This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2021.3073462, IEEE Transactions on Industrial Informatics

An Uncertainty-aware Hybrid Approach for Sea State Estimation Using Ship Motion Responses

Peihua Han, *Student Member, IEEE,* Guoyuan Li, *Senior Member, IEEE,* Xu Cheng, *Member, IEEE,* Stian Skjong, and Houxiang Zhang, *Senior Member, IEEE*

***Abstract*—Understanding the current environmental conditions is essential for autonomous ships, among which real-time estimation of sea conditions is a key aspect. Considering the ship as a large wave buoy, the sea state can be estimated from motion responses without extra sensors installed. This task is challenging since the relationship between the wave and the ship motion is hard to model. Existing methods include a wave buoy analogy (WBA) method, which assumes linearity between wave and ship motion, and a machine learning (ML) approach. Since the data collected from a vessel in the real world is typically limited to a small range of sea states, the ML method might fail when the encountered sea state is not in the training dataset. This paper proposes a hybrid approach that combined the two methods above. The ML method is compensated by the WBA method based on the uncertainty of estimation results and, thus, the failure can be avoided. Real-world historical data from the Research Vessel (RV) Gunnerus are applied to validate the approach. Results indicate that the hybrid approach improves the estimation accuracy.**

***Index Terms*—Sea state estimation, autonomous ship, supervised machine learning, hybrid method.**

I. Introduction

REMOTELY operated and autonomous ships are a topic of increasing interest in the maritime industry [1]. These ships have the potential to reduce human-based errors, lower fuel consumption, and extend the operational window [2]. Efforts have been made in the recent years to develop modern control [3] and path planning algorithm [4] for marine vehicles. Nonetheless, these autonomous systems must be able to process the current environmental conditions for safe and effective decision making. For marine vessels, the external sea loads are crucial for their control and operation [5]. Real-time estimation of sea states is therefore of key importance for autonomous vessels.

The sea state refers to the general condition of the ocean with respect to wind waves and swell at a certain location in oceanography. A sea state is usually characterized by statistical parameters, e.g., significant wave height, average wave frequency, and peak frequency [6]. The primary tool nowadays to collect accurate statistical wave data is floating wave buoys. However, wave buoys are deployed at fixed locations and they are not practical for a vessel in maneuvering operations. Other methods include meteorological satellite and wave radar. The meteorological satellite image quality is often subjected to a time delay of several hours and could be affected by cloudy weather. A wave radar satisfies the need, but it is expensive to install, requires frequent calibration [7], and is only equipped to a limited number of vessels.

Nowadays, the majority of marine vessels are equipped with sensors that measure the ship motions in 6 degrees of freedom. The motion responses reflect the sea state conditions and therefore a ship can be considered as a large wave buoy. From this perspective, a vessel is essentially equipped with an environmental condition estimation system [8]. Estimating the sea state based on the ship motion responses is of interest and has been investigated in the literature. Several challenges exist to estimating the sea state using motion data: (1) ocean waves are stochastic processes and they are usually described by statistical parameters; (2) It is difficult to model the relationship between wave and ship motion; (3) extra complexity is added due to the moving of the vessel. Previous works involve model-based methods that use response amplitude operators (RAOs) to relate the sea state to vessel responses. RAOs are complex-valued transfer functions that are calculated using strip theory and sometimes computational fluid dynamics. Ship responses are, in general, non-linearly related to the wave excitation. However, the transfer functions are linear and therefore only valid for light and moderate sea states [9]. Besides, the RAOs are difficult to estimate exactly and might need to be tuned with real-world data. On the other hand, this task can be posed as a supervised machine learning problem and several data-driven approaches have been employed to learn the mapping from measured ship motion responses to an actual sea state [10]. The advantage of these approaches is that they are able to discover the pattern between ship motions and sea states based on historical experience.

However, machine learning methods often require an extensive amount of training data and they only perform well when the training and testing data are sampled independently and identically from the same distribution [11]. In other words, models in deployment can fail catastrophically when the test data distribution differs from the distribution of the training data [12]. The vessel is usually operated in the same route for a specific period, the historical data collected in the real world, therefore, contains limited number of sea state and can not

Manuscript received January 14, 2021; revised March 24, 2021; accepted April 9, 2021. This work was supported by a grant from the Research Council of Norway through the Knowledge-Building Project for industry "Digital Twins For Vessel Life Cycle Service" (Project nr. 280703) and a grant from the IKTPLUSS Project "Remote Control Centre for Autonomous Ship Support" (Project nr: 309323). *(Corresponding author: Guoyuan Li, Houxiang Zhang.)*

Peihua Han, Guoyuan Li, Xu Cheng, and Houxiang Zhang are with the Department of Ocean Operations and Civil Engineering, Norweigian University of Science and Technology (NTNU), Aalesund 6009, Norway. (e-mail: peihua.han@ntnu.no; guoyuan.li@ntnu.no; xu.cheng@ieee.org; hozh@ntnu.no).

Stian Skjong is with SINTEF Ocean, Trondheim 7010, Norway. (Stian.skjong@sintef.no).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2021.3073462, IEEE Transactions on Industrial Informatics

cover the entire range of possible sea states. When the vessel is deployed into a new route or experience a new sea state, the machine learning model trained with historical data is likely to fail. A failing on the sea state estimation might cause severely operational and financial costs. Ideally, a machine learning model should be able to provide not only the predictions but also how much confidence it has in the predictions. There are existing models that can directly provide or approximate the uncertainty [13], [14]. When the predictive uncertainty can be accounted for, the model-based method can be utilized to compensate for the prediction with high uncertainty.

In this paper, the feasibility of the hybrid approach for sea state estimation using ship motion responses will be investigated. The ML model estimates the current sea state with predictive uncertainty, while, in parallel, the wave buoy analogy method provides the estimation results using the same ship motion responses. The estimation results from both methods are then fused together. Specifically, the wave buoy method results compensate the ML results based on its predictive uncertainty. This work will focus on the estimation of the significant wave height and the mean wave period. Real-world data are collected from the research vessel R/V Gunnerus. Currently, there is no wave radar installed on the ship and the sea state is manually observed by the captain based on experience. The proposed method aims to provide an onboard support tool for estimating the sea state and further support the development of autonomous vessels and operations. The main contributions can be highlighted as follows:

- A hybrid model is developed for sea state estimation using measured ship motion responses.
- The developed hybrid model has the ability to estimate a broad range of sea states when the training data is limited.
- The performance of the developed hybrid model is verified through real-world data collected from a research vessel.

