	PICP	RMSE	RRMSE	PICP
S175	60	0.15	0.36	10%	—	0.54	15%	100%	1.41	39%	100%
S175	90	0.15	0.66	41%	—	0.37	23%	100%	0.46	29%	100%
S175	120	0.15	0.85	81%	—	0.71	68%	100%	0.48	46%	100%
S175	150	0.15	0.91	65%	—	0.82	59%	40%	0.89	64%	100%
S175	180	0.15	0.47	83%	—	0.41	72%	80%	1.57	276%	100%
HSVA	60	0.23	1.69	24%	—	1.65	23%	60%	1.53	21%	86%
HSVA	90	0.23	2.02	59%	—	1.76	51%	64%	2.41	70%	61%
HSVA	120	0.23	1.37	39%	—	1.66	47%	91%	1.21	34%	91%
HSVA	150	0.23	2.44	76%	—	2.32	72%	90%	1.65	51%	80%
DTC	60	0.05	0.71	20%	—	0.57	16%	83%	1.26	35%	100%
DTC	120	0.05	1.69	227%	—	0.9	121%	83%	2.09	281%	100%
DTC	150	0.05	1.63	227%	—	0.64	89%	100%	1.51	210%	100%
DTC	180	0.05	0.74	95%	—	0.36	46%	100%	2.2	281%	100%
KVLCC2	60	0.05	0.76	26%	—	0.72	25%	92%	0.81	28%	100%
KVLCC2	90	0.05	1.01	73%	—	0.58	42%	87%	0.54	39%	100%
KVLCC2	120	0.05	1.02	223%	—	0.42	92%	100%	0.55	120%	100%
KVLCC2	150	0.05	0.71	173%	—	0.16	39%	100%	0.17	41%	100%
170k	60	0.13	0.58	13%	—	1.02	22%	100%	1.06	23%	81%
170k	90	0.13	1.23	40%	—	1.01	33%	100%	1.34	44%	100%
170k	120	0.13	1.77	165%	—	1.56	145%	90%	1.83	170%	100%
170k	150	0.13	0.96	77%	—	0.84	67%	100%	1.46	116%	100%

**Fig. 10.** Comparison of RMSE, RRMSE and PICP of different methods.

values, even exceeding 200%, can be observed. This is because, when the wave angle approaches following sea conditions, the value of C_{aw} itself becomes very small, and in many cases its absolute value is less than 1. In addition, although the collected data correspond to the same hull type, the experiments were conducted at different facilities and under different test conditions, so the measured results may vary considerably. This is one of the major limitations of building models using towing tank data collected from published literature. For the unseen DTC hull, the extrapolation advantage of Phi-HGPR can be observed. Due to its limited extrapolation capability, the basic GPR performs very poorly for this hull type. In addition, the PICP metric provides a useful indication of the coverage of the predictive intervals. Except for a few individual cases, such as S175 at 150°, where the interval coverage is relatively poor, Phi-HGPR provides reasonable predictive intervals in most cases.

The performance of different models on wave resistance prediction for various hull types is summarized in Table 3. Overall, the basic GPR fails to learn the physical trend of added wave resistance under oblique sea conditions. The few “good” cases are incidental, driven by strong adherence to nearby samples (over-interpolation). Semi-empirical models have acceptable performance in many scenarios, but further enhancement is hard to achieve without risking a loss in established accuracy. By incorporating the semi-empirical prediction as a prior, the Phi-HGPR achieves better accuracy in most cases while preserving physically plausible curve shapes. This is expected: the semi-empirical model alone cannot resolve several complex wave-ship interaction mechanisms, whereas the learning component can capture the residual regularities that are omitted. Empirically, Phi-HGPR produces curves that slightly adjust the semi-empirical baseline upward or downward (with downward corrections prevailing), thus reducing prediction error without violating physical constraints. However, the semi-empirical formula itself may exhibit systematic bias for certain hull forms. If these hull forms are also sparsely represented in the training data, the residual GP may be unable to fully compensate for this bias, and the resulting predictions may remain strongly influenced by the prior and exhibit poor performance. In such cases, recalibration of the semi-empirical parameters, either offline or through online updating, may be necessary. However, the development and recalibration of the semi-empirical formula itself are beyond the scope of the present study.

The performance gap between Phi-HGPR and the basic GPR is from limited data: the basic GPR lacks sufficient local support for effective interpolation and has poor extrapolation, whereas Phi-HGPR is anchored by the physics-based prior in data-sparse regimes, preventing extrapolation drift. In most cases, the basic GP has an almost constant confidence band across the domain; in contrast, the Phi-HGPR exhibits input-dependent posterior standard deviations that reflect variability in the hydrodynamic mechanisms under different wave and speed conditions. Its predictive intervals narrow where the training constraints are

strong and residuals are small, and widen near regime changes, in data-sparse regions, or where residuals are large. Moreover, in most cases, the Phi-HGPR predictive intervals widen slightly near resonance, expand in the very-short-wave regime, and contract toward zero in the long wave limit. By contrast, Basic GPR's PIs are uniformly over-broad across the entire wavelength range and largely insensitive to the encounter angle.

