Ocean Engineering 369 (2026) 128869

Contents lists available at ScienceDirect

Ocean Engineering

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

Research paper

Data-driven modelling and prediction of ship manoeuvring motions via Dynamic Bayesian Networks

Cihad Çelik [a,b], Huanhuan Li [c,*], Lu Zou [d,e], Musa Bashir [f], Zaili Yang [a,**]

[a] Liverpool Logistics, Offshore and Marine Research Institute, Liverpool John Moores University, Liverpool, UK
[b] Faculty of Naval Architecture and Ocean Engineering, Istanbul Technical University, Istanbul, Türkiye
[c] School of Engineering, University of Southampton, Southampton, UK
[d] School of Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, China
[e] State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai, China
[f] Department of Civil and Environmental Engineering, School of Engineering, University of Liverpool, UK

ARTICLE INFO

Keywords:
Data-driven modelling
System identification
Dynamic Bayesian Networks
Ship manoeuvring
Probabilistic prediction

ABSTRACT

Accurate modelling and prediction of ship manoeuvring motions are essential for navigation safety and control system development. Conventional physics-based manoeuvring models typically depend on predefined hydrodynamic formulations and extensive captive or free-running tests, which constrain their adaptability to varying operating conditions. This study develops a data-driven Gaussian Dynamic Bayesian Network (DBN) framework with time-lagged dependency learning for modelling ship manoeuvring dynamics and recursively predicting their evolution under new manoeuvring scenarios. The DBN explicitly characterises the temporal and conditional relationships between surge, sway, yaw motions, and rudder actions, allowing both dynamic modelling and interpretable assessment of underlying motion couplings. Free-running model tests of the KCS benchmark ship in deep and calm water are used for model training and testing. The DBN demonstrates strong predictive capability for both zig-zag and turning manoeuvres, with its performance evaluated against Long Short-Term Memory (LSTM) and Gaussian Process Regression (GPR) models. The results show competitive prediction performance across the investigated motion variables, with the relative accuracy varying among the three models, while the DBN requires substantially lower training time than the others. In addition to its predictive performance, the DBN provides a transparent representation of time-lagged dynamic dependencies, facilitating engineering interpretation of manoeuvring behaviour. Overall, the results demonstrate the potential of the proposed framework for recursive ship manoeuvring prediction.

1. Introduction

The modelling of ship manoeuvring motions continues to play a central role in advancing maritime autonomy and control. Over the past decades, numerous approaches have been developed to describe the complex nonlinear dynamics governing vessel motions, which can be broadly classified into two principal categories: physics-based and data-driven modelling. Physics-based formulations remain the traditional foundation of the field, while data-driven approaches, enabled by modern computational power and data availability, have gained increasing attention in recent years (Xu and Guedes Soares, 2025).

Physics-based manoeuvring models express ship dynamics through physically interpretable hydrodynamic derivatives, typically determined via captive model experiments, Computational Fluid Dynamics (CFD) simulations, or System Identification (SI) techniques (Araki et al., 2012). These models, which follow the classical Abkowitz formulation (Abkowitz, 1964) or the Maneuvering Modeling Group (MMG) model (Yasukawa and Yoshimura, 2015), have been successfully applied to a wide variety of vessel types and operating conditions. Guo and Zou (2017), for example, determined the hydrodynamic derivatives of the ONR tumblehome model using RANS-based CFD simulations for 4 Degrees of Freedom (DoF). Similarly, Yang and el Moctar (2025) extended the traditional Abkowitz model to a four-quadrant formulation that accounts for both forward and reverse motions, incorporating propeller revolutions as an active control input. Such studies demonstrate the robustness and precision achievable with physics-grounded

* Corresponding author.
** Corresponding author.
E-mail addresses: huanhuan.li@soton.ac.uk (H. Li), z.yang@ljmu.ac.uk (Z. Yang).

https://doi.org/10.1016/j.oceaneng.2026.128869
Received 8 June 2026; Received in revised form 17 September 2026; Accepted 13 October 2026
Available online 17 October 2026
0029-8018/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

models when supported by extensive hydrodynamic testing or high-fidelity numerical simulation.