The remainder of this paper is organized as follows: a introduction to sea state estimation is given in Section II. Section III introduces the proposed hybrid estimation approach. The experiments are discussed in Section IV. Section V concludes the paper.

II. RELATED WORK

Research has been conducted on estimating the sea state based on the motion response. Most of them focuses on the field of frequency domain analysis. Through Fast Fourier Transform (FFT) or autocorrelation analysis, the ship motion response is first transformed into the frequency domain. The RAOs are then used to relate the wave spectrum to the motion spectrum. The fundamental idea is to minimize the difference between the measured ship spectrum and the calculated ship spectrum [15]. If a wave spectra, e.g., JONSWAP, Bretschneider, is assumed, the wave parameters are obtained in the nonlinear optimization process [16]. Otherwise, a Bayesian approach can be applied [17], in which the wave spectrum is represented in a discrete frequency-directional domain and the original least square problem is transformed into the maximization of posterior. The methods are initially developed for dynamically positioned (DP) vessels, Iseki and Ohtsu [17] extend this method to ship with forward speed by incoperating the Doppler shift function.

The above methods depend on the spectral analysis, which may cause a certain degree of errors, the estimation of sea state based on ship motion response can also be solved in the time domain. Pascoal and Soares [18] treated the wave components as state variables and proposed an estimation algorithm based on the Kalman filter. This method is further extended to account for ships with forward speed and validation is performed through sea trials [19]. A similar observer-based approach is developed by Belleter et al. [20] to estimate the wave frequency using the measured roll or pitch angle. Nevertheless, these two methods, either frequency or time domain, are dependent on RAOs to relate the wave to the ship motion. RAOs are simplified linear transfer functions and hard to tune for a broad use (tuning with real-world data is often needed). In addition, RAOs only hold for mild and moderate sea state [9].

Machine learning methods are alternative methods that learn the mapping between ship motions and sea states directly. The advantage of these methods is that they do not rely on an explicit model to link waves to ship motions. Tu et al. [21] extract time and frequency domain statistical information of the measured motion data and apply a three-layer classifier to classify the sea state. Han et al. [10] extract statistical, temporal, spectral and wavelet features from ship motion responses. An ensemble machine learning model is then developed to estimate the sea state. A concern is how to extract useful features. An end-to-end deep learning method has also been developed. Cheng et al. [22] treat it as a time series classification problem and combine convolutional neural network (CNN), Long Short-Term-Memory (LSTM), and FFT to classify the sea state. Further, they develop a CNN with skip-connection and demonstrate its superior performance [23]. Mak and Duz [24] regard it as a regression problem and compared the performance of three network architectures: CNN, LSTM-CNN, and sliding puzzle. However, the collection of a large dataset that covers possible sea states is the foundation for these approaches, which is usually hard to archive and the model might fail catastrophically when the encountered sample is not in the training set.

III. HYBRID SEA STATE ESTIMATION

Since the proposed method in this study is a combination of model-based and data-based methods, the following will outline how they are constructed and how they cooperate to estimate the sea state from measured ship motions.

A. Overview of the hybrid approach

The machine learning model is only good at interpolation but generally cannot extrapolate well. The data collected for sea state estimation purposes is usually limited and can not cover the entire range of possible sea state conditions. Therefore, a model-based method is used to compensate for the ML results when the sample is out of training distribution. Fig. 1 shows the schematic illustration of the proposed hybrid

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2021.3073462, IEEE Transactions on Industrial Informatics

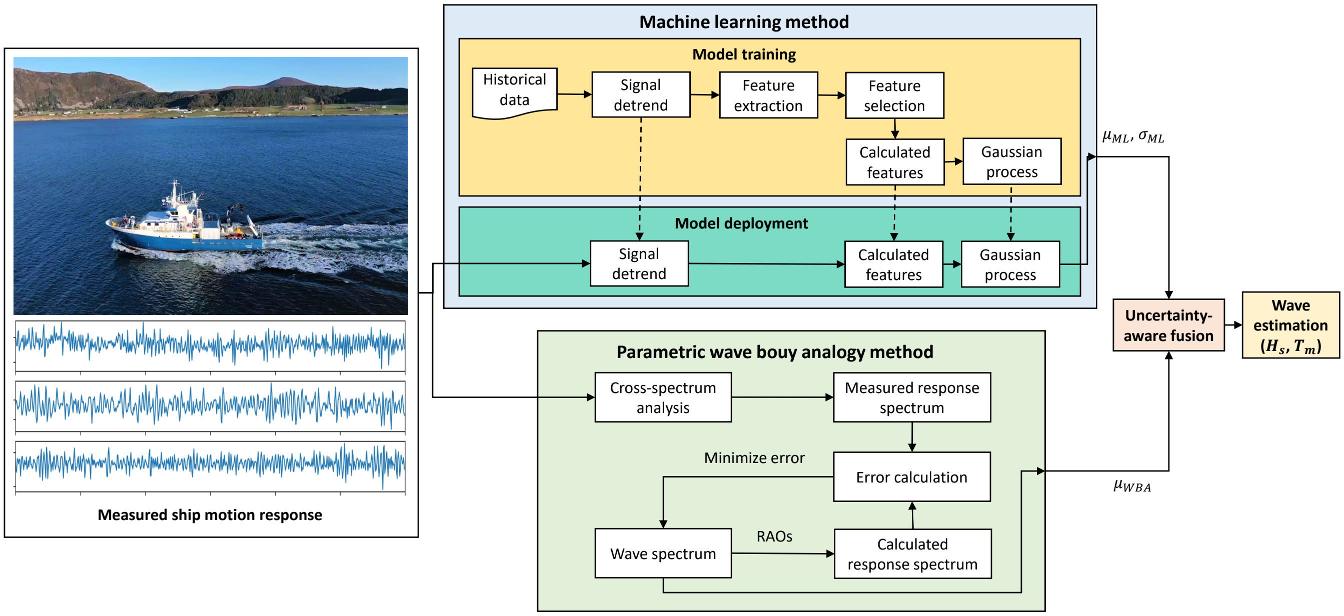

Fig. 1. Schematic illustration of the proposed hybrid approach. The upper rectangle is the ML model and the lower rectangle is the model-based method.

