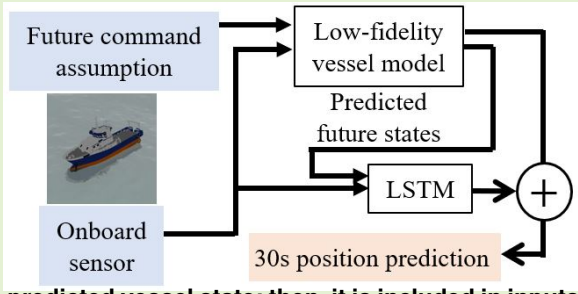
A multiple-output hybrid ship trajectory predictor with consideration for future command assumption

Motoyasu Kanazawa, Robert Skulstad, Guoyuan Li *Senior Member, IEEE,*, Lars I.Hatledal, Houxiang Zhang *Senior Member, IEEE,*

***Abstract*— Onboard sensors contribute to data-driven understanding of complex and nonlinear ship dynamics in real time. By using sensors, precise ship trajectory prediction plays a key role in intelligent collision avoidance. A hybrid predictor makes prediction based on a mathematical model of which error is compensated by a black-box model. A Multiple-output Hybrid Predictor (MHP), which makes a long-horizon prediction at a time based on onboard sensor data, was developed in the previous study. However, it can not handle a time series of future command assumption. This limitation hinders an MHP from being applied to the evaluation of future command assumption in the predictive decision making. A novel architecture of MHP presented in this study converts a long time series of future command assumption into a fixed-length model-based-predicted vessel state; then, it is included in inputs of a black-box error compensator. This idea is robust to multidimensionality of commands and long control horizon. Assuming a low-fidelity vessel model and a limited data are available, simulation experiments are conducted. The effect of environmental disturbances and maneuverings on the prediction performance is examined comprehensively for the first time. The present study successfully incorporates a long time series of future command assumption in an MHP. It reduces the mean error by 81.8% compared to a model-based predictor and by 45.6% compared to a data-driven predictor under Beaufort wind force scale 4 wave, wind, and ocean current. Present study expands the application of MHP to the predictive decision making of future autonomous ships.**



***Index Terms*— collision avoidance, onboard sensors, recurrent neural network, ship trajectory prediction**

I. Introduction

DATA-DRIVEN technologies by exploiting onboard ship sensor data have gathered significant momentum [1] recently. In particular, the basis of intelligent marine transportation is better understanding of highly nonlinear and complex ship dynamics. The intelligent systems lab in NTNU Ålesund is currently working on the development of intelligent autonomous ships through the establishment of "Remote Control Centre for Autonomous Ship Support: AuReCo". In AuReCo, real sensor data are shared with real vessel, digital model of the vessel, and the remote control centre. By collecting real sensor data from real ship operation, we aim at the integration of data-driven measures and domain knowledge, that is referred to as the hybrid modeling. In particular, by integrating two measures in the hybrid modeling, the accurate trajectory prediction of an own ship (OS) forms a basis of predictive decision making of autonomous ships. Model Predictive Control (MPC), that has been widely used in this domain [2], [3], makes trajectory predictions by using a model based on the assumption of a time series of commands $\mathbf{U} = [U(t|t), ..., U(t+k|t)..., U(t+N|t)]$ where N is the number of steps of the control horizon and $U(t+k|t)$ is the assumption of command at the k-step future when the prediction is made at step t. Optimization algorithms of MPC optimize $\mathbf{U}$ so that pre-defined cost function is minimized and only $U(t|t)$ is applied to the plant. It should be noted that $U(t+k|t) = U(t+k|t+k)$ is not guaranteed since the controller updates $U(t|t)$ in real time. In this study, a Multiple-output Hybrid Predictor (MHP) that handles a time series $\mathbf{U}$ is developed with the aim of better predictive decision making of autonomous maneuvering.

Algorithms for trajectory predictors by using sensor data are grouped into model-based, data-driven, and hybrid predictors. Trajectory prediction made by a mathematical vessel model is referred to as a model-based prediction [4]. Its advantage is its physical interpretability. It enables us to develop a predictor efficiently based on physical understanding of vessel behavior through experiments. However, a model-based predictor has

Manuscript received Sep 06, 2021; revised Oct 06, 2021. This work was supported by a grant from NRF, IKTPLUSS project No.309323 "Remote Control Center for Autonomous Ship Support" in Norway.

M.Kanazawa, R.Skulstad, G.Li, and H.Zhang are with the Department of Ocean Operations and Civil Engineering, Faculty of Engineering, Norwegian University of Science and Technology (NTNU), 6009 Aalesund, Norway (e-mail: (motoyasu.kanazawa, robert.skulstad, guoyuan.li, and hozh) @ntnu.no).

L.I.Hatledal is with the Department of ICT and Natural Sciences, NTNU, 6009 Aalesund, Norway (e-mail: laht@ntnu.no).

three big problems by nature. First, uncertainties of a mathematical model exist to some degree. As a mathematical model is a simplification of highly nonlinear and complex phenomena in reality, errors can not be removed completely. Second, uncertain environmental disturbances, such as wave and current, are not able to be accurately incorporated into a mathematical vessel model due to the lack of accurate measurements since most vessels are not equipped with accurate wave radars. Even if it is equipped with them, accuracy of the calculation of wave forces in real time is yet to be fully developed [5]. Last, the development of a mathematical model can be prohibitively expensive.

Machine learning (ML) has an advantage in dealing with trends and pattern recognition by exploiting a large number of training data with uncertainty and non-linearity. As modern vessels are equipped with many sensors including accelerometers and GPS, predictors using ML technology have been gaining increasing attention in recent years [6]. They are referred to as a data-driven predictor. The data-driven predictor is supposed to work well if the given dataset is sufficient, of good quality, and well-distributed. However, this is an optimistic assumption in reality. The quality and distribution of real data fetched from maneuvering history are not tailored for ML training. The limited amount of dataset could be a hindrance when a predictor of ships with less maneuvering experience is developed.

In order to compensate disadvantages of model-based and data-driven predictors with each other and minimize their total cost & effort, Skulstad et al. [4] presented a hybrid predictor of a dynamic vessel model and a Neural Network (NN). Inputs to the NN are onboard sensor data and outputs compensate the error between model-based predicted and true trajectories. According to the definitions of [7], their hybrid predictor f_h is grouped into multiple-output strategy. It makes a multiple-step-ahead prediction at a time $[\phi_{t+H}, ...\phi_{t+1}] = f_h(\phi_t, ..., \phi_{t-d+1}, U(t+H-1|t), ..., U(t|t)) + w$ where t is time step when making prediction, d is the length of a time window, ϕ is a state of a system, H is the length of the prediction horizon, and w is a noise. Conversely, single-output strategy makes one-step-ahead prediction $\phi_{t+1} = f_h(\phi_t, ..., \phi_{t-d+1}, U(t|t)) + w$ and iterates the prediction H times to make H-step prediction. Its structure is simple, however, its performance for long-horizon prediction is not guaranteed due to error propagation over the prediction horizon [7]. Moreover, time-consuming iterative prediction would not be preferable if it is used for real-time decision making. Skulstad et al. [4] reported their MHP reduced mean prediction error made by a model-based predictor almost by half at 30s future. However, they assume a vessel keeps current command over the prediction horizon $U(t+H-1|t) = ... = U(t|t)$. This limitation hinders an MHP from being applied to the evaluation of the effect of future command assumption on trajectory prediction in the predictive decision making. This study enables an MHP to handle a time series of future command assumption. The biggest challenge is "curse of dimensionality" due to having multidimensional future command assumption over the control horizon in the NN. A novel architecture of an MHP presented in this study converts a long time series of future command assumption into a fixed-length predicted vessel state; then, it is included in inputs of a black-box error compensator. This idea is robust to multidimensionaity of commands and the long control horizon that have been a challenge of handling future command assumption in an MHP.

Furthermore, this study performs a comprehensive validation study. Real data employed by Skulstad et al. [4] are taken at ports isolated from strong environmental disturbances and wave measurements are not given in their dataset. Therefore, the effect of environmental disturbances on the prediction performance is yet to be investigated. This study utilizes virtual onboard sensor data under different environmental disturbances and maneuverings by using the co-simulation platform Vico [8]. The experiment in this study assumes that only a low-fidelity vessel model and a limited number of data are available. The virtual R/V Gunnerus, which is a Norwegian University of Science and Technology's research vessel, is employed. With consideration for future command assumption, a proposed hybrid predictor consists of a low-fidelity vessel model and an NN. Contributions of this work are summarised as follows.