4.2. Ablation study

An ablation study is conducted to further examine the role of the main components in the proposed framework. Four models are compared: Phi-HGPR with resampling, Phi-GPR with resampling, Phi-HGPR without resampling, and standard GPR without resampling. The parameters of the GPR component are kept the same for all models to reduce the influence of model settings on the results. In addition to RMSE, two additional metrics, mean absolute error (MAE) and the coefficient of determination (R^2), are introduced to evaluate the point-prediction performance,

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - a_i| \quad (23)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (p_i - a_i)^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \quad (24)$$

The results are presented in Table 4, which shows that the resampling strategy has a clear impact on prediction accuracy. When Phi-HGPR is trained without resampling, its RMSE, MAE, and R^2 are 1.52, 0.77, and 0.63, respectively. After introducing the resampling, the RMSE and MAE decrease by approximately 32% and 42%, respectively, while R^2 increases to 0.79. This indicates that the original imbalanced dataset indeed limits the prediction accuracy of the model, and that hybrid resampling can improve the model's ability to learn useful information from sparse oblique and following sea conditions. The role of heteroscedastic modelling can be observed by comparing Phi-HGPR with resampling and Phi-GPR with resampling. Their point-prediction performance is very similar, which indicates that heteroscedastic modelling has only a limited effect on the accuracy of the predicted mean. Its main contribution is to provide input-dependent predictive intervals, which is more evident in the case analysis in Section 4.1. Since the introduction of heteroscedasticity has only a minor influence on prediction accuracy, the effect of the physics-informed prior can be examined by comparing Phi-HGPR without resampling and GPR without resampling. Compared with the latter, the former reduces RMSE by approximately 14% and MAE by approximately 9%, while increasing R^2 by 0.09 in absolute terms. This shows that the physics-informed prior can also improve prediction accuracy in wave heading regions where the original data are relatively sparse.

5. Probabilistic assessment of ship arrival time and fuel consumption

Real-time estimation of a ship's ETA and fuel consumption is crucial for informed operational decision-making, effective berth scheduling, and optimized voyage planning under EEXI and CII constraints [42–44]. Waves are among the dominant influencing factors of variability of ETA

Table 3
Model performance of different methods.

Hull type	Model	Accuracy	Shape (physic)	Uncertainty
S175	Phi-HGPR	✓	✓	✓
	GPR	✓	✓	✓
	CTH	✓	✓	✓
HSVA	Phi-HGPR	✓	✓	✓
	GPR	✓	✓	✓
	CTH	✓	✓	✓
DTC	Phi-HGPR	✓	✓	✓
	GPR	✓	✓	✓
	CTH	✓	✓	✓
KVLCC2	Phi-HGPR	✓	✓	✓
	GPR	✓	✓	✓
	CTH	✓	✓	✓
170k	Phi-HGPR	✓	✓	✓
	GPR	✓	✓	✓
	CTH	✓	✓	✓

Table 4
Performance metrics of different models.

Model	RMSE	R^2	MAE
Phi-HGPR with resampling	1.03	0.79	0.44
Phi-GPR with resampling	1.08	0.78	0.49
Phi-HGPR without resampling	1.52	0.63	0.77
GPR without resampling	1.78	0.54	0.85

and fuel consumption. In previous practice, added wave resistance is treated as a point estimate, with uncertainty neglected. As a result, the reliability of that estimate cannot be quantified. In this study, the proposed PhI-HGPR model provides a probabilistic characterization of the added wave resistance, which serves as a stochastic input to the simulation. By propagating these uncertainties through the ship's power-speed relationship, a distribution for ETA and fuel use can be obtained, thus enabling the optimization of the trade-off between speed and fuel consumption under the given constraints.

The following case study demonstrates this integration using continuous full-scale monitoring data from a chemical tanker on a trans-Atlantic route, as shown in Fig. 11 and Table 5. The voyage is discretized into multiple legs, and for each leg the mean metocean conditions are combined with the PhI-HGPR-predicted distribution of added wave resistance. The observed sailing time for each leg is used in the timing likelihood to sequentially reweight the candidate power profiles. The resulting ensemble is then used to construct predictive distributions of ETA and fuel consumption conditioned on the observations, while the parameters of the ship performance model remain fixed. The resulting framework can support risk-aware and data-adaptive decision making for ship operators and traffic coordinators.

5.1. Modelling for ship performance

In the applications, the ship's ETA and fuel consumption are evaluated under realistic environmental conditions as functions of the operating speed and the encountered weather [45–48]. A ship performance model that accounts for uncertainty in wave resistance needs to be developed [49]. Since ETA can be easily computed when the speed and voyage distance are known, this section mainly describes the estimation of fuel consumption. The purpose is to define a function f_{SP} that can compute fuel consumption based on weather conditions and sailing speed:

$$FC(V, \mathbf{W}) = f_{SP}(V, \mathbf{W}) \quad (25)$$

where V denotes the ship speed and $\mathbf{W}$ collects the metocean variables, including significant wave height, wave period, mean wave direction, relative wind speed, and relative wind direction.

$$FC(V, \mathbf{W}) = P(V, \mathbf{W}) \times SFOC \quad (26)$$

where P is the main engine power and the term $SFOC$ represents the specific fuel oil consumption. The corresponding effective power P_E can be obtained as:

$$P_E(V, \mathbf{W}) = P(V, \mathbf{W}) \times \eta = (R_{calm}(V) + R_{AA}(V, \mathbf{W}) + R_{aw}(V, \mathbf{W})) \times V \quad (27)$$

where η is a power transmission factor that aggregates several efficiencies (shaft transmission, propeller open-water efficiency, rotative efficiency, and hull efficiency). In Eq. (27), the effective power is expressed in terms of the total resistance, which is decomposed into calm water resistance R_{calm} , added wind resistance R_{AA} , and added wave



Fig. 11. The ship for the case study.

Table 5

Main parameters for chemical tanker.