method. Historical data containing ship motion and corresponding sea state information is collected to train a machine learning model. The machine learning pipeline consists of feature extraction, feature selection, and model training. The Gaussian process is chosen since it not only provides predictions but also uncertainty. The wave buoy analogy method builds on a comparison between measurements of response spectrum and calculated ones. By minimizing the discrepancy between the measured and calculated spectrum, the sea state is determined. Then the uncertainty-aware confusion module receives the sea state estimation results from these two methods. The WBA estimation results are used to compensate for the ML results according to its uncertainty. In this way, the hybrid estimation results are the combination of the estimation results made by the ML model and the WBA method.

The detailed data-driven method, model-based method and the fusion of both methods will be illustrated in the following sections.

B. Data-driven sea state estimation

The machine learning model is established based on the procedure described in Han et al. [10].

1) Signal detrending: The measured ship motion might be affected by the measurement offset. In order to ensure that the ship motions fluctuates around zero, the average value of measured signal is subtracted. This step is important for the robustness of the extracted features.

2) Feature extraction: Considering a signal is a discrete time series data $(x_1, x_2, \dots, x_n)$ with length n, four categories of features are constructed to describe the sea state pattern, namely, statistical, temporal, spectral, and wavelet features.

Statistical features Seven basic statistical features are extracted from each DOF measurement. Six standard features of the signal including *maximum*, *minimum*, *mean*, *variance*, *skew*, and *kurtosis* are considered. Additionally, the q *quantile* information of the signal is extracted, which is the value greater than q of the ordered values from the signal. The variable q is selected as 0.2, 0.4, 0.6, and 0.8.

Temporal features Firstly five temporal features are considered, which include: *absolute sum of change* ($\sum_{i=1}^{n-1} |x_{i+1} - x_i|$), *absolute energy* ($\sum_{i=1}^{n} x_i^2$), *mean second derivative center* ($\frac{1}{2(n-2)} \sum_{i=1}^{n-2} \frac{1}{2}(x_{i+2} - 2x_{i+1} + x_i)$), *zero cross* (the number of the signal crossing zero), *longest strike above mean* (the length of the longest consecutive subsequence in a signal that is larger than its mean). *Autocorrelation*: This feature measures the similarity between observations as a function of the time lag between them. For a discrete process, the autocorrelation is obtained as $\frac{1}{(n-k)\sigma^2} \sum_{i=1}^{n-k} (x_i - \mu)(x_{i+k} - \mu)$, where μ and σ^2 are the mean and variance respectively. k denotes the time lag. Five different time lags (10, 20, 30, 40, 50) are used to extract this feature.

Welch spectral features The Welch method is an approach of converting a signal from the time domain to the frequency domain and estimating the power of a signal at different frequencies. The method is based on the fast Fourier transform (FFT) and the Hamming window. After the signal is transformed into the frequency domain, four basic spectral features including *max power spectrum*, *fundamental frequency*, *max frequency*, and *median frequency* are extracted. Additionally, five features related to the shape of the spectrum [25] is also extracted: *centroid*, *variation*, *spread*, *skewness*, *kurtosis*.

Wavelet features The wavelet transform is a time-frequency analysis method which selects the appropriate frequency band adaptively based on the characteristics of the signal. A signal can be split into different frequency sub-bands and therefore the signal can be analyzed with multi-scales in the time and frequency domain. The Daubechies wavelet of order 1 (db1) is selected as the basis function and the decomposition level

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2021.3073462, IEEE Transactions on Industrial Informatics

is five, which results in five approximation components and five detail components in total. For each components, the *mean*, *variance*, *median*, *skewness*, *kurtosis*, *absolute energy*, *absolute sum of changes*, and *zero cross* are extracted.

3) Feature selection: In order to select salient features from the constructed multi-domain features, mRMR [26] feature selection framework is utilized. The mRMR criterion is a filter-based feature selection method which can effectively reduce the redundant features while keeping the relevant features for the model. The mRMR criterion can be expressed as:

$$f_{mRMR}(x_i) = I(y, x_i) - \frac{1}{|S|}\sum_{x\in S} I(x_s, x_i) \tag{1}$$

where the function $I(\cdot,\cdot)$ denotes the mutual information (MI). $|S|$ is the size of the feature set and $x_s \in S$ is one feature out of the feature set. The first term in Eq.(1) represents the relevant to the target y while the second term measures the redundancy. Since the MI is computationally expensive for continuous variables, the redundancy is replaced with correlation. The MI used to measure the relevance is normalized to $[0, 1]$ to have a same range as the correlation.

4) Gaussian process regression: The data-driven predictive model is built based on the Gaussian process (GP) model [13], [27]. A Gaussian Process is a probability distribution over functions. The advantage of GP is that it provides a well-calibrated uncertainty of the prediction. We assume either exact or independent normally distributed measurement errors, i.e. the evaluation of $y(x)$ at point x satisfies:

$$y(x)|f(x) \sim \mathcal{N}(\mu(x), \sigma^2(x)) \tag{2}$$

where σ^2 is a known function describing the variance of the measurement errors and $\mu(x)$ is the mean.