- Present novel MHP architecture successfully handles a long time series of future command assumption. It expands the application of MHP to the predictive decision making of autonomous ships. In the simulation experiment, present study reduces the mean prediction error by 81.8% compared to the model-based predictor and by 45.6% compared to the data-driven predictor at 30s future under Beaufort wind force scale 4 level wind, wave, and ocean current.
- For the first time, this study reveals the effect of environmental disturbances and maneuverings on the prediction accuracy of the hybrid predictor through a comprehensive validation study.

This study unfolds as follows. In Section II, related works are presented. Section III introduces a new structure of an MHP that enables efficient consideration of future command assumption. Section IV describes a validation study. Conclusions are given in Section V.

II. Related works

A. Model-based prediction

Since a model-based predictor is based on physical interpretation of vessel behavior, its parameters are identified efficiently through the selection of a mathematical model, numerical calculations and experiments, and algorithms of parameter identifications. The most straightforward mathematical expression is a holonomic model of which assumption is that the vessel moves freely on the North-East plane without considering kinematic and kinetic constraints. It is applied to many collision avoidance algorithms [9]. Kinematic models are used in [10]. Aiming at better prediction accuracy, tremendous research effort has been dedicated to developing kinetic expressions of vessel dynamics that are referred to as dynamic models. They are grouped into the response model, the Abkowitz model, the Mathematical Modeling Group (MMG) model, and the vectorial representation model

[11]. The application of dynamic models in control tasks can be seen in [3]. Although the response model has a limited number of parameters, it is only applicable to simple course-keeping tasks as it focuses only on the yaw motion. The other dynamic models have a large number of parameters to be identified. Previous researches reported that those parameters have multicollinearity with each other. Therefore, parameter identification without losing generality is very challenging [12]. ML technology is used for efficient parameter identification of a mathematical vessel model [13]. Triantafyllou et al. [14] used Kalman Filter (KF) to estimate and predict ship states online based on a mathematical vessel model. Regardless of the type of a mathematical expression, increasing cost of parameter identification and preciseness of a model are in the relationship of the trade-off. Even though efficient parameter identification using ML technology is on-trend, discrepancy from reality is inevitable as complex nonlinear terms are ruled out from formulations through simplifications. In addition, it is a big problem that the effect of environmental disturbances is not able to be accurately incorporated into a model-based prediction due to (1) the lack of accurate real-time measurement of wave and current in most cases (2) its mathematical representation has discrepancy from reality [5].

B. Data-driven prediction

In contrast to a model-based predictor, a data-driven predictor attempts to surrogate vessel behavior without prior knowledge of its physical formulation by using black-box models. Thanks to the development of computational processing power, data-driven predictors have been gaining increasing attention. Previous research employed NNs [6], Support Vector Machines [15], Gaussian Process Regressions [16], and Auto-Regressive models [17]. The biggest advantage of a data-driven predictor is that it is able to deal with non-linearity and uncertainty through offline and online experiences. On the other hand, regardless of the type of a data-driven predictor, its disadvantage is that prediction performance depends on the quality and amount of collected data.

C. Hybrid predictor

Aiming at compensating disadvantages of model-based and data-driven predictors with each other, a hybrid predictor consists of both of them. In order to deal with the low fidelity of a mathematical model and the effect of environmental disturbance, NNs have been used to surrogate related terms in the mathematical model [18], [19].

In addition, it is gaining increasing attention to compensate the error made by a model-based predictor by using black-box models. Mei et al. [20] proposed a hybrid model of a reference vessel model and a random forest. First, their algorithm searches a reference vessel of which particulars are similar to those of a targeting vessel in their database; then the random forest is trained so that it compensates the error between the true acceleration of the targeting vessel and the predicted acceleration based on the reference vessel model. Skulstad et al. [21] made single-step-ahead prediction of acceleration by using a feed-forward NN which compensates the error between true and predicted acceleration made by a model-based predictor. According to [7], these models are grouped into single-output prediction strategy which makes multi-step-ahead prediction by iterating single-step-ahead prediction many times. Its advantage is the flexibility of its simple structure. However, some disadvantages are known when it is used for long-horizon prediction [7]. First, a single-output prediction is tuned for one-step-ahead prediction. It takes a predicted value as an input instead of measurements in the iteration. Therefore, the prediction performance in the distant future is not guaranteed. Second, time-consuming iterative prediction could be an obstacle of real-time decision making.

Skulstad et al. [4] proposed multiple-output prediction strategy of a hybrid predictor of which NN compensates the error between the 30s true and model-based predicted trajectories. They employed real sensor data of 88 separate docking operations and reported their hybrid predictor reduced the average distance error almost by half compared to a model-based predictor at 30s future. Since it makes 30s prediction at a time, long-horizon prediction performance is checked in the training process and the error does not propagate over the prediction horizon. It is notable they assume thruster commands at the current time are kept over the prediction horizon. Therefore, it can not evaluate the effect of future command assumption on trajectory prediction. This is a practical and key limitation of their study when it is used for predictive decision making.

TABLE I
SENSOR DATA USED IN THE INSIDE VESSEL MODEL AND NN

Input name	Unit	Range (train-val)	Range (test)
Heading	°	0.0/360.0	0.0/360.0
Surge speed	m/s	0.05/4.10	0.16/4.34
Sway speed	m/s	-0.83/0.84	-0.69/0.70
Yaw rate	°/s	-3.7/3.4	-2.6/3.9
Wind direction	°	2.3/356.7	2.2/348.7
Wind speed	m/s	5.0/7.0	5.0/7.0
Thruster revolution	RPM	26.0/128.0	27.0/128.0
Thruster angle	°	-19.6/19.5	-19.7/19.8

III. AN MHP WITH CONSIDERATION FOR FUTURE COMMAND CHANGE

Aiming at consideration of future command assumption, a new structure of an MHP is introduced in this section. In order to make use of trajectory prediction for the sake of better predictive decision making, the performance should be evaluated in terms of the prediction horizon that is longer than the time to make a change of course and speed [3]. This study evaluates 30s prediction performance that is enough for making a change of course and speed of the R/V Gunnerus. The hybrid predictor makes 30s North-East position prediction $[\hat{N}_1, ..., \hat{N}_k, ..., \hat{N}_{30}]$ and $[\hat{E}_1, ..., \hat{E}_k, ..., \hat{E}_{30}]$ where $\hat{N}_k$ and $\hat{E}_k$ are predicted North and East positions made by the hybrid predictor at ks future, respectively. The real-time measurements of wave, current, and swell are not given to a predictor since their accurate measurements are limited in most cases. Therefore, an NN compensates the error made by the model-based predictor

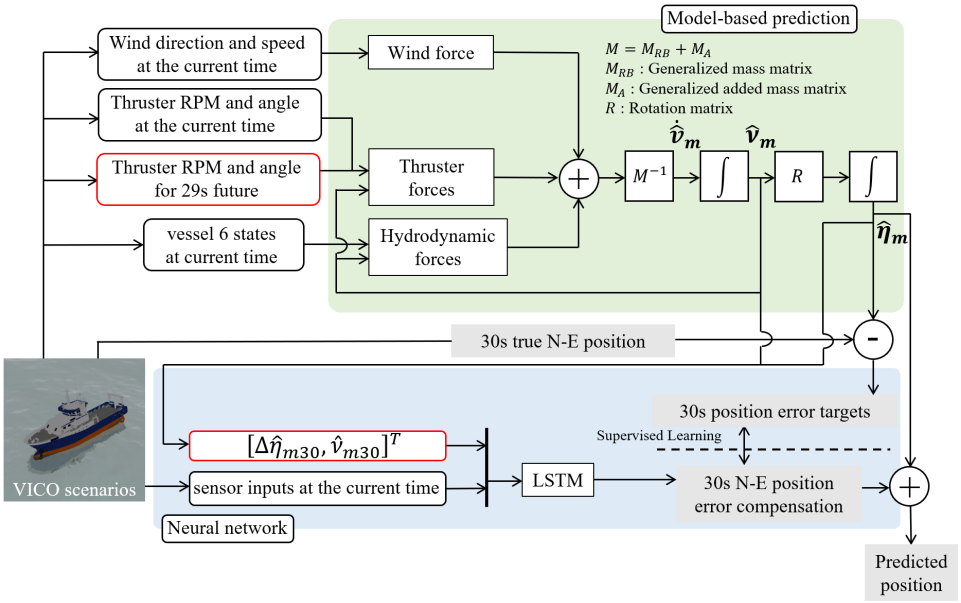


Fig. 1. A structure of the proposed hybrid predictor with consideration for future command change

based on a time series of velocities that are informed of how the vessel is oscillated by 1st-order wave forces and drifted by 2nd-order wave forces. We do not consider the effect of ocean current in the predictor. In order to deal with those, one can develop a hybrid predictor with those sensors and models based on this study. However, that is out of scope of this study. Loading conditions are not included in the input features of the NN as the NN is trained for a specific loading condition of a specific ship.