Main parameters	Value
Length L [m]	176
Breadth B [m]	32
Designed draft T [m]	11
Maximum continuous power [kW]	8200
Deadweight DWT [tons]	46067
Wetted surface [m ²]	8166
Non-dimensional radius of gyration in yaw K_{yy} [-]	0.25
Block coefficient C_B [-]	0.80

resistance R_{aw} . The calm water component is dominated by viscous friction and wave-making effects around the hull. Following the semi-empirical ITTC recommendation [50], the calm water resistance is written as:

$$R_{calm} = \frac{1}{2} C_f \rho_w S_{wet} V_{cw}^2 \quad (28)$$

where C_f is the resistance coefficient derived from towing-tank tests, ρ_w is the seawater density, S_{wet} is the wetted surface area, and V_{cw} is the ship speed in calm water.

The additional resistance due to wind acting on the hull above the waterline and on the superstructure is estimated using the ISO [51] formulation, given as:

$$R_{AA} = \frac{1}{2} \rho_A [C_A(\psi) A_{XV} V_{wr}^2 - C_A(0) A_{XV} V_G^2] \quad (29)$$

where ρ_A is the air density, A_{XV} is the projected transverse area above the waterline, V_{wr} is the relative wind speed, ψ is the relative wind angle with respect to the ship heading, and V_G is the ship's speed over ground. The wind-resistance coefficient C_A is obtained from wind-tunnel measurements supplied by the shipowner.

For waves, the mean resistance in an irregular sea state, $\bar{R}_{aw}$, is evaluated by superposing the contributions from regular-wave components. Specifically, the transfer function of added resistance in regular waves, $R_{aw}(\omega)$, is first determined with the proposed PhI-HGPR model at different speeds. This transfer function is then combined with the wave energy spectrum $S(\omega)$, which characterizes the actual sea state through the significant wave height H_s and a characteristic period T . By integrating the regular-wave responses over all frequencies, the added wave resistance can be calculated as:

$$\bar{R}_{aw}(V, \mathbf{W}) = 2 \int_0^\infty S(\omega | \mathbf{W}) \frac{R_{aw}(\omega | V, \mathbf{W})}{\zeta_a(\omega)^2} d\omega \quad (30)$$

where R_{aw} can be calculated as:

$$R_{aw}(\omega | V, \mathbf{W}) = C_{aw}(\omega | V, \mathbf{W}) \rho_w g \zeta_a^2 B^2 / L \quad (31)$$

where C_{aw} denotes the nondimensional transfer function of added resistance in regular waves, B is the ship's breadth, and ζ_a is the wave amplitude. The JONSWAP spectrum is adopted to describe $S(\omega)$:

$$S(\omega) = \frac{320 H_s^2}{T_p^4 \omega^5} \exp\left(\frac{-1950}{T_p^4 \omega^4}\right) \gamma \exp\left[\frac{-(\omega - \omega_p)^2}{2 \sigma_J^2 \omega_p^2}\right] \quad (32)$$

The spectrum is parameterized by the significant wave height H_s , peak period T_p , peak enhancement factor γ , spectral peak frequency ω_p , and spectral width parameter σ_J . The calculation for ship performance can be summarized in Fig. 12.

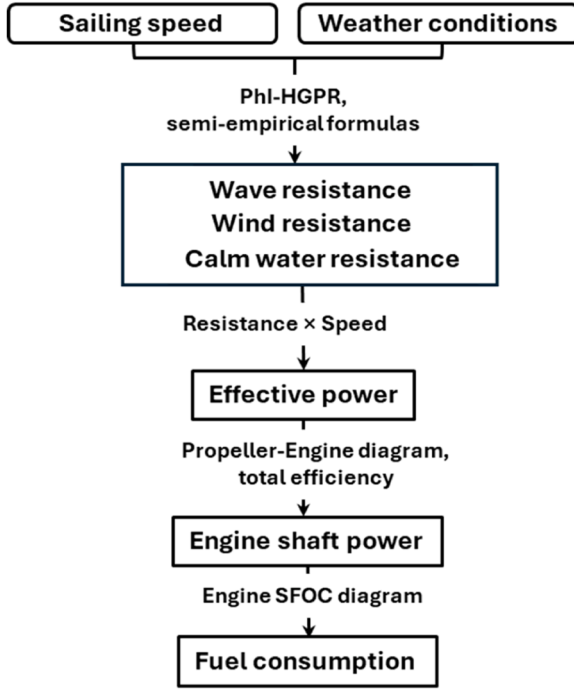


Fig. 12. Ship fuel consumption calculation flowchart.

5.2. Sequential Bayesian reweighting for probabilistic ETA and fuel-consumption estimation

In this case study, we apply the proposed added wave resistance prediction model together with the formulas discussed in Section 5.1 to a North Atlantic passage [52]. Continuous full-scale monitoring data are available from 07:00 on 14 November 2018 to 03:30 on 24 November 2018 (UTC), from 41.386° N, 71.416° W to 35.959° N, 5.472° W, for a total distance of about 3110 nm. The track is discretized into nineteen legs, as shown in Fig. 13. For each leg, the local wind, wave and current conditions are approximated by the mean of the values at its two bounding waypoints, and the Phi-HGPR model provides the corresponding distribution of added wave resistance. A sequential Bayesian reweighting scheme is constructed, in which the main engine power on each leg is treated as a latent operational variable. The observed leg-wise sailing times are assimilated through the timing likelihood to update the posterior weights of candidate power profiles. The resulting reweighted ensemble is then used to derive predictive distributions of total ETA and total fuel consumption, while the parameters of the ship performance model remain fixed.

5.2.1. Real-time ETA estimation

The ETA is central to berth-window allocation and pilotage scheduling. It also affects schedule reliability, demurrage costs, post-call planning, and just-in-time arrival regimes for anchorage waiting.