GP is characterized by a mean function $m(x)$ and a covariance kernel function $\kappa(x, x')$. Given the training set at n points with input as $x_{1:n} \triangleq \{x_1, x_2, \cdots, x_n\}$ and target as $y_{1:n} \triangleq \{y_1, y_2, \cdots, y_n\}$, the posterior can be obtained by combining these observed values with prior:

$$\begin{aligned}
&\mu(x) = m(x)\\
&+ \kappa(x, x_{1:n})[\kappa(x_{1:n}, x_{1:n}) + \sigma_n^2 I]^{-1}(y_{1:n} - m(x_{1:n}))\\
&\sigma^2(x) = \kappa(x, x)\\
&- \kappa(x, x_{1:n})[\kappa(x_{1:n}, x_{1:n}) + \sigma_n^2 I]^{-1}\kappa(x_{1:n}, x)
\end{aligned} \tag{3}$$

where σ_n^2 is a additive noise level. The $\mu(x)$ can be viewed as the prediction of the function value, while the σ^2 is a measure of uncertainty of the prediction. In this work a constant mean function $m(x) = 0$ is used and the rational quadratic kernel is used:

$$\kappa(x, x') = \left(1 + \frac{(x - x')^2}{2\alpha l^2}\right)^{-\alpha} \tag{4}$$

where α and l are parameters of the kernel. These parameters are obtained by maximizing the log marginal likelihood.

C. Parametric wave buoy analogy method

Assuming linearity between waves and ship response, the cross-spectrum of ship responses are related to the direction wave spectrum through the following integral:

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

where $\Phi(\omega_e, \theta)$ denotes the response amplitude operators (RAOs) in terms of a complex-valued transfer function and $\overline{\Phi(\omega_e, \theta)}$ is the complex conjugate. $E(\omega_e, \theta)$ is the directional wave spectrum, ω_e and θ are the encounter wave frequency and the relative wave direction, respectively. It is noteworthy that the wave spectrum is advantageously estimated in the wave frequency domain. The encounter frequency ω_e is related to the absolute frequency ω through the Doppler shift:

$$\omega_e = \omega - \omega^2\psi, \quad \psi = \frac{v}{g}\cos\theta \tag{6}$$

where g is the acceleration of gravity and v is the forward speed of the vessel.

The parametric directional wave spectrum is usually based on a 10-parameter bi-model spectrum. Since the shape parameter λ has a weak influence on wave-induced loads and ship motion [16], its value has been fixed as 1. Therefore the wave spectrum is given by:

$$\begin{aligned}
E(\omega, \theta) =& \frac{1}{4}\sum_{i=1}^{2}\frac{5}{4}\omega_{mi}^4\frac{H_{si}^2}{\omega^5}\exp{[-\frac{5}{4}(\frac{\omega_{mi}}{\omega})^4]}\\
&\times A(s_i)\cos^{2s_i}(\frac{\theta - \theta_{mi}}{2})
\end{aligned} \tag{7}$$

where H_s is the significant wave height, θ_m is the mean wave direction and ω_m is the model frequency. The spectrum in Eq.(7) can be referred to as a Pierson-Moskowitz spectrum with the $\cos^{2s}$ spreading model. Since the model considers two separated wave components ($i = 1, 2$), it is capable of representing a variety of spectrum shapes. The constant $A(s)$ in the $\cos^{2s}$ model is difined as:

$$A(s) = \frac{2^{2s-1}\Gamma^2(s+1)}{\pi\Gamma(2s+1)} \tag{8}$$

where Γ demotes the Gamma function and s is the spreading parameter.

The estimation problem can be established through Eq.(5), where the left-hand side is estimated by measured ship motion response and the right-hand side is obtained through theoretical calculations. By minimizing the difference between the two sides in Eq.(5), the sea state parameters can be obtained. In this way, a minimization problem is formulated through the following objective function:

$$\min_x \quad \sum_{i=1}^{n}\sum_{j=1}^{n}(S_{ij} - \hat{S_{ij}}(x))^2 \tag{9}$$

where S_{ij} is the cross spectrum from measured ship motion responses and $\hat{S_{ij}}$ is the cross spectrum from theoretical calculation with wave parameters x. The wave parameters in this paper is representing by a 8-component vector $x =$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2021.3073462, IEEE Transactions on Industrial Informatics

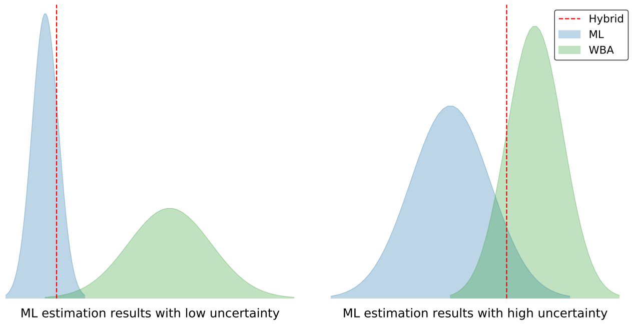

Fig. 2. Illustration of the uncertainty-aware fusion.

$[H_{s1}, \omega_{m1}, \theta_{m1}, s_1, H_{s2}, \omega_{m2}, \theta_{m2}, s_2]$. n is the number of used ship motion components and it is set as 3 in this paper since only the sway velocity, roll, heave are used.

This leads to a non-linear optimization problem. We randomly sample 20 initial points from the wave parameters space and then the L-BFGS-B algorithm is used, the estimated wave parameters are selected as the one with the lowest value in Eq.(9). In this way, a near-optimal result is achieved.

When the 8-component wave parameters are determined, the 2D directional wave spectrum is obtained. Then the estimated significant wave height $\hat{H}_s$ and the mean wave period $\hat{T}_m$ can be calculated as follows:

$$\begin{aligned} \hat{H}_s &= 4\sqrt{m_0} \\ \hat{T}_m &= m_{-1}/m_0 \end{aligned} \tag{10}$$

where m_0 and m_{-1} represents the moment of wave with order 0 and -1, respectively. Specifically, $m_n = \iint \omega^n E(\omega, \theta) d\omega d\theta$ with order n.