A. A structure of the proposed MHP

A structure of the proposed MHP is shown in Fig. 1. It consists of a model-based predictor highlighted in green and a Long Short-Term Memory (LSTM), which is one of the most popular architectures of Recurrent NNs (RNNs), highlighted in blue. At each time step of the operation scenario, the relative wind speed at the current time V_{rw}, the relative wind angle from bow at the current time γ_{rw}, thruster revolution n_0 and angle δ_0 at the current time, a thruster revolution vector $\mathbf{n} = [n_1, n_2, ..., n_{29}]$ and an angle vector $\boldsymbol{\delta} = [\delta_1, \delta_2, ..., \delta_{29}]$ for 29s future, and vessel state $[N_0, E_0, \psi_0, u_0, v_0, r_0]^T$ at the current time are provided to a model-based predictor where n_k is thruster revolution, δ_k is thruster angle at ks future. Please note that n_{30} and δ_{30} are not included in the inputs since they do not determine the position at 30s future that is determined by the vessel state and commands at 29s future. In the application of this study, $[\hat{N}_1, ..., \hat{N}_{30}]$ and $[\hat{E}_1, ..., \hat{E}_{30}]$ are calculated to evaluate the decision making $[n_0, ..., n_{29}]$ and $[\delta_0, ..., \delta_{29}]$. $\hat{\eta}_{mk} = [\hat{N}_{mk}, \hat{E}_{mk}, \hat{\psi}_{mk}]^T$ is a model-based predicted position vector of North, East, and heading at ks future. $\hat{\nu}_{mk} = [\hat{u}_{mk}, \hat{v}_{mk}, \hat{r}_{mk}]^T$ is a model-based predicted velocity vector of surge, sway, and yaw velocities at ks future. η_0 is a true position vector and ν_0 is a true velocity vector at the current time. A model-based predictor f_m produces $[\hat{\eta}_{m1}, ..., \hat{\eta}_{m30}, \hat{\nu}_{m1}, ..., \hat{\nu}_{m30}] = f_m(\eta_0, \nu_0, V_{rw}, \gamma_{rw}, n_0, \mathbf{n}, \delta_0, \boldsymbol{\delta})$. A model-based predicted trajectory deviates from a true trajectory due to the low fidelity of the vessel model, the ignorance of the measurement of wave and ocean current, and wave-frequency motion in $[\eta_0, \nu_0]^T$. Therefore, LSTM f_n aims to compensate the error in North and East positions made by the model-based predictor. The error compensation is described as $[\Delta\hat{N}_{n1}, ..., \Delta\hat{N}_{n30}]$ and $[\Delta\hat{E}_{n1}, ..., \Delta\hat{E}_{n30}]$ in North and East positions, respectively. A function of the error compensation is approximated by the LSTM f_n as $[\Delta\hat{N}_{n1}, ..., \Delta\hat{N}_{n30}, \Delta\hat{E}_{n1}, ..., \Delta\hat{E}_{n30}] = f_n(\sin\psi_0, \cos\psi_0, \nu_0, V_{gw}, \beta_{gw}, n_0, \delta_0, \Delta\hat{\eta}_{m30}, \hat{\nu}_{m30})$ where $\Delta\hat{\eta}_{m30} = [\hat{N}_{m30} - N_0, \hat{E}_{m30} - E_0, \hat{\psi}_{m30} - \psi_0]^T$ with consideration for future command assumption, V_{gw} is the global wind speed at the current time, and β_{gw} is the global wind direction at the current time. ψ_0 is mapped to $\sin\psi_0$ and $\cos\psi_0$ in order to avoid a jumping phenomenon of ψ_0 around $0°$. The number of inputs of the NN is 15 and that of outputs is 60. Finally, predicted North position $\hat{N}_k = \hat{N}_{mk} + \Delta\hat{N}_{nk}$ and East position $\hat{E}_k = \hat{E}_{mk} + \Delta\hat{E}_{nk}$ are calculated in the prediction phase.

B. A vessel model

The advantage of using hybrid predictor is to enhance prediction performance by combining a low-fidelity vessel model and a limited number of operation data. Therefore, a 3 Degrees of Freedom (DOF) maneuvering model (1)-(3) is chosen as a low-fidelity vessel model in the model-based predictor instead of highly nonlinear and complex 6DOF maneuvering and seakeeping model used for simulation scenario generation. The time step of numerical integration is 1s. A model-based prediction is independent of the Vico system.

$$(M_{RB} + M_A)\dot{\nu} + C(\nu)\nu + D(\nu)\nu = q \tag{1}$$

$$\dot{\eta} = R(\psi)\nu \tag{2}$$

$$q = q_{\text{wind}} + q_{\text{prop}} + q_{\text{man}} \tag{3}$$

M_{RB} is the generalized mass matrix, M_A is the generalized added mass at infinite wave frequency, C is the coriolis-centripetal matrix, D is the damping matrix for the model-based predictor. R is the rotation matrix between body-fixed coordinate and inertial coordinate. The prediction made by (1)-(3) deviates from a true trajectory made by a highly nonlinear and complex vessel model due to the following reasons. This aims at representing the deviation of a model-based prediction from a true trajectory in reality.

- Coefficients in (1) are introduced by removing 3rd-order velocity terms of the highly nonlinear and complex vessel model used for virtual scenario generation [22]. Therefore, (1)-(3) produce much deviation from a true trajectory when a vessel maneuvers rapidly.
- (1) has no retardation function.
- Wave forces acting on the vessel are not considered.
- The effect of ocean current is not modeled due to the ignorance of its measurement.
- The predicted trajectory oscillates around a true trajectory due to the wave-frequency component in $[\eta_0, \nu_0]^T$.

Three-dimensional wind force q_{wind} is given by (4).

$$q_{\text{wind}} = \frac{1}{2}\rho_a V_{rw}^2 \begin{bmatrix} C_X(\gamma_{rw})A_{Fw} \\ C_Y(\gamma_{rw})A_{Lw} \\ C_N(\gamma_{rw})A_{Lw}L_{oa} \end{bmatrix} \tag{4}$$

where ρ_a is the air density, C_X, C_Y, C_N are the wind coefficients tuned for the R/V Gunnerus, A_{Fw} is the frontal projected area, A_{Lw} is the lateral projected area, and L_{oa} is the length of the R/V Gunnerus. Mathematical thruster models provided by manufacturers calculate thruster forces (q_{prop} and q_{man}) based on vessel state, n_0, $\mathbf{n}$, δ_0, and $\boldsymbol{\delta}$. By summing up wind, thruster, and hydrodynamic forces in the surge, sway, and yaw direction, total forces and moment acting on the vessel are calculated. A numerical integration starts with vessel state at the current time; then $[\hat{\eta}_{m1}, ..., \hat{\eta}_{m30}, \hat{\nu}_{m1}, ..., \hat{\nu}_{m30}]$ is calculated.

C. An LSTM

An NN of the present hybrid predictor has an input layer, an LSTM layer/layers, and an output layer. The number of LSTM layers is optimized by hyperparameter tuning framework. An LSTM consists of three gates; namely, an input gate, an output gate, and a forget gate. These gates enable an LSTM to keep valuable information in the recurrent memory and forget unnecessary information by using a forget gate. It is good at dealing with long-term time-series information. An LSTM in Pytorch [23] calculates following functions.

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \tag{5}$$
$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \tag{6}$$
$$g_t = \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \tag{7}$$
$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \tag{8}$$
$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \tag{9}$$
$$h_t = o_t \odot \tanh(c_t) \tag{10}$$

where i_t is the input gate, f_t is the forget gate, g_t is the cell input function, o_t is the output gate, b terms are bias vectors, W terms are weight matrices, c_t is the cell state at time t, h_{t-1} is the hidden state at time $t-1$, x_t is the input at time t, σ is the sigmoid function, and $\odot$ is the element-wise product of vectors. The linear activation function is applied to the output layer.