Accordingly, predictive intervals for the ETA should be estimated to support subsequent decision making. In the present estimation framework, the main engine power on each leg is treated as a latent variable. Let P_k denote the main engine power on leg k and collect them in the vector $P = (P_1, \dots, P_{19})$. For any candidate power vector P , the ship performance model of Section 5.1 is combined with a stochastic prediction of the added wave resistance from the Phi-HGPR model. On each leg, a random draw of added resistance is taken from the Phi-HGPR posterior for the prevailing metocean conditions. Together with calm water and wind resistance, this determines the total resistance, the achievable steady speed, the leg time, and the required power. Candidates that would violate the maximum continuous power or other operational limits are ruled out and thereby shape the ETA distribution. It should be noted that the present study focuses on the uncertainty of wave resistance and its influence on voyage states, while other sources of uncertainty are beyond the scope of the present work.

At departure an informative joint prior $p(P)$ is specified, centred around the most common operating power and with correlation between neighbouring legs. A correlation coefficient ρ is introduced to control the smoothness of the candidate power profiles between neighbouring legs. A Monte Carlo ensemble of candidate pairs $(\rho^{(s)}, P^{(s)})$ is drawn from this prior. For each sample, the ship performance and resistance models are used together with the leg-wise metocean conditions and a sampled added resistance on each leg to compute the sailing times t_k^s . In this way, uncertainty in the added wave resistance is propagated into the simulated leg times.

After completing leg k , the realised travel time on that leg, denoted t_k , is assimilated. If we denote the observations up to waypoint k by $D_k = \{t_1, \dots, t_k\}$, the posterior distribution of P and ρ based on Bayesian theory can be written as:

$$p(P, \rho | D_k) = \frac{p(t_k | P, \rho) p(P, \rho | D_{k-1})}{p(t_k | D_{k-1})} \quad (33)$$

The per-leg timing likelihood $\mathcal{L}_k^{(s)}$ is modelled as a Gaussian error on the leg time. For the s -th sample this likelihood is

$$\mathcal{L}_k^{(s)} = \mathcal{N}(t_k^{\text{obs}} | t_k^{(s)}, \sigma_t^2) \quad (34)$$

where t_k^{obs} is the measured sailing time on leg k , $t_k^{(s)}$ is the simulated time for the s -th candidate pair, and σ_t^2 is a variance term collecting observation error and model discrepancy. The posterior weight of sample s after assimilating leg k is then updated as:

$$w_k^{(s)} = \frac{C_k^{(s)} \mathcal{L}_k^{(s)} w_{k-1}^{(s)}}{\sum_{r=1}^S C_k^{(r)} \mathcal{L}_k^{(r)} w_{k-1}^{(r)}} \quad (35)$$

where $w_{k-1}^{(s)}$ is the previous weight and $C_k^{(s)} \in \{0, 1\}$ discards candidates which, for the sampled added resistance on leg k , would violate the maximum continuous power. The weighted ensemble $\{\rho^{(s)}, P^{(s)}, w_k^{(s)}\}$ approximates the posterior distribution $p(P, \rho | D_k)$ of the engine powers on the legs and the correlation factor, with the influence of added wave resistance uncertainty embedded through the sampled sailing times. For

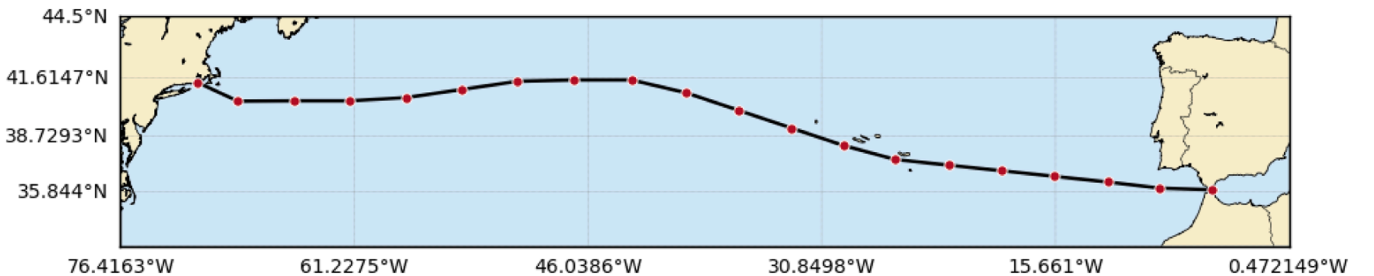


Fig. 13. Voyage for ETA and fuel consumption inference.

each sample s , the corresponding total arrival time after finishing leg k is

$$\text{ETA}_k^{(s)} = t_1 + \dots + t_k + t_{k+1}^{(s)} + \dots + t_{19}^{(s)} \quad (36)$$

The posterior ETA distribution at waypoint k is given by the weighted empirical distribution of $\text{ETA}_k^{(s)}$. Under this scheme, the ETA distribution changes for two reasons: observed sailing times replace simulated values for the completed legs, and the timing likelihood reweights the candidate power profiles for the remaining legs. The stochastic evaluation of future leg times continues to represent the propagated Phi-HGPR added-resistance uncertainty, as shown in Fig. 14.

It should be noted that the Bayesian updating in this study is formulated only as voyage operational state inference. The parameters in the ship performance model are kept fixed throughout the voyage and are not recalibrated online using the assimilated observations. This means that, when a discrepancy arises between the simulated and observed leg-wise sailing times, it is not represented by modifying the physical model parameters. Instead, the discrepancy is accounted for by updating the posterior weights of the candidate power profiles through the timing likelihood function.