D. Uncertainty-aware fusion

As shown in Fig. 2, the estimation results from the machine learning model and the wave buoy analogy method are assumed to follow a distribution as $P(y|ML)$ and $P(y|WBA)$, respectively. Since $P(y|ML)$ and $P(y|WBA)$ are independent, the final result can be obtained through eq.(11). In this way, the hybrid estimation results would move towards the WBA results if the uncertainty of the ML results are high.

$$P(y|ML, WBA) = P(y|ML) \cdot P(y|WBA) \tag{11}$$

The $P(y|ML)$ follows a Gaussian distribution with mean μ_{ML} and variance σ_{ML}^2, which can be calculated by eq.(3). For the wave buoy analogy method, the uncertainty is not easy to measure directly, a Gaussian distribution is also assumed for $P(y|WBA)$ with mean μ_{WBA} (calculated by eq.(10)) and variance σ_{WBA}^2. Then the final estimation result is:

$$y_{ML,WBA} = \mu_{ML} + \frac{\sigma_{ML}^2(\mu_{WBA} - \mu_{ML})}{\sigma_{ML}^2 + \sigma_{WBA}^2} \tag{12}$$

Here σ_{WBA}^2 is a parameter which can be tuned to adjust the final results towards ML or WBA results. The computational complexity of the hybrid approach is $\mathcal{O}(n^3 + m^2)$, where n is the number of samples used in the ML model and m is the number of undetermined parameters in the WBA method.

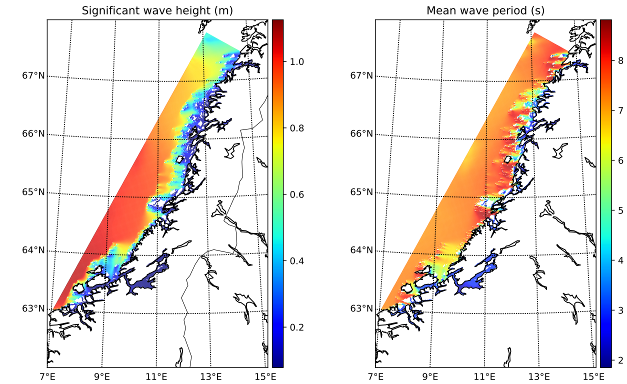

Fig. 3. Sea State information in the middle Norway at 12:00, 13th, June, 2018 reported by the Norwegian Meteorological Institute.

IV. EXPERIMENT

A. Data

The experiment was conducted based on historical data acquired through log files created by a data acquisition system onboard the RV Gunnerus. The one-year time period from June 2017 and ending in October 2018 was selected. For all measurements in the data set, a sampling rate of 1 Hz was observed.

The maneuvering data that the vessel is cruising with a constant speed and constant heading is obtained, which results in a total of 47 trajectories. The cruising speed of the vessel is about 10 knots. The trajectories are then cut into 20 minutes segment without overlapping since the sea state usually remains unchanged for 20 minutes. Three sensor measurements related to the vessel motion were obtained: *sway velocity*, *roll*, and *heave*. These measurements are responsible for estimating the sea state. Two additional variables *longitude* and *latitude* are obtained, which is for matching the target sea state into the motion responses. Table. I gives all the input variables used in this study. Ranges are given as maximum and minimum values observed in the time series of each variable.

TABLE I
SHIP MOTION RESPONSES USED IN THIS STUDY AS INPUT

Variable name	Range	Unit
Sway velocity	[-2.64, 3.23]	knots
Roll	[-13.00, 12.01]	deg
Heave	[-2.03, 2.13]	m

The sea state information is collected from the weather forecast system provided by the Norwegian Meteorological Institute (MET). Since the vessel is only operating in the west coastal region of Norway, the coastal data is used. The coastal wave data is obtained by a numerical wave model which is run on an 800-meter grid with ECMWF and AROME atmospheric force. Two sea state characteristics are considered: Significant wave height H_s and mean wave period T_m. Fig. 3 shows the contour plot of the significant wave height in the coastal region of middle Norway on a specific day. The two sea state characteristics are then matched to the ship motion data

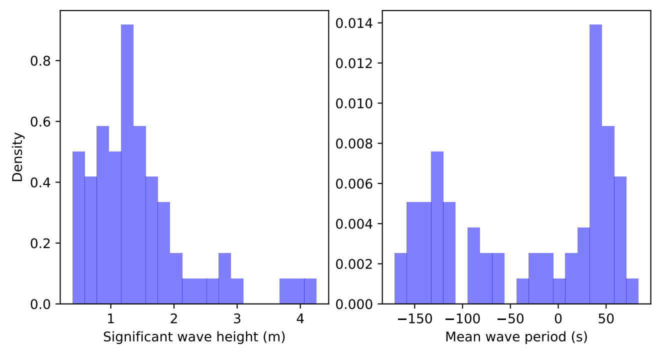

Fig. 4. Distribution of the collected sea state characteristics.

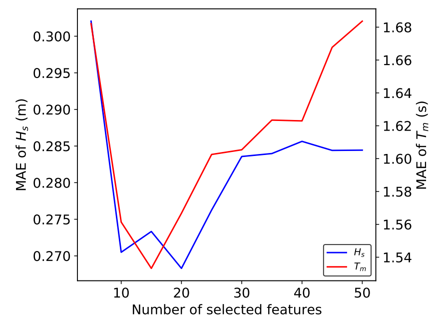

Fig. 6. Number of selected features VS. MAE of significant wave height and mean wave period.

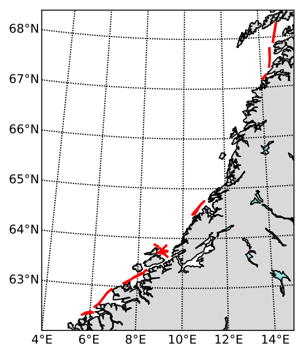

Fig. 5. Illustration of the data we collected from R/V Gunnerus operating on the west coast of Norway. The red lines denote the trajectory of the vessel.

through position information. Specifically, the *longitude* and *latitude* corresponded to the ship motion data are used to query the nearest sea state information. The process is done by utilizing a ball tree with the Haversine distance.

Fig. 4 shows the sea state distribution of the collected data. It is shown that the significant wave height is mostly distributed around 1m. The reason is that the vessel is usually operated near the west coast of Norway and it is not likely to go far away from the shore, as shown in Fig. 5.