As explained in the previous subsection, inputs to the LSTM are 8 sensor measurements at the current time as shown in Table. I in addition to $\delta\hat{\eta}_{m30}$ and $\hat{\nu}_{m30}$. Note that RPM is an abbreviation of Revolution Per Minute in the table. The mean Squared Error $e = \frac{1}{2}\sum_{k=1}^{30}\{(N_k - \hat{N}_k)^2 + (E_k - \hat{E}_k)^2\}$ of the output vector and true deviation of the model-based predictor is defined at each time step of each scenario. N_k and E_k are true North and East positions at ks future, respectively. Weights and biases of the LSTM are updated so that a summation of e over all time steps and scenarios becomes smaller. Hyperparameters are a set of parameters tuned before the training of an NN. Since the performance of an NN is significantly affected by hyperparameters, 5 hyperparameters (The number of LSTM layers *nlayer*, the number of units in the hidden layer *nunit*, learning rate *lr*, dropout rate of the input layer *dropin*, and dropout rate of recurrent information *droprr*) are automatically tuned by using hyperparameter tuning framework *optuna* [24] as shown in Table. II. *Optuna* searches optimal learning rate in the log domain as its range of search is wide. The validation loss converges around 50 trials. Therefore, the number of trials of parameter search of *optuna* is set to 50. Adam [25] is selected as an optimizer. The LSTM is developed in Pytorch [26] framework in Python. Each input variable and target values are z-score normalized in the training dataset. Target values are normalized since it contributed to better prediction performance in the experiment. The corresponding mean and standard deviation are applied to the test dataset.

$n_{\text{all}} = 200$ scenarios are generated in this study. The advantage of a hybrid predictor over a data-driven predictor is expected to deal with a limited number of dataset by having an inside vessel model. Therefore, n_{all} is fixed as we examine the prediction performance provided that the number of dataset is limited. After shuffling the scenarios, $n_{\text{test}} = 60$ scenarios of $n_{\text{all}} = 200$ scenarios generated by Vico are randomly selected as a test dataset and it is kept untouched in the training process. A training dataset consists of the remaining $n_{\text{train}} = 140$ scenarios and it is split into 5 datasets; then cross validation is performed. 4 datasets are used for the training and the other dataset is used for the validation of the trained LSTM. If prediction accuracy to the validation dataset does not improve for $n_e = 200$ epochs, the training process is stopped in order to avoid the over-fitting to the training dataset; then the best model is used for prediction. We have checked larger $n_e > 200$ contributes to the marginal improvement of prediction accuracy in the validation dataset as shown in Section IV. By switching the validation dataset 5 times, 5 LSTMs are trained independently. Their hyperparameters are the same. The final prediction to the test dataset is the average of predictions made by 5 LSTMs. Please note that scenarios in the training dataset are independent of those in the test dataset.

Training of the LSTM is performed offline in this study.

TABLE II
HYPERPARAMETER SEARCH BY *optuna* [24]

	Search range		Optimized	
Hyperparameter	Min	Max	Hybrid	Data-driven
nlayer	1	3	3	2
nunit	10	500	270	410
lr	0.0001	0.1	5.1×10^{-3}	1.1×10^{-3}
dropin	0.0	0.5	3.0×10^{-3}	1.7×10^{-3}
droprr	0.0	0.5	0.41	0.07

D. Consideration of future command assumption

A contribution of this study, which is the efficient consideration of future command assumption in an MHP, is highlighted herein. Being different from the work by Skulstad et al. [4], the proposed structure makes model-based prediction with consideration for a 29s time series of future command assumption. Subsequently, a model-predicted state only at 30s future $[\Delta\hat{\eta}_{m30}, \hat{\nu}_{m30}]^T$ and sensor inputs at the current time are given to the LSTM. Being different from a straightforward formulation of error compensation by the LSTM $[\Delta\hat{N}_{n1}, ..., \Delta\hat{N}_{n30}, \Delta\hat{E}_{n1}, ..., \Delta\hat{E}_{n30}] = f_n(\sin\psi_0, \cos\psi_0, \nu_0, V_{gw}, \gamma_{gw}, n_0, \mathbf{n}, \delta_0, \boldsymbol{\delta})$, an approximation f_n in this study can incorporate the effect of the future command assumption without expanding the number of input variables of the LSTM drastically by making use of model-based predicted future vessel state. By avoiding the curse of dimentionality, it contributes to the high training efficiency. This idea is robust to diverse application since the number of input variables are independent of the number of commands and the length of the control horizon.

In the training phase, $\boldsymbol{n}$ and $\boldsymbol{\delta}$ are given by a time history of a training dataset. In the prediction phase, $\boldsymbol{n}$ and $\boldsymbol{\delta}$ are given by a controller. In this study, commands are independent of real-time trajectory prediction since we focus only on prediction performance given that commands are given by a controller. Therefore, $\boldsymbol{n}$ and $\boldsymbol{\delta}$ are given by a time history of a test dataset in the prediction phase as well. By using this study, one can evaluate possible evasive actions based on their predicted consequences by changing future command assumptions. Subsequently, a controller can change its algorithm from a global path following to a local evasive action if the collision risk becomes critical. In order to incorporate this study into a controller and change control signals based on predicted states, stability issues will arise. That would be a future challenge in the control applications of this study.

IV. EXPERIMENT

A. Experimental setting

1) A vessel model in Vico: This study employs the virtual R/V Gunnerus in the validation study. Its physical parameters are shown in Table. III. Further details can be found in `https://www.ntnu.edu/oceans/gunnerus`. We conduct the simulation experiments in Vico, which is a Entity-Component-System-based co-simulation platform. Vico enables a vessel model to be constructed by assembling independent black-box models of a hull model, thruster models, and an user-defined controller. We generate realistic virtual operation scenarios by using a vessel model in Vico. The hull model used in Vico is a 6DOF high-fidelity maneuvering and seakeeping vessel model of the R/V Gunnerus, which was developed in the SimVal project [22], [28]. Environmental disturbances (wave, wind, and current) can be manipulated by users; then, their forces acting on the vessel are calculated in the black-box hull model. Thruster models are provided by thruster manufactures. Further details of Vico and the hull model can be found in [8], [22], [28].

TABLE III
SPECIFICATIONS OF THE R/V GUNNERUS

Specifications	Value
Mass	370t
Deadweight	107t
Length between perpendiculars	28.9m
Breadth middle	9.6m
Draught	2.7m

TABLE IV
TWO SEA STATES USED IN THE SCENARIO GENERATION INSPIRED BY BEAUFORT WIND FORCE SCALE [27]

	Beaufort wind force scale	
	3: Gentle breeze	4 : Moderate breeze
Wind speed (m/s)	5.0	7.0
Significant wave height (m)	0.6	1.0
Significant wave period (s)	3.9	5.0
Ocean current speed (m/s)	0.1	0.2

2) Simulation setup: An original scenario is a 215s time history. Waves and winds are ramped up from 0s to 50s. The position and velocity of the vessel are reset to zero at 50s. A 160s time history from 55s to 215s is extracted for an experiment in this study. Hereinafter, the beginning of the extracted scenario is defined as $t = 0$s. As each time step requires a 30s trajectory in the future in the training process, 130s from $t = 0$s to $t = 130$s is defined as one scenario. The virtual R/V Gunnerus is equipped with two azimuth thrusters. The same thruster revolution and angle are given to them in this study. Maximum change rates of thruster revolution and angle are set in order to avoid sudden change of commands. Thruster revolution is set to $n_{\max}$ until $t = T_n$, and then turned back to $0.5n_{\max}$ with the maximum change rate. Thruster angle is set to $\delta_{\max}$ until $t = T_\delta$, and then turned back to zero with the maximum change rate. In this study, commands are simultaneously applied to

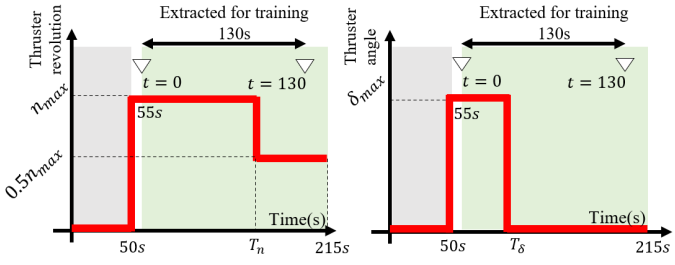


Fig. 2. Pre-defined time series of thruster revolution and angle

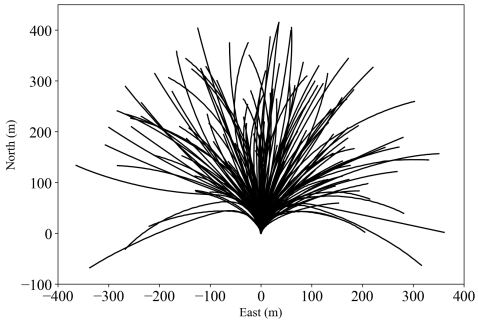


Fig. 3. Trajectories in the training dataset

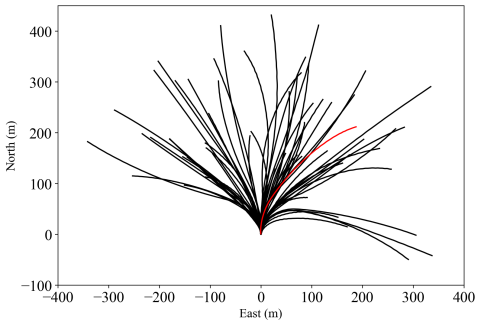


Fig. 4. Trajectories in the test dataset.