Fig. 15 presents the posterior distributions of total ETA at each waypoint along the 19-leg route. At departure, the distributional spread mainly reflects the prior variation among the candidate leg-wise power profiles and the sampled Phi-HGPR added-resistance uncertainty. As the vessel progresses, observed sailing times replace simulated values for the completed legs, while the timing likelihood reweights the candidate power profiles for the remaining legs. These two effects cause the ETA distribution to become progressively narrower. The slight downward shift in the median suggests that the pre-voyage ETA estimates are mildly conservative. Near the destination, the distribution becomes highly concentrated within the uncertainty sources represented by the present framework

5.2.2. Real-time fuel consumption estimation

As in the ETA estimation, the engine power on each leg is treated as a latent variable, and the added wave resistance predicted by the Phi-HGPR model is sampled and propagated through the performance model. In this section, the same stochastic ingredients are used, but the focus is on mapping the resulting powers and times to fuel consumption instead of arrival time [53]. Given one joint candidate sample $(\rho^{(s)}, P^{(s)})$ on all legs, the resistance model is evaluated jointly on the legs to obtain a multivariate normal posterior for the vector of added wave resistances. This posterior is written as

$$R_{aw}^{(s)} \sim N(m(X^{(s)}), \Sigma(X^{(s)})) \quad (37)$$

where $R_{aw}^{(s)}$ stacks the per-leg added resistances for sample s , $m(X^{(s)})$ gives their joint means, and $\Sigma(X^{(s)})$ denotes their joint covariance (diagonal

terms represent per-leg variance and off-diagonal terms encode cross-leg correlation). A correlated draw $R_{aw}^{(s)}$ from this posterior is then combined with the power vector $P^{(s)}$. For each leg k , the hydrodynamic model of Section 5.1 and the power-resistance balance are used to obtain the total resistance $R_{t,k}^{(s)}$, the steady speed $V_k^{(s)}$, and the leg time $t_k^{(s)}$ associated with sample s and the sampled added resistance. The engine power on leg k is computed according to:

$$P_k^{(s)} = \frac{R_{t,k}^{(s)} V_k^{(s)}}{\eta} \quad (38)$$

In this way, both the simulated leg times $t_k^{(s)}$ and the powers required to overcome the total resistance inherit the uncertainty in the added wave resistance as well as the uncertainty in the engine powers. Fuel consumption on leg k for sample s is then evaluated using the engine fuel-consumption map described earlier, which is written as:

$$F_k^{(s)} = f(P_k^{(s)}, t_k^{(s)}, \phi) \quad (39)$$

where $f(\cdot)$ denotes the fuel model, namely specific fuel oil consumption applied to power over time, and ϕ collects its parameters. The sequential updating of the power samples and their weights over waypoints is identical to that in the ETA estimation. After each leg, the observed sailing time is assimilated using the Gaussian timing likelihood, the weights $w_k^{(s)}$ are updated and normalised, and, if necessary, resampling and rejuvenation are applied. When propagating the ensemble to predict fuel consumption on the remaining legs, the Phi-HGPR posterior is re-evaluated at the corresponding metocean conditions and a new correlated draw $R_{aw}^{(s)}$ is taken for each sample. Thus, at every waypoint, resistance uncertainty is newly injected into the prediction of future fuel use.

At waypoint k , the posterior predictive distribution for total fuel consumption is obtained from the weighted ensemble by combining the fuel already burned on the completed legs with the predicted fuel on the remaining legs. For sample s , the total fuel is given by:

$$F^{(s)} = F_0^{(s)} + \dots + F_k^{(s)} + F_{k+1}^{(s)} + \dots + F_{19}^{(s)} \quad (40)$$

where $F_0^{(s)}$ is the observed fuel on the sailed legs. The weighted empirical distribution of the $\{F^{(s)}\}$ provides the posterior distribution of total fuel consumption. The posterior distributions of total fuel for the voyage are shown in Fig. 16. At departure the distribution is already centred close to the value obtained near the destination, and this central estimate changes only slightly along the route. As more legs are completed and the observed fuel use accumulates, the violins gradually become narrower, showing a steady reduction in uncertainty. Near the destination the distributions are much more concentrated than at the start but still

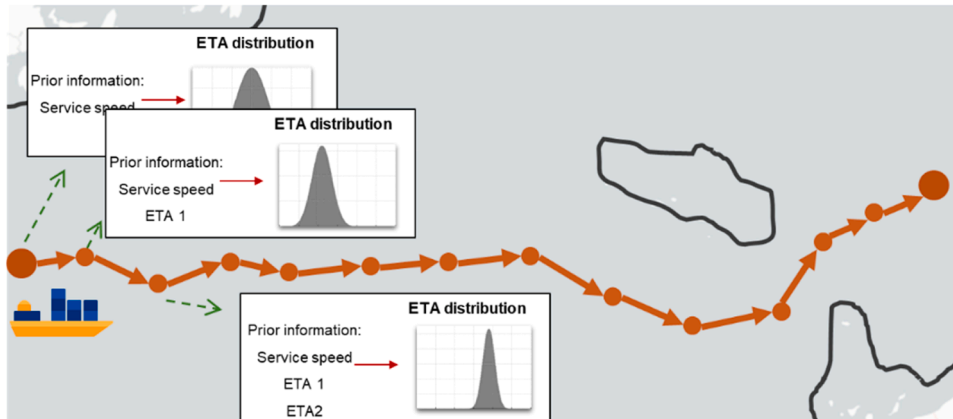


Fig. 14. The theory of real-time ETA estimation.