B. Evaluation Metrics

As presented in Section IV-A, the ship motion data is cut into segments of 20 minutes. For the ML method, the segments are divided into 5 subsets without shuffling. In this way, the segments that come from the same trajectory would not end up in different folds to prevent data leakage. Among the 5 subsets, a single subset is retained as the validation data, and the remaining 4 subsets are used as training data. The process is then repeated 5 times and the out-of-fold predictions are used. In parallel, the WBA method is utilized to provide the same kinds of predictions. The RAOs used in the WBA method is calculated through a hydrodynamic workbench ShipX. The hybrid predictions are given by combining the out-of-fold predictions from the data-driven model and the predictions from the parametric wave buoy analogy method. To evaluate the performance of the methods, the mean absolute error (MAE) is used. The MAE is calculated as follows:

$$y_{err} = \sum_{i=1}^{N} |y_i^{pred} - y_i^{actual}| \tag{13}$$

where N is the number of sample. y^{pred} and y^{actual} denotes the predicted and actual value, respectively.

C. Machine Learning Model Development

The development of the machine learning model consists of feature extraction, feature selection, and model training. As described in Section III, four different kinds of features are extracted and only the salient features are selected. Once the salient features are determined, these features can be constructed and used in the Gaussian process model in the deployment stage. The training takes around 3 seconds using the Intel Xeon W-2225 CPU. Fig. 6 shows the mean absolute error (MAE) versus the number of used features for the Gaussian process model. The features are ranked by the mRMR criterion. The blue line indicates the significant wave height while the red line indicates the mean wave period.

It is shown that the performance of the model first increases with the number of features and then the performance degrades. The reason is that some of the features are similar and therefore a certain degree of feature redundancy exists. When the selected features exceed a certain value, the MAE of the model starts to increase. The optimal number of features uses for H_s and T_p is 15 and 20, respectively.

In order to understand what kind of features are used on the developed model. The features used in our ML model are shown in Fig. 7 and Fig. 8 for significant wave height and mean wave period, together with the corresponding score from mRMR criterion. The motion is in the first bracket while the feature extracted from this motion is in the second bracket. The approximation and detail component from wavelet transform is denoted as “approx” and “detail”, respectively. It is shown that for the significant wave height, the features related to the amplitude or the strength of the signal is favored. As for the mean wave period, the focus is given to the spectral and wavelet features. The selected features fit our intuition since the wave height is related to the magnitude of ship displacement, and the response spectrum shape and the signal in different frequency range is sensitive to the wave period. The development of the ML model is finished.

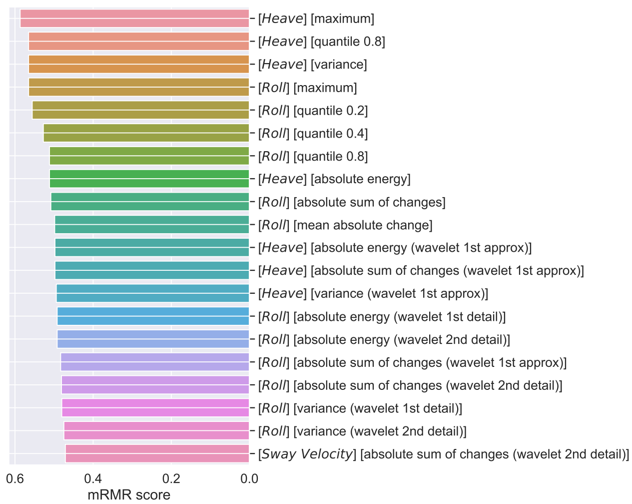

Fig. 7. Features ranked by mRMR criterion for significant wave height.

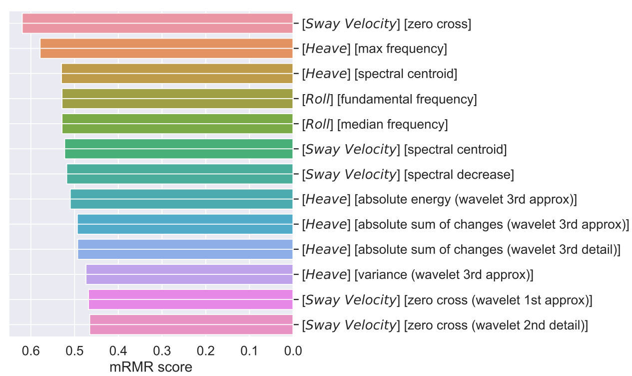

Fig. 8. Features ranked by mRMR criterion for mean wave direction.

D. Effect of σ^2_{WBA}

To develop the hybrid model, the uncertainty of the WBA method is required. Since the uncertainty of the model-based WBA method can not be directly represented, a constant parameter σ^2_{WBA} is then introduced to express the uncertainty. Generally, larger σ^2_{WBA} suggests that we have less confidence in the WBA estimation and vice versa. Fig. 9 and Fig. 10 show the MAE versus σ^2_{WBA} in terms of significant wave height and mean wave period, respectively.

In these two figures, the MAE of the hybrid method first drops and then steadily increase with the increase of σ^2_{WBA}. The MAE of the hybrid method is similar to the WBA method when σ^2_{WBA} is small and it is similar to ML predictions when σ^2_{WBA} is large. The MAE of the hybrid method can be lower than the ML method when σ^2_{WBA} exceeds a certain value. This phenomenon is more obvious for the significant wave height as shown in Fig. 9. The hybrid method only reduces the MAE for the mean wave period in a small range. The reason is that the mean wave period estimated by the WBA method has similar or even higher errors when comparing with the results with high uncertainty from the ML method. From the sensitivity analysis, the optimal values for σ^2_{WBA} for the

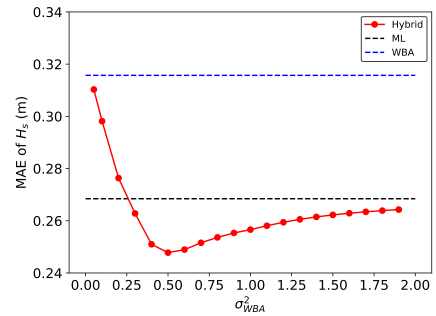

Fig. 9. Effect of σ^2_{WBA} on the estimation error of the significant wave height.