thruster models. This is a reasonable simplification since the difference between the command and feedback values of the R/V Gunnerus is very small. In the prediction phase, a time series of future command values are given to a predictor. Each scenario is parameterized by a scenario parameter vector $P = [\text{Beaufort wind force scale}, n_{\max}, \delta_{\max}, \theta_d, T_n, T_\delta]$. $T_n \in [20s, 50s]$, $T_\delta \in [80s, 110s]$, $n_{\max} \in [50\text{RPM}, 130\text{RPM}]$, $\delta_{\max} \in [-20°, 20°]$, and $\theta_d \in [0°, 360°)$ are randomly selected for each scenario; then, $n_{\text{all}} = 200$ unique scenarios are generated. The time step of scenario generation is 0.05s that is the default value of the vessel model simulated in Vico and saved in 1Hz for the training. The vessel in the scenario is exposed to wind, wave, and ocean current. A calculation of the first-order wave force, the second-order wave force, wind force, and the effect of ocean current are done in the vessel model simulated in Vico by defining the type of wave and wind spectrum, mean wind speed U_w, significant wave height H_s, significant wave period T_s of the spectrum, ocean current speed U_c, and the global direction of environmental disturbances θ_d to the north. JONSWAP spectrum [29] is chosen for the type of wave spectrum. The static wind is applied to the vessel. We generate scenarios under two levels of environmental disturbances as shown in Table. IV based on Beaufort wind force scale [27]. Beaufort wind force scale is level 3: Gentle breeze or level 4: Moderate breeze and defines the combination of (H_s, T_s, U_w, U_c) of the scenario. For the sake of simplicity, θ_d of wind and ocean current is the same as that of the wave. Note that θ_d does not change over time in one scenario. Half of 200 scenarios are under Gentle breeze disturbances and the others are under Moderate breeze disturbances.

Trajectories in the training dataset are shown in Fig. 3 and those in the test dataset are shown in Fig. 4. We can see that generated trajectories are different with each other since T_n, T_δ, $n_{\max}$, $\delta_{\max}$, and θ_d are randomized. A trajectory $P = [\text{Beaufort wind force scale} = 4, n_{\max} = 102.2\text{RPM}, \delta_{\max} = -9.0°, T_\delta = 38s, T_n = 96s, \theta_d = 14.4°]$ is shown in red in Fig. 4. We can see that the vessel is exposed to Moderate breeze disturbance from the north and the vessel gradually turns clockwise.

B. Evaluation metrics

We introduce some metrics for plausible evaluations of prediction performance. We have $n_{\text{test}} = 60$ scenarios in the test dataset, which is 30% of n_{all}. Prediction error $l_{ijk} = \sqrt{(N_{ijk} - \hat{N}_{ijk})^2 - (E_{ijk} - \hat{E}_{ijk})^2}$ of the scenario i, at a js time step, and ks future is defined where N_{ijk} is a true North position, $\hat{N}_{ijk}$ is a predicted North position, E_{ijk} is a true East position, and $\hat{E}_{ijk}$ is a predicted East position. We introduce mean prediction error for the prediction horizon $\overline{l_k}$ as (11) where $T = 130$s is the length of a scenario.

$$\overline{l_k} = \frac{1}{n_{\text{test}}} \frac{1}{T} \sum_{i=1}^{n_{\text{test}}} \sum_{j=0}^{T-1} l_{ijk} \tag{11}$$

In order to examine the distribution of prediction error, we define 9th decile $l_k^{90\%}$ which is the boundary value of the top 10 percentile of l_{ijk} in the test dataset. We look into how a summation of prediction error for 30s changes over time of scenario i as $S_{ij} = \sum_{k=1}^{30} l_{ijk}$. It should be noted that $\overline{S}_j = 1/n_{\text{test}} \sum_{i=1}^{n_{\text{test}}} S_{ij}$ and $\overline{S}_i = 1/T \sum_{j=0}^{T-1} S_{ij}$ in the following sections.

C. Reference predictors

In a validation study, two reference predictors are used in addition to the proposed hybrid predictor for the sake of better understanding of the results. In this study, we assume only a low-fidelity vessel model and a limited number of data are available. The prediction performance of this study is evaluated by comparing with that of the model-based predictor based solely on the given low-fidelity vessel model and the data-driven predictor trained solely by the given limited dataset. A model-based prediction is $[\hat{N_{m1}}, ..., \hat{N_{m30}}]^T$ and $[\hat{E_{m1}}, ..., \hat{E_{m30}}]^T$ without the error compensation made by the LSTM. A data-driven predictor is a multiple-output LSTM f_d which produces $[\hat{N}_1 - N_0, ..., \hat{N}_{30} - N_0, \hat{E}_1 - E_0, ..., \hat{E}_{30} - E_0]^T = f_d(\sin\psi_0, \cos\psi_0, \nu_0, V_{gw}, \beta_{gw}, n_0, \delta_0, \mathbf{n}, \boldsymbol{\delta})$ without the help of a model-based predictor. Therefore, its dimension of inputs is 67 and that of outputs is 60. One should note this experiment does not show the general quantitative limitation of the model-based and data-driven measures. Their prediction performance becomes better if a high-fidelity vessel model and a large dataset are available, however, that is not the scope of the hybrid modeling.

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/JSEN.2021.3119069, IEEE Sensors Journal

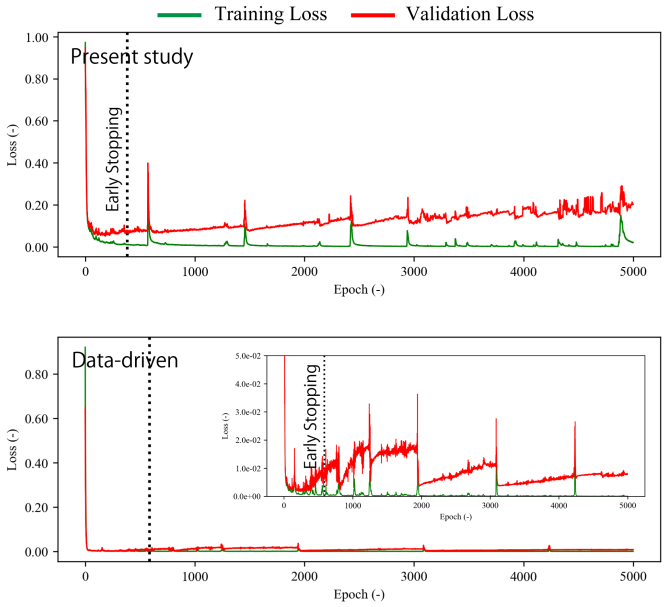


Fig. 5. Histories of validation and training losses of the present hybrid and data-driven predictors with optimized hyperparameters in the 1st fold of the cross validation

D. Performance validation

In Fig. 5, the top panel shows how validation loss (shown in red) and training loss (shown in green) of the present hybrid predictor with optimized hyperparameters changes over epochs in the 1st fold of the cross validation. The bottom panel shows that of the data-driven predictor with optimized hyperparameters.

The small figure in the panel shows the enlarged graph. The dotted lines show the number of epochs where early stopping is invoked. The validation loss increases after early stopping is invoked whereas the training loss keeps decreasing in both panels. It indicates both of the LSTMs are well-trained while avoiding overfitting to the training dataset as much as possible by using appropriate setting of n_e.

Fig. 6 - Fig. 8 show results of a scenario of which $\overline{S}_i$ of the present study is the largest among the scenarios in the test dataset. $P = [\text{Beaufort wind force scale} = 4, n_{\max} = 102.2\text{RPM}, \delta_{\max} = -9.0°, T_\delta = 38s, T_n = 96s, \theta_d = 14.4°]$ in this scenario and its trajectory is shown in red in Fig. 4. Time histories of S_{ij} of present hybrid, model-based, and data-driven predictors are shown in Fig. 6. Time histories of heading, surge speed, sway speed, yaw rate, thruster revolution and thruster angle are shown in Fig. 7. Global wind direction and speed are not shown in Fig. 7 as they do not change in one scenario. In Fig. 6, we can see that the prediction performance of the present hybrid predictor outperforms that of the model-based and data-driven predictors. The model-based prediction stays unstable due to the low fidelity of the model and wave-frequency components in $[\eta_0, \nu_0]$. In order to remove the wave-frequency components from the model-based prediction, one can implement a low-pass filter. However, the tuning of the low-pass filter takes effort and the time delay of the processed signal is inevitable. In the beginning of the scenario, the performances of the hybrid and data-driven predictors are not as good as that of the hybrid predictor after $t = 30$s as they have no sufficient recurrent information. The present hybrid predictor maintains its prediction error at low levels whereas the data-driven predictor produces more error. It should be noted that we do not see the increment of prediction error of the present study around $t = T_\delta$ and $t = T_n$. It indicates that the present hybrid predictor successfully handles a time series of future command assumption while reducing the error made by the low-fidelity model-based predictor.