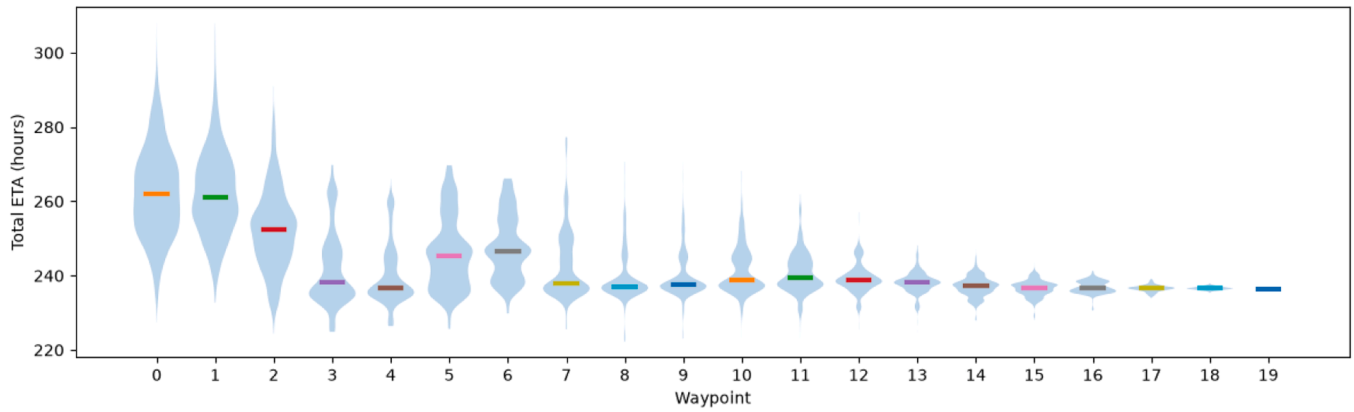


Fig. 15. Result of ETA inference.

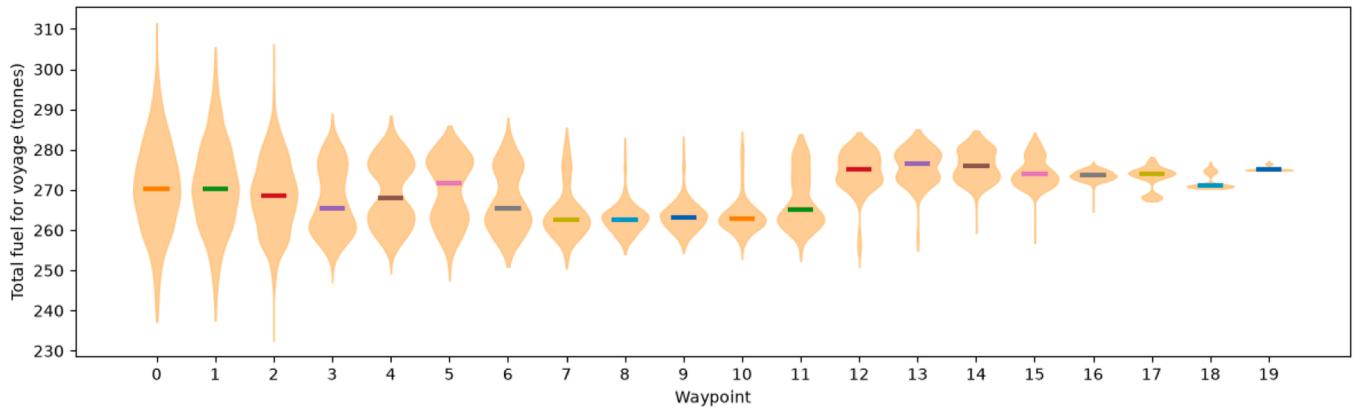


Fig. 16. Result of fuel consumption inference.

span several tonnes of fuel, which reflects the fact that total fuel is an accumulated quantity and therefore retains the effect of remaining uncertainties in power demand and added resistance over all legs.

6. Conclusions

This study presented a physics-informed heteroscedastic Gaussian Process Regression (Phi-HGPR) for predicting added wave resistance at arbitrary encounter angles and propagating this uncertainty to ETA and fuel consumption estimates. By embedding a semi-empirical baseline in the GP mean and learning residual corrections with input-dependent noise, the model remains physically consistent while improving accuracy and uncertainty calibration. Across five representative hulls, i.e., S175, HSVA, DTC, KVLCC2, 170k bulk carrier, Phi-HGPR achieved the lowest RMSE in 14 of 21 hull-angle-speed cases, while preserving expected curve shapes. Notable improvements were observed for KVLCC2 at 150° and for DTC at 120° and 180°; under these conditions, Phi-HGPR achieved much lower RMSE values than the other methods. For the S175 hull, Phi-HGPR matched or beat alternatives at most angles and produced PIs that covered most measurements, whereas the basic GPR often lost shape fidelity (e.g., nearly flat predictions at oblique headings). For the 170k bulk carrier, Phi-HGPR avoided the forward-shifted peaks seen with GPR and maintained plausible trend behavior. The principal exception was HSVA, where the basic GPR, likely aided by similar hulls in the training set, performed best at most angles. Even there, Phi-HGPR retained tighter and more informative uncertainty bands than other methods and provided adaptive corrections (including upward adjustments where warranted). In the voyage application, observed leg-wise sailing times were used to sequentially reweight candidate power profiles, while the parameters of the ship performance model remained

fixed. The reweighted ensemble yielded observation-conditioned predictive distributions of ETA and fuel consumption. These distributions became narrower along the route because completed-leg observations replaced simulated values and the remaining candidate profiles were reweighted through the timing likelihood. Near arrival, the ETA distribution became highly concentrated within the uncertainty sources represented by the present framework, while the fuel consumption distribution retained some spread because of the accumulated uncertainty over the remaining legs.