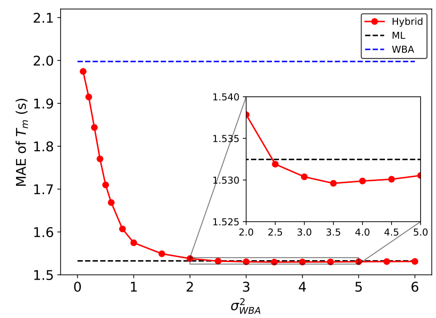

Fig. 10. Effect of σ^2_{WBA} on the estimation error of the mean wave period.

significant wave height and the mean wave period are around 0.5 and 3.5 in this study. These values yield the lowest error for the hybrid method.

E. Performance Evaluation

In this part, the performance of different methods is evaluated. A baseline model named SeaStateNet [22] is implemented here for comparison. SeaStateNet is an end-to-end deep learning model that directly uses the raw sensor as input. In order to distinguish between this ML model with our ML model, SeaStateNet and GP are used as the notation in this part.

Fig. 11 shows the significant wave height for each sample, where MET stands for the "actual" value from the Norwegian Meteorological Institute. Fig. 12 presents the same graph for the mean wave period. The value of σ^2_{WBA} are selected as 0.5 and 3.5 for significant wave height and mean wave period, respectively (see Section IV-D). The GP model provides fairly accurate results in terms of the significant wave height. For the mean wave period, the predictions are mostly distributed in the range of $5s$ to $8s$, therefore it provides relatively bad results for low and high wave periods. Similar results are observed for the WBA method and the SeaStateNet model. The reason might be that the vessel itself is a filter and its motions are only sensitive in a specific range of the wave frequency. The hybrid model predictions are the GP model predictions corrected by the WBA method. As shown in Fig. 11, the GP model predictions with high uncertainty are corrected, which can be easily observed for samples 17, 51, and 52. The GP predictions and hybrid predictions in Fig. 12 is quite similar since we put

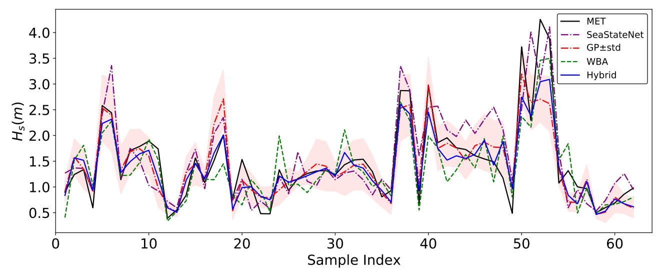

Fig. 11. Estimation of significant wave height by different approaches.

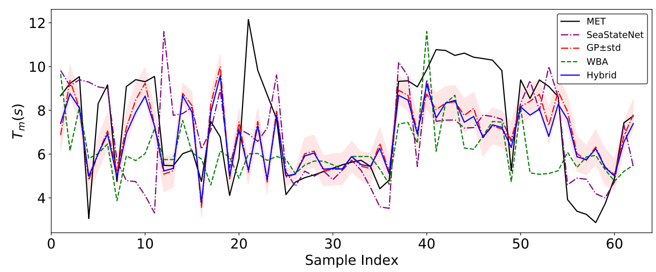

Fig. 12. Estimation of mean wave period by different approaches.

a relatively large σ^2_{WBA}. The reason is that the results from WBA for the mean wave period are relatively less accurate compared with the significant wave height.

Table. II summarized the overall performance in terms of MAE. The GP model performs better than the SeaStateNet model. The reason might be that our data is limited. It is shown that the GP predictions provide an overall low error when comparing with the WBA method. The hybrid method reduces the MAE in terms of significant wave height by about 10% when comparing with the GP method. For the mean wave period, the hybrid method gives a similar error with the GP method. From the experiment, the hybrid method can reduce the estimation errors by correcting the high uncertainty GP predictions with the WBA predictions. Compared with the rest of the models, the hybrid model has the smallest error.

TABLE II
MAE OF DIFFERENT SEA STATE ESTIMATION METHOD

Sea State	SeaStateNet	WBA	GP	Hybrid (GP+WBA)
$H_s(m)$	0.392	0.316	0.268	0.248
$T_m(s)$	1.758	1.998	1.533	1.529

F. Discussion

The proposed hybrid method consists of a data-driven method and a model-based method. The error of the model-based wave buoy analogy method comes from the following aspects: (1) the assumption of parametric wave spectrum; (2) the errors from spectral analysis; (3) the uncertainty of the transfer function; (4) the nonlinear optimization procedure. Even though the data-driven approach does not subject to the limitation above, it is prone to be failed when the new sample is not from the same distribution as the training data. The wave buoy method is used to compensate for the results from the data-driven method when the sample is unlikely from the training data, which is represented by the outputs with uncertainty from the data-driven method. The uncertainty should accurately characterize the confidence of the results. Therefore the success of the proposed method relies on the accuracy of both methods. The upper bound error of the hybrid method is the method with a higher error, which is the WBA method in this case. To summarize, the proposed hybrid method tries to eliminate the disadvantage of the data-driven method with the model-based method. However, the correctness of the uncertainty representation and the accuracy of the model-based method are two key aspects for this approach.

V. CONCLUSIONS

Estimating the sea state from measured ship motion response is a complex and challenging task. As a way to reduce the possibility of failure in the ML model when the encountered sea state is not in the training set, estimation results from the ML model were combined with the results from the model-based wave buoy analogy method. This results in a hybrid estimation approach. In the ML model, the Gaussian process is used, which allows obtaining not only the estimation results but also the uncertainty of the estimation results. When the uncertainty of the ML model results can be obtained, the WBA results are used to compensate for the ML results based on its uncertainty. Specifically, the more uncertainty present in the ML model, the more the final results will be relying on the WBA method. This is accomplished by the proposed fusion module. A substantial decrease in the mean absolute error was observed for the significant wave height, with a reduction of error of nearly 10%. For the mean wave period, the hybrid approach shows a similar performance compared to the pure ML model in this case.