Snapshots of predictions at (A) $t = 35$s and (B) $t = 91$s are shown in Fig. 8. At (A), the vessel turns back the thruster angle from $\delta = -9.0°$ when making prediction to $\delta = 0°$ at $t = T_\delta$ after making prediction. The dotted line, which is a true trajectory for 30s, shows that the vessel turns clockwise in the beginning of the prediction, however, it gradually goes straight in the last half of the 30s prediction horizon. One can see that the present hybrid predictor captures this behavior and its prediction is more accurate than the other predictors. At (B), the vessel deaccelerates at $t = T_n$ after making prediction. The model-based prediction ends up with a longer trajectory than the true trajectory, whereas the prediction performance of the present hybrid predictor is very good.

The overall prediction performance in the test dataset is investigated herein. Fig. 9 shows the mean prediction error $\overline{l_k}$ in the left panel and the 90% boundary value of prediction error $l_k^{90\%}$ in the right panel over 30s prediction horizon. Solid lines show the prediction performance under lower disturbance (Beaufort wind force scale: 3) and dotted lines show that under higher disturbances (Beaufort wind force scale: 4). In the left figure, LSTM-based predictors (the data-driven and hybrid predictors) outperform the model-based predictor. In particular, the prediction performance of the model-based predictor under the higher disturbance deteriorates significantly whereas its effect on LSTM-based predictors are marginal. In addition, it is discerned that the present hybrid predictor outperforms the data-driven predictor notably especially at the longer prediction horizon. In terms of $l_k^{90\%}$, we see the same trend in the right figure. It indicates that the present hybrid predictor contributes to reducing not only mean prediction error but also the frequency of the large prediction error.

Fig. 10 shows the time history of the mean 30s error summation $\overline{S}_j$ in the test dataset. The range of $T_\delta \in [20s, 50s]$ and $T_n \in [80s, 110s]$ are highlighted in light blue and yellow in the figure. Although we see the increment of prediction error of the data-driven and hybrid predictors in the beginning of the scenario due to the insufficient recurrent information, $\overline{S}_j$ of the present hybrid predictor is significantly smaller than that of the model-based and data-driven predictors. It is notable that the prediction performance of the present study does not deteriorate due to the future command change as we see no increment of error in the range of $T_\delta \in [20s, 50s]$ and $T_n \in [80s, 110s]$.

Fig. 11 compares the mean prediction error $\bar{l}_k$ of three predictors under the disturbances of Beaufort wind force scale 4 at 30s future. The present hybrid predictor reduces the prediction error by 81.8% compared to the model-based predictor based solely on the low-fidelity vessel model and by 45.6% compared to the data-driven predictor trained solely by

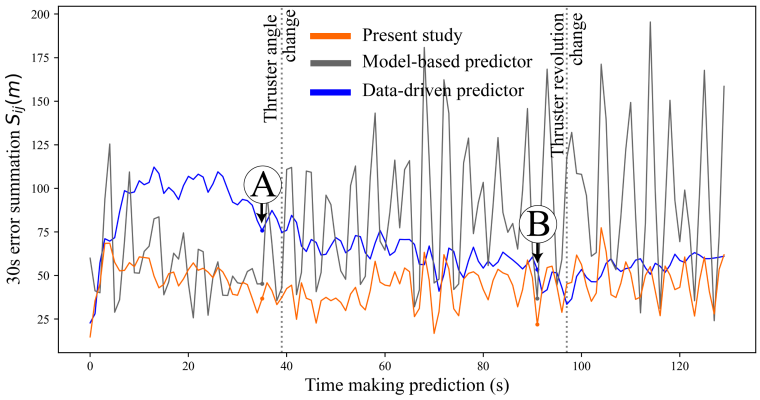


Fig. 6. Time histories of S_{ij} of 3 predictors

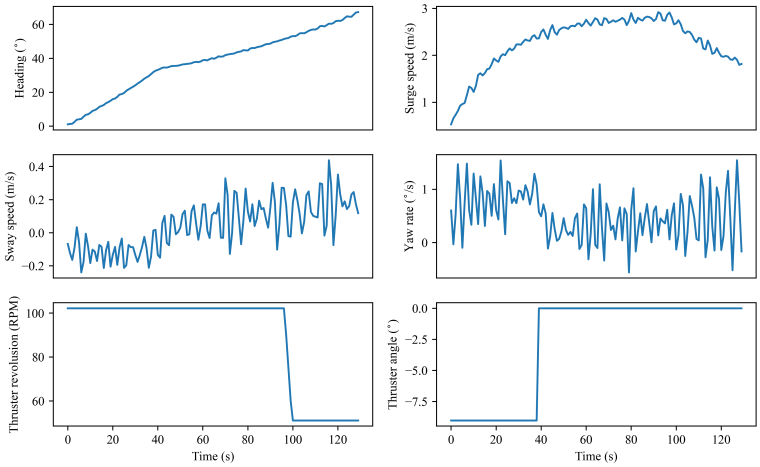


Fig. 7. Time histories of sensor measurements

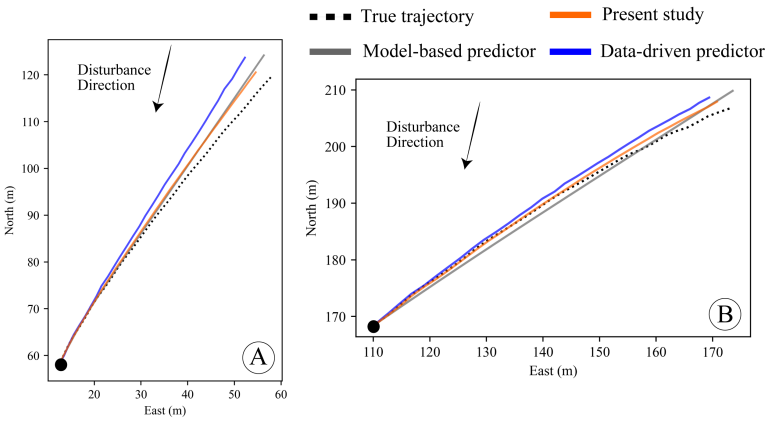


Fig. 8. Snapshots of predictions at (A)$t = 35$s and (B)$t = 91$s of the scenario shown in Fig. 6 and Fig. 7

a limited number of dataset.

In Fig. 12, S_{ij}s made by a model-based predictor, present study, and the data-driven predictors are plotted. The clockwise angle of a plot shows the local direction of environmental disturbances (0° for the head sea and 180° for the following wave) when making a prediction. The distance from a plot to the origin shows the value of S_{ij}. It is clear that the prediction performance of the present study outperforms that of the other predictors regardless of the local direction of environmental disturbances. In Fig. 13, S_{ij} and surge speed when making prediction of the three predictors are plotted. The present study makes prediction accurately and stably in the whole range of surge speed. However, the model-based and data-driven predictors produces much error in the whole range of surge speed. The results in Fig. 12 and Fig. 13 show the robustness of the present study under different environmental disturbances

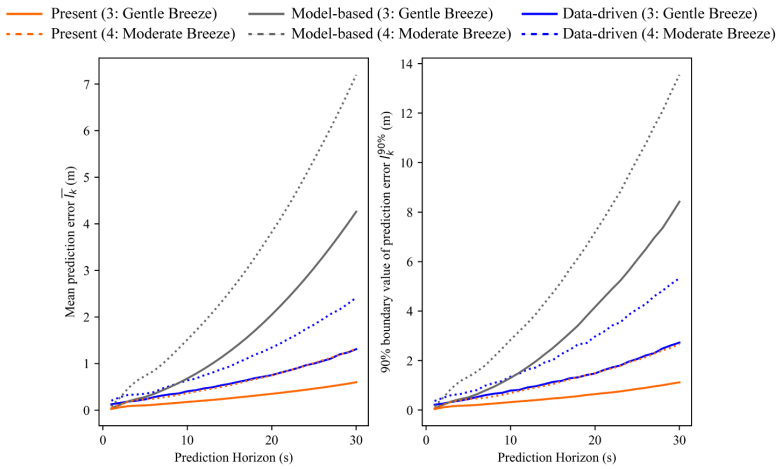


Fig. 9. Mean and 90% boundary value of prediction error over the 30s prediction horizon in the test dataset