Several limitations should be acknowledged. Publicly available oblique wave model test data remain scarce, and resampling can only partially alleviate this imbalance. In addition, using a single semi-empirical prior over the full wavelength range may limit performance, since different formulations may be more suitable for short and long wave regimes. The ship performance model is kept fixed during Bayesian updating, meaning that discrepancies between simulated and observed values are not used to recalibrate model parameters online. Finally, the ETA and fuel consumption assessment does not consider the uncertainty of the wave and metocean data themselves.

Future work will focus on enriching the database with validated high fidelity CFD results. Different resampling strategies will be compared. Accounting for other uncertainty sources in metocean inputs, considering the geographical and seasonal coverage of the full-scale voyage data, extending the current framework to calibrate both the semi-empirical prior and the ship performance model through offline or online hierarchical Bayesian methods using voyage observations would also be interesting directions for future investigation.

CRediT authorship contribution statement

Chi Zhang: Writing – original draft, Validation, Methodology, Investigation. **Shanshan Fu:** Investigation, Conceptualization. **Min-gyang Zhang:** Validation, Methodology, Investigation. **Wengang Mao:** Writing – review & editing, Supervision, Funding acquisition. **Tsoula-kos Nikolaos:** Validation, Data curation.

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.

Acknowledgements

This work was supported by Trafikverket (the Swedish Transport Administration) under Grant No. TRV2023/98101 and by the IAMU Project FY2026-4. The authors are grateful to the ship owners for providing the full-scale measurement data.

Data availability

The authors do not have permission to share data.