This study suggests that the proposed hybrid method offers better performance compared with the pure ML or the pure WBA method. The major drawback of this approach is that if the model-based method is inaccurate and the trust in this method is high. Also, the parameter σ^2_{WBA} needs to be determined by expertise or trial and error. Since the ML model is expected to get better if more data is available, the proposed hybrid model could be a transition from a pure model-based method to a pure data-driven method. Future research will focus on developing a machine learning model to estimate the 2D wave spectrum instead of wave characteristics. In addition, incorporating the estimation method into control or path planning of marine vehicle will be investigated.

REFERENCES

[1] G. Kavallieratos, V. Diamantopoulou, and S. Katsikas, "Shipping 4.0: Security requirements for the cyber-enabled ship," *IEEE Transactions on Industrial Informatics*, 2020.

[2] R. Jalonen, R. Tuominen, and M. Wahlström, "Remote and autonomous ships—the next steps: Safety and security in autonomous shipping—challenges for research and development," *Rolls-Royce, Buckingham Gate, London: The Advanced Autonomous Waterborne Applications (AAWA)*, pp. 56–73, 2016.

[3] Y. Wang, B. Jiang, Z.-G. Wu, S. Xie, and Y. Peng, "Adaptive sliding mode fault-tolerant fuzzy tracking control with application to unmanned marine vehicles," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020.

[4] G. Li, H. P. Hildre, and H. Zhang, "Toward time-optimal trajectory planning for autonomous ship maneuvering in close-range encounters," *IEEE Journal of Oceanic Engineering*, vol. 45, no. 4, pp. 1219–1234, 2019.

[5] Z. Peng, J. Wang, D. Wang, and Q.-L. Han, "An overview of recent advances in coordinated control of multiple autonomous surface vehicles," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 732–745, 2020.

[6] T. I. Fossen, *Handbook of marine craft hydrodynamics and motion control*. John Wiley & Sons, 2011.

[7] D. C. Stredulinsky and E. M. Thornhill, "Ship motion and wave radar data fusion for shipboard wave measurement," *Journal of ship research*, vol. 55, no. 2, pp. 73–85, 2011.

[8] A. H. Brodtkorb, U. D. Nielsen, and A. J. Sørensen, "Sea state estimation using vessel response in dynamic positioning," *Applied Ocean Research*, vol. 70, pp. 76–86, 2018.

[9] U. D. Nielsen, "Estimation of directional wave spectra from measured ship responses," in *12th International Congress of the International Maritime Association of the Mediterranean: Maritime Transportation and Exploitation of Ocean and Coastal Resources*, 2005, pp. 1103–1112.

[10] P. Han, G. Li, S. Skjong, B. Wu, and H. Zhang, "Data-driven sea state estimation for vessels using multi-domain featuresfrom motion responses," accepted by International Conference on Robotics and Automation (ICRA) 2021.

[11] O. Bousquet, S. Boucheron, and G. Lugosi, "Introduction to statistical learning theory," in *Summer School on Machine Learning*. Springer, 2003, pp. 169–207.

[12] J. Ren, P. J. Liu, E. Fertig, J. Snoek, R. Poplin, M. Depristo, J. Dillon, and B. Lakshminarayanan, "Likelihood ratios for out-of-distribution detection," in *Advances in Neural Information Processing Systems*, 2019, pp. 14 707–14 718.

[13] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

[14] K. Lee, K. Lee, H. Lee, and J. Shin, "A simple unified framework for detecting out-of-distribution samples and adversarial attacks," in *Advances in Neural Information Processing Systems*, 2018, pp. 7167–7177.

[15] U. D. Nielsen, "Estimations of on-site directional wave spectra from measured ship responses," *Marine Structures*, vol. 19, no. 1, pp. 33–69, 2006.

[16] E. A. Tannuri, J. V. Sparano, A. N. Simos, and J. J. Da Cruz, "Estimating directional wave spectrum based on stationary ship motion measurements," *Applied Ocean Research*, vol. 25, no. 5, pp. 243–261, 2003.

[17] T. Iseki and K. Ohtsu, "Bayesian estimation of directional wave spectra based on ship motions," *Control Engineering Practice*, vol. 8, no. 2, pp. 215–219, 2000.

[18] R. Pascoal and C. G. Soares, "Kalman filtering of vessel motions for ocean wave directional spectrum estimation," *Ocean Engineering*, vol. 36, no. 6-7, pp. 477–488, 2009.

[19] R. Pascoal, L. P. Perera, and C. G. Soares, "Estimation of directional sea spectra from ship motions in sea trials," *Ocean Engineering*, vol. 132, pp. 126–137, 2017.

[20] D. J. Belleter, R. Galeazzi, and T. I. Fossen, "Experimental verification of a global exponential stable nonlinear wave encounter frequency estimator," *Ocean Engineering*, vol. 97, pp. 48–56, 2015.

[21] F. Tu, S. S. Ge, Y. S. Choo, and C. C. Hang, "Sea state identification based on vessel motion response learning via multi-layer classifiers," *Ocean Engineering*, vol. 147, pp. 318–332, 2018.

[22] X. Cheng, G. Li, R. Skulstad, S. Chen, H. P. Hildre, and H. Zhang, "Modeling and analysis of motion data from dynamically positioned vessels for sea state estimation," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 6644–6650.

[23] X. Cheng, G. Li, A. L. Ellefsen, S. Chen, H. P. Hildre, and H. Zhang, "A novel densely connected convolutional neural network for sea state estimation using ship motion data," *IEEE Transactions on Instrumentation and Measurement*, 2020.

[24] B. Mak and B. Düz, "Ship as a wave buoy: Estimating relative wave direction from in-service ship motion measurements using machine learning," in *International Conference on Offshore Mechanics and Arctic Engineering*, vol. 58882. American Society of Mechanical Engineers, 2019, p. V009T13A043.

[25] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The timbre toolbox: Extracting audio descriptors from musical signals," *The Journal of the Acoustical Society of America*, vol. 130, no. 5, pp. 2902–2916, 2011.

[26] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

[27] R. R. Richardson, C. R. Birkl, M. A. Osborne, and D. A. Howey, "Gaussian process regression for in situ capacity estimation of lithium-ion batteries," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 127–138, 2018.