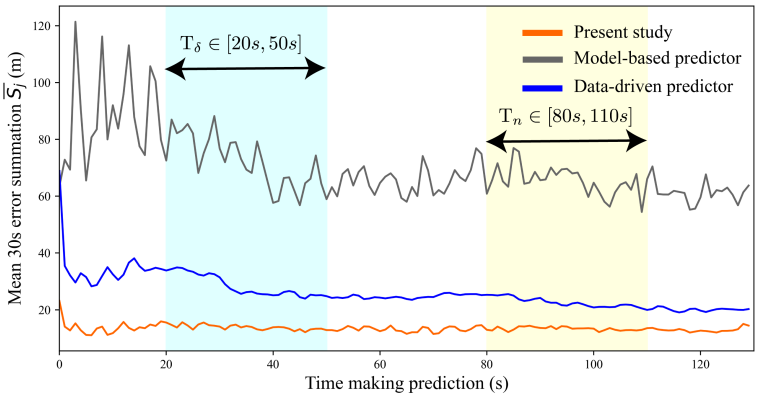


Fig. 10. Time histories of $\overline{S}_j$

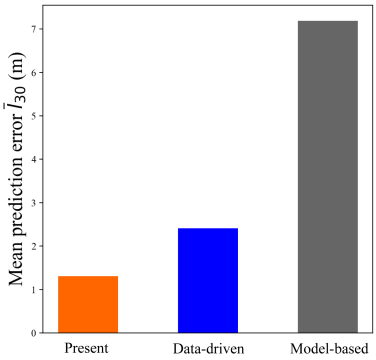


Fig. 11. Mean prediction error $\overline{l}_k$ under Moderate Breeze disturbances at 30s future.

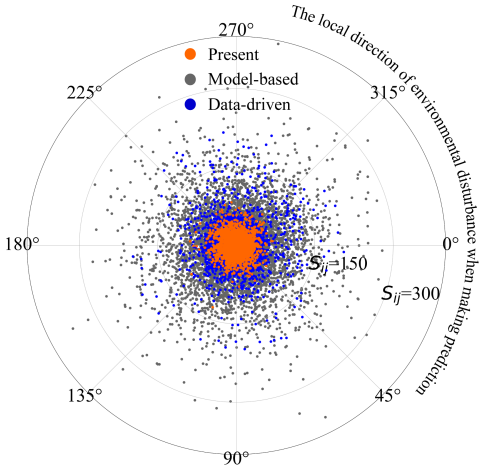


Fig. 12. Prediction error S_{ij} and local direction of environmental disturbance

and maneuverings.

E. Discussion

The presented results indicate the validity and benefit of proposed predictor that handles multidimensional future command assumption over the control horizon when only the low-fidelity vessel model and a limited number of data are available compared to the model-based predictor based solely on the low-fidelity vessel model and the data-driven predictor trained solely by a limited number of data. It is notable that no excessive increment of prediction error is seen even though future command assumption changes significantly over the control horizon in the experiment. It indicates present work overcomes the limitation of the conventional MHP. Thereby, a controller can evaluate the future command assumption

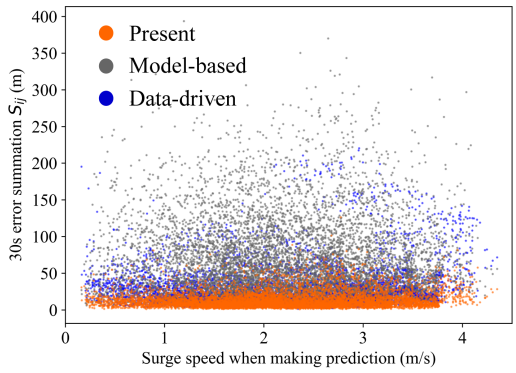


Fig. 13. Prediction error S_{ij} and surge speed at the current time

based on hybrid-predicted future trajectory. It contributes to better predictive decision making (e.g., MPC, the evaluation of evasive actions based on its predicted consequences) based on the hybrid modeling of ship dynamics. In the experiment of present work, prediction performance of present study in the test dataset is robust to the local direction of environmental disturbance and surge speed although the interaction between environmental disturbances and the ship in the test dataset is not explicitly included in the training dataset since we select $n_{\text{text}} = 60$ scenarios that is independent of the other $n_{\text{train}} = 140$ scenarios in the training dataset. In practice, this result is very important as it is impossible to enumerate all possible interactions of environmental disturbances and ships in the training dataset.

In this study, a best practice of a hybrid predictor is still an open question. For example, by using a hybrid predictor, how much can we compromise the fidelity of a vessel model? In which situation is a hybrid predictor superior to model-based and data-driven predictors? Future research needs to answer these questions in order to make this idea more practical to the industry. In this validation study, thruster commands are simultaneously applied to actual thruster revolution and angle since their difference of the R/V Gunnerus is very small. If they have a large difference and time delay, a function f_c that calculates feedback values from command values is required in order to apply the present study. As f_c can be incorporated into the NN's approximation f_n or be developed according to the characteristics of thrusters, the framework of the present study will be valid in that case as well.

V. Conclusion

This study proposed a new multiple-output hybrid ship trajectory predictor, that plays a key role in integrating data-driven measures based on sensor data and domain knowledge, and examine the performance of its methodology. It can handle the effect of future command assumption on trajectory prediction, that has been a practical limitation of a conventional multiple-output hybrid predictor. The proposed idea enables consideration of future command assumption without a drastic expansion of the input dimension of an internal neural network due to having its multidimensional time series over the control horizon.

Assuming only a low-fidelity vessel model and a limited number of data are available, prediction performance of present work is compared to that of the model-based predictor based solely on a low-fidelity vessel model and the data-driven predictor trained solely by a limited number of data. Prediction performance under different types/levels of environmental disturbances is validated for the first time by using comprehensive simulation scenarios generated by the co-simulation platform Vico. The proposed hybrid predictor works accurately and stably regardless of the local direction of environmental disturbances and the surge speed. It reduces the mean prediction error by 81.8% compared to the model-based predictor and by 45.6% compared to the data-driven predictor at 30s future under Beaufort wind force scale 4 level wind, wave, and ocean current. It is notable that no excessive increment of prediction error is seen even though future command assumption changes significantly over the control horizon in the experiment.

A presented comprehensive validation study provides the evidence of the effectiveness of the proposed hybrid predictor by handling a time series of future command assumption, that plays a key role in the predictive decision making. In the future, we plan to share real measurements of onboard sensor data with the real vessel, its digital model, and the remote control centre. Present study will be a key algorithm of utilizing sensor data for better decision making.

Our future work will revolve around three research topics. First online learning of a hybrid predictor will be a key technology in order to deal with a change of loading condition in real time. Second a best practice of a hybrid predictor needs to be investigated. The effect of the amount / quality of data and the fidelity of mathematical model on prediction performance is yet to be fully understood. Last the implementation of present work into a model predictive controller is needed.

References

[1] P. A. Lindahl, D. H. Green, G. Bredariol, A. Aboulian, J. S. Donnal, and S. B. Leeb, "Shipboard Fault Detection Through Nonintrusive Load Monitoring: A Case Study," *IEEE Sensors Journal*, vol. 18, no. 21, pp. 8986–8995, 2018.

[2] I. N. Controller, "Investigation of a Multitasking System for Automatic Ship Berthing in Marine Practice Based on an Integrated Neural Controller," no. 1997, 2020.

[3] T. A. Johansen, T. Perez, and A. Cristofaro, "Ship collision avoidance and COLREGS compliance using simulation-based control behavior selection with predictive hazard assessment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3407–3422, 2016.

[4] R. Skulstad, G. Li, T. I. Fossen, B. Vik, and H. Zhang, "A Hybrid Approach to Motion Prediction for Ship Docking - Integration of a Neural Network Model into the Ship Dynamic Model," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021.

[5] M.-G. Seo, B. W. Nam, and Y.-G. Kim, "Numerical Evaluation of Ship Turning Performance in Regular and Irregular Waves," *Journal of Offshore Mechanics and Arctic Engineering*, vol. 142, no. 2, 2020.

[6] G. Li, B. Kawan, H. Wang, and H. Zhang, "Neural-network-based modelling and analysis for time series prediction of ship motion," *Ship Technology Research (Schiffstechnik)*, vol. 64, no. 1, pp. 30–39, 2017. [Online]. Available: http://dx.doi.org/10.1080/09377255.2017.1309786

[7] S. Ben Taieb, A. Sorjamaa, and G. Bontempi, "Multiple-output modeling for multi-step-ahead time series forecasting," *Neurocomputing*, vol. 73, no. 10-12, pp. 1950–1957, 2010.