References

- [1] Wang H, Yan R, Au M, Wang S, Jin Y. Federated learning for green shipping optimization and management. *Adv Eng Inform* 2023;56:101994.
- [2] IMO (2023). IMO strategy on reduction of GHG emissions from ships. 2023.
- [3] Boni E, Montagni M, Moreschi A, Pugi L. Energy storage system optimization for an autonomous SailBoat. In: Proceedings of the 2019 5th international forum on research and technologies for society and industry: innovation to shape the future, RTSI 2019; 2019. p. 231–5.
- [4] Godet A, Wallner L, Panagakos G, Barfod M. Developing correction factors for weather's influence on the energy efficiency indicators of container ships using model-based machine learning. *Ocean Coast Manag* 2024;258:107390.
- [5] Moreno-Gutiérrez J, Calderay F, Saborido N, Boile M, Rodríguez Valero R, Durán-Grados V. Methodologies for estimating shipping emissions and energy consumption: a comparative analysis of current methods. *Energy* 2015;86:603–16.
- [6] Wang K, Liang H, Liu X, Li Z, Jing Z, Chi Y, et al. Investigations and applications of carbon control technology: the roadmap to zero-carbon shipping. *Ocean Eng* 2025; 339:121898.
- [7] Fan A, Wang Y, Yang L, Yang Z, Hu Z. A novel grey box model for ship fuel consumption prediction adapted to complex navigating conditions. *Energy* 2025; 315:134436.
- [8] IMO (2012). Draft guidelines for the calculation of the coefficient f_w for decrease in ship speed in a representative sea condition. In: MEPC 64/4/38. 2.
- [9] Yuan J, Nian V, He J, Yan W. Cost-effectiveness analysis of energy efficiency measures for maritime shipping using a metamodal based approach with different data sources. *Energy* 2019;189:116205.
- [10] Liu S, Papanikolaou A. Fast approach to the estimation of the added resistance of ships in head waves. *Ocean Eng* 2016;112:211–25.
- [11] Lang X, Mao W. A semi-empirical model for ship speed loss prediction at head sea and its validation by full-scale measurements. *Ocean Eng* 2020;209:107494.
- [12] Kim Y, Esmailian E, Steen S. A meta-model for added resistance in waves. *Ocean Eng* 2022;266:112749.
- [13] Cepowski T. The prediction of ship added resistance at the preliminary design stage by the use of an artificial neural network. *Ocean Eng* 2020;195:106657.
- [14] Cepowski T. The use of a set of artificial neural networks to predict added resistance in head waves at the parametric ship design stage. *Ocean Eng* 2023;281: 114744.
- [15] Duan W, Yang K, Huang L, Jing Y, Ma S. A DFN-based method for fast prediction of ships' added resistance in heading waves. *Ocean Eng* 2022;245:110484.
- [16] Zhang C, Vergara D, Zhang M, Tsoulakos N, Mao W. A machine learning method to evaluate head sea induced weather impact on ship fuel consumption. *Energy* 2025; 328:136533.
- [17] Zhu K, Huang C, Li S, Lin X. Physics-informed Gaussian process for tool wear prediction. *ISA Trans* 2023;143:548–56.
- [18] Xu Y, Kohtz S, Boakye J, Gardoni P, Wang P. Physics-informed machine learning for reliability and systems safety applications: state of the art and challenges. *Reliab Eng Syst Saf* 2023;230:108900.
- [19] Kanazawa M, Skulstad R, Wang T, Li G, Hatledal L, Zhang H. A physics-data co-operative ship dynamic model for a docking operation. *IEEE Sens J* 2022;22(11): 11173–83.
- [20] Deo I, Venkateshwaran A, Jaiman R. A physics-guided probabilistic surrogate modeling framework for digital twins of underwater radiated noise. *Ocean Eng* 2026;358:125461.
- [21] Jiang C, Zhong T, Choi H, Youn B. Physics-informed Gaussian process probabilistic modeling with multi-source data for prognostics of degradation processes. *Reliab Eng Syst Saf* 2025;258:110893.
- [22] Ko D, Nam B. Modeling the influence of forward speed on ship maneuverability with a physics-guided gaussian process regression. In: Proceedings of the ASME 2026 45th international conference on ocean, offshore and arctic engineering; 2026. OMAE2026.
- [23] Li M, Jia G, Mahmoud H, Yu Y, Tom N. Physics-constrained Gaussian process model for prediction of hydrodynamic interactions between wave energy converters in an array. *Appl Math Model* 2023;119:465–85.
- [24] Elshafei B, Popov A, Giddings D. Enhanced offshore wind resource assessment using hybrid data fusion and numerical models. *Energy* 2024;310:133208.
- [25] Marrel A, Iooss B. Probabilistic surrogate modeling by Gaussian process: a review on recent insights in estimation and validation. *Reliab Eng Syst Saf* 2024;247: 110094.
- [26] Zhang M, Kujala P, Musharraf M, Zhang J, Hirdaris S. A machine learning method for the prediction of ship motion trajectories in real operational conditions. *Ocean Eng* 2023;283:114905.
- [27] Xu Y, Bansal P, Wang P, Li Y. Physics-informed machine learning for system reliability analysis and design with partially observed information. *Reliab Eng Syst Saf* 2025;254:110598.
- [28] Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. *Nat Rev Phys* 2021;3:422–40.
- [29] Fujii H, Takahashi T. Experimental study on the resistance increase of a large full ship in regular oblique waves. *J Soc Nav Archit Jpn* 1975;1975(137):132–7.
- [30] ITTC (2014). Preparation, conduct and analysis of speed/power trials (incl. Appendix D: Wave correction STAwave-1/2). ITTC Recommended Procedures and Guidelines.
- [31] ITTC. Preparation, conduct and analysis of speed/power trials (Revision 07). ITTC Recommended Procedures and Guidelines; 2017.
- [32] Faltinsen O. Prediction of resistance and propulsion of a ship in a seaway. In: Proceedings of the 13th symposium on naval hydrodynamics; 1980. p. 505–29.
- [33] Jinkine V, Ferdinand V. A method for predicting the added resistance of fast cargo ships in head waves. *Int Shipbuild Prog* 1974;21(238):149–67.
- [34] Tsujimoto M, Shibata K, Kuroda M, Takagi K. A practical correction method for added resistance in waves. *J Jpn Soc Nav Archit Ocean Eng* 2008;8:177–84.
- [35] Lang X, Mao W. A practical speed loss prediction model at arbitrary wave heading for ship voyage optimization. *J Mar Sci Appl* 2021;20(3):410–25.
- [36] Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Mach Learn* 2021;110(3): 457–506.
- [37] Goldberg P, Williams C, Bishop C. Regression with input-dependent noise: a Gaussian process treatment. *Adv Neural Inf Process Syst* 1998;10:493–9.
- [38] Kersting K, Plagemann C, Pfaff P, Burgard W. Most likely heteroscedastic Gaussian process regression. In: Proceedings of the 24th international conference on machine learning; 2007. p. 393–400.
- [39] Xia M, Fan A, Hu Z, Yi Q, Vladimir N, Mao W. Data augmentation-based approach to enhance the accuracy, generalization, and reliability of ship fuel consumption prediction. *Ocean Eng* 2025;341:122558.
- [40] Chua P, Chieng R, Wang R, Lee K, Xu H, Fu X, Yan R. Development of a knowledge-based noise test for robustness of vessel fuel consumption prediction models. *Marit Policy Manag* 2025;52(8):1178–207.
- [41] Kou P, Liang D, Gao L, Lou J. Probabilistic electricity price forecasting with variational heteroscedastic Gaussian process and active learning. *Energy Convers Manag* 2015;89:298–308.
- [42] Karatug Ç, Tadmor M, Ventura M, Soares C. Decision support system for ship energy efficiency management based on an optimization model. *Energy* 2024;292:130318.
- [43] Luo X, Yan R, Xu L, Wang S. Accuracy and applicability of ship's fuel consumption prediction models: a comprehensive comparative analysis. *Energy* 2024;310: 133187.
- [44] Yan R, Chu Z, Wu L, Wang S. Predicting vessel service time: a data-driven approach. *Adv Eng Inform* 2024;62:102718.
- [45] Fossen T. Handbook of marine craft hydrodynamics and motion control. John Wiley & Sons; 2011.
- [46] Tillig F, Ringsberg J, Mao W, Ramne B. A generic energy systems model for efficient ship design and operation. *Proc Inst Mech Eng Part M J Eng Marit Environ* 2017;231(2):649–66.
- [47] Wang K, Liu X, Guo X, Wang J, Wang Z, Huang L. A novel high-precision and self-adaptive prediction method for ship energy consumption based on the multi-model fusion approach. *Energy* 2024;310:133265.
- [48] Zhang C, Zhang D, Zhang M, Zhang J, Mao W. A three-dimensional ant colony algorithm for multi-objective ice routing of a ship in the Arctic area. *Ocean Eng* 2022;266:113241.
- [49] ISO. ISO 15016:2025 Ships and marine technology—Guidelines for the assessment of speed and power performance by analysis of speed trial data. International Organization for Standardization; 2025.
- [50] ITTC (2002). Resistance uncertainty analysis example for resistance test, recommended procedures 7.5-02-02-02.
- [51] ISO (2015). ISO 15016:2015 - annex B: beaufort scale for wind velocity.
- [52] Chi H, Pedrielli G, Ng SH, Kister T, Bressan S. A framework for real-time monitoring of energy efficiency of marine vessels. *Energy* 2018;145:246–60.
- [53] Zhang M, Tsoulakos N, Kujala P, Hirdaris S. A deep learning method for the prediction of ship fuel consumption in real operational conditions. *Eng Appl Artif Intell* 2024;130:107425.