[8] L. I. Hatledal, Y. Chu, A. Styve, and H. Zhang, "Vico: An entity-component-system based co-simulation framework," *Simulation Modelling Practice and Theory*, vol. 108, no. September 2020,

p. 102243, 2021. [Online]. Available: https://doi.org/10.1016/j.simpat.2020.102243

[9] L. Zhao and M.-i. Roh, "COLREGs-compliant multiship collision avoidance based on deep reinforcement learning," *Ocean Engineering*, vol. 191, no. May, p. 106436, 2019. [Online]. Available: https://doi.org/10.1016/j.oceaneng.2019.106436

[10] L. P. Perera, "Navigation vector based ship maneuvering prediction," *Ocean Engineering*, vol. 138, no. November 2016, pp. 151–160, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.oceaneng.2017.04.017

[11] T. I. Fossen, *Handbook of marine craft hydrodynamics and motion control*. Wiley, 2011.

[12] W. Luo and X. Li, "Measures to diminish the parameter drift in the modeling of ship manoeuvring using system identification," *Applied Ocean Research*, vol. 67, pp. 9–20, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.apor.2017.06.008

[13] M. Zhu, A. Hahn, Y. Q. Wen, and A. Bolles, "Identification-based simplified model of large container ships using support vector machines and artificial bee colony algorithm," *Applied Ocean Research*, vol. 68, pp. 249–261, 2017. [Online]. Available: https://doi.org/10.1016/j.apor.2017.09.006

[14] M. S. Triantafyllou, M. Bodson, and M. Athans, "Real Time Estimation of Ship Motions Using KalmanFiltering Techniques," *IEEE Journal of Oceanic Engineering*, vol. 8, no. 1, pp. 9–20, 1983.

[15] B. Kawan, H. Wang, G. Li, and K. Chhantyal, "Data-driven Modeling of Ship Motion Prediction Based on Support Vector Regression," in *Proceedings of the 58th Conference on Simulation and Modelling (SIMS 58) Reykjavik, Iceland, September 25th – 27th, 2017*, vol. 138, 2017, pp. 350–354.

[16] H. Rong, A. P. Teixeira, and C. Guedes Soares, "Ship trajectory uncertainty prediction based on a Gaussian Process model," *Ocean Engineering*, vol. 182, no. December 2018, pp. 499–511, 2019. [Online]. Available: https://doi.org/10.1016/j.oceaneng.2019.04.024

[17] W. Y. Duan, L. M. Huang, Y. Han, and R. Wang, "IRF - AR model for short-term prediction of ship motion," in *Proceedings of the International Offshore and Polar Engineering Conference*, vol. 2015-Janua. Kona, Big Island, Hawaii, USA: International Society of Offshore and Polar Engineers, 2015, pp. 59–66.

[18] P. W. van de Ven, T. A. Johansen, A. J. Sørensen, C. Flanagan, and D. Toal, "Neural network augmented identification of underwater vehicle models," *Control Engineering Practice*, vol. 15, no. 6, pp. 715–725, 2007.

[19] R. Li, J. Huang, X. Pan, Q. Hu, and Z. Huang, "Path following of underactuated surface ships based on model predictive control with neural network," *International Journal of Advanced Robotic Systems*, vol. 17, no. 4, pp. 1–10, 2020. [Online]. Available: https://doi.org/10.1177/1729881420945956

[20] B. Mei, L. Sun, and G. Shi, "White-Black-Box Hybrid Model Identification Based on RM-RF for Ship Maneuvering," *IEEE Access*, vol. 7, pp. 57 691–57 705, 2019.

[21] R. Skulstad, G. Li, T. I. Fossen, T. Wang, and H. Zhang, "A Co-operative Hybrid Model For Ship Motion Prediction," *Modeling, Identification and Control: A Norwegian Research Bulletin*, vol. 42, no. 1, pp. 17–26, 2021.

[22] A. Ross, V. Hassani, Selvik, and D. Fathi, "Identification of Nonlinear Manoeuvring Models for Marine Vessels Using Planar Motion Mechanism Tests," in *Proceedings of the ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering*, St.John's, Newfoundland, Canada, 2015.

[23] H. Sak, A. Senior, and F. Beaufays, "Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition," no. Cd, 2014. [Online]. Available: http://arxiv.org/abs/1402.1128

[24] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Anchorage, AK, USA: Association for Computing Machinery, 2019, pp. 2623–2631. [Online]. Available: https://doi.org/10.1145/3292500.3330701

[25] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15, 2015.

[26] A. Paszke, S. Gross, J. Bradbury, Z. Lin, Z. Devito, F. Massa, B. Steiner, T. Killeen, and E. Yang, "PyTorch : An Imperative Style , High-Performance Deep Learning Library," no. NeurIPS, 2019.

[27] Met Office, "Beaufort wind force scale used in Met Office." [Online]. Available: https://www.metoffice.gov.uk/weather/guides/coast-and-sea/beaufort-scale

[28] V. Hassani, D. Fathi, A. Ross, F. Sprenger, Selvik, and T. E. Berg, "Time domain simulation model for research vessel Gunnerus," in *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE*, vol. 7, St.John's, Newfoundland, Canada, 2015.

[29] K. Hasselmann, T. P. Barnett, E. Bouws, H. Carlson, D. E. Cartwright, K. Eake, J. A. Euring, A. Gicnapp, D. E. Hasselmann, P. Kruseman, A. Meerburg, P. Mullen, D. J. Olbers, K. Richren, W. Sell, and H. Walden, "Measurements of wind-wave growth and swell decay during the joint North Sea wave project (JONSWAP)." *Deutsche Hydrographische Zeitschrift*, vol. 12, no. Deutches Hydrographisches Institut, 1973.

Motoyasu Kanazawa received the M.Sc. degree in environment from the University of Tokyo, Chiba, Japan in 2019. After his graduation, he joined Japan Marine United Corporation, Yokohama, Japan as a structural engineer. He is currently pursuing the Ph.D. degree with the Intelligent Systems Laboratory, Department of Ocean Operations and Civil Engineering, the Norwegian University of Science and Technology, Ålesund, Norway.

Robert Skulstad received his M.Sc. degree at the department of Engineering Cybernetics and Ph.D. degree at the department of Ocean Operations and Civil Engineering from the Norwegian University of Science and Technology, Norway, in 2014 and 2021, respectively. He now works as a post-doctoral research fellow at the department of Ocean Operations and Civil Engineering at NTNU in Ålesund. His research interests include ship motion prediction, machine learning and ship motion control.

Guoyuan Li (M'14-SM'19) received the Ph.D. degree in computer science from the Department of Informatics, Institute of Technical Aspects of Multimodal Systems, University of Hamburg, Hamburg, Germany, in 2013. In 2014, he joined the Department of Ocean Operations and Civil Engineering, Intelligent Systems Laboratory, Norwegian University of Science and Technology (NTNU), Ålesund, Norway. He is currently a Professor of ship intelligence at NTNU. He has published more than 70 articles in the areas of his research interests, including modeling and simulation of ship motion, autonomous navigation, intelligent control, optimization algorithms, and locomotion control of bioinspired robots.

Lars Ivar Hatledal Dr. Lars Ivar Hatledal received a B.Sc. degree in automation and M.Sc degree in simulation and visualization from NTNU, Ålesund, Norway, in 2013 and 2017 respectively. From 2013 to 2017 he worked as a part-time research assistant with the Intelligent System lab at NTNU Ålesund, Department of Ocean Operations and Civil Engineering. In 2021 he received a Ph.D degree from NTNU and is currently working as an associate professor with the Department of ICT and Natural Sciences in NTNU Ålesund.

Houxiang Zhang (M'04- SM'12) received the M.Sc. degree and Ph.D. degree both on Mechanical and Electronic Engineering from Robotics Institute, Beihang University in 2000 and 2003 respectively. From 2004, he worked as Postdoctoral fellow, senior researcher at the Institute of Technical Aspects of Multimodal Systems (TAMS), Department of Informatics, Faculty of Mathematics, Informatics and Natural Sciences, University of Hamburg, Germany. In Feb. 2011, he finished the Habilitation on Informatics at University of Hamburg. Dr. Zhang joined the NTNU, Norway in April 2011 where he is a Professor on Mechatronics. Dr. Zhang has engaged into two main research areas including biological locomotion control, and digitalization and virtual prototyping in demanding marine operation. He has applied for and coordinated more than 30 projects supported by Norwegian Research Council (NFR), German Research Council (DFG), EU, and industry. In these areas, he has published over 200 journal and conference papers as author or co-author. Dr. Zhang has received four best paper awards, and four finalist awards for best conference paper at International conference on Robotics and Automation. He has been elected to the Norwegian Academy of Technological Sciences in 2019.