Despite their interpretability, the development of physics-based models remains time-intensive and resource-demanding. Each new vessel requires a set of experimental or numerical tests, which limits transferability and scalability required by today's global ship fleet's fast developments. Moreover, these models are typically tuned for calm-water conditions, whereas real-world manoeuvring is often influenced by waves, currents, and wind. To address this, many researchers have incorporated external excitations within the hydrodynamic framework. For instance, Sarigul et al. (2025) estimated manoeuvring performance in regular waves through a system-based approach combining captive tests with wave-induced force models. Other efforts, such as Wang et al. (2025a); (2025b); (2025c) expanded the physical analysis to planning hulls undergoing turning manoeuvres. System identification has also evolved significantly, integrating both model-based and data-assisted perspectives. Luo and Li (2017) employed free-running model tests of the KVLCC2 tanker and extracted hydrodynamic coefficients using support vector machines, demonstrating that modern learning tools can improve parameter estimation within a physics-based framework. Likewise, Xue et al. (2020) applied Bayesian reasoning to estimate 4-DoF hydrodynamic derivatives with uncertainty quantification. These trends reveal that while physics-based models provide insight and consistency, they face limitations in adaptability, motivating a growing interest in data-driven alternatives that can automatically learn complex behaviour directly from observations.

In recent years, data-driven approaches (Guo et al., 2025), also known as non-parametric modelling, have emerged as powerful complements to traditional physics-based models. Such methods eliminate the need for the explicit derivation of hydrodynamic coefficients and instead infer nonlinear relationships directly from measured or simulated data. The major advantage lies in their ability to inherently account for environmental effects, such as waves, currents, and wind, since these factors are naturally embedded in the training data. Consequently, data-driven models often capture a vessel's realistic response under operational conditions without explicit disturbance modelling. Early examples include He et al. (2022), who developed a non-parametric manoeuvring model for the KVLCC vessel using a fully connected neural network trained on free-running model test data. Building on this line of research, Zhou et al. (2023) introduced a hybrid model that employed recurrent long short-term memory (R-LSTM) networks to represent the manoeuvring behaviour of both KCS and ONR benchmark ships, later extending it to handle regular wave conditions through a sliding time-window technique for online learning. In a relevant study, Liu et al. (2024a); (2024b) adopted an Echo State Network (ESN) for the KVLCC and KCS, demonstrating strong prediction capability and faster training compared with conventional deep neural networks.

Following these advances, research attention has been increasingly shifted toward real-time and small-sample learning, addressing practical limitations in data availability and computational demand. Zhou et al. (2024, 2025) developed Bi-LSTM-based hybrid models capable of predicting full-scale ship motions in real time under random rudder actions, representing a step toward onboard deployment. Meng et al. (2025) proposed a pre-training and ensemble learning strategy for multi-step online prediction of 4-DoF manoeuvres, while Tian et al. (2025) introduced a hybrid model designed to handle disturbed water conditions with incremental online learning. Similarly, Zhang et al. (2025) employed a mutual information-based feature selection and sparse regression framework to identify optimal structures from small datasets, improving both efficiency and interpretability. Collectively, these studies underline a growing consensus: modern neural and recurrent networks can achieve high accuracy and adaptability, but their performance is closely tied to data quality and model design.

Parallel to deep learning, probabilistic and kernel-based models have offered a complementary direction with strong theoretical guarantees and uncertainty quantification. Gaussian Process Regression (GPR), in particular, has gained traction due to its capability to model nonlinear relationships while providing confidence intervals for predictions. A series of studies by Ouyang and colleagues (Ouyang et al., 2022, 2023a-c) refined the GPR framework for ship manoeuvring, introducing local and adaptive hybrid-kernel models to improve flexibility and robustness. Liu et al. (2023) combined GPR with wavelet-based denoising, effectively mitigating measurement noise in training data. These probabilistic approaches are particularly valuable in cases where data are sparse or noisy, providing interpretable results with quantified uncertainty, an advantage over opaque deep neural models.

Support Vector Regression (SVR) and sparse-learning methods remain widely used for system identification and low-data scenarios. The earlier application of SVMs by Luo and Li (2017) has motivated several subsequent studies. For instance, Zhang et al. (2022) applied ν-SVR with random excitation to better capture manoeuvring dynamics, while Yuan et al. (2025) developed an Autoencoder-Multivariate SVR framework (AE-MSVR) to model the coupled sway-yaw behaviour more accurately. These models, though simpler than neural networks, exhibit strong generalisation capabilities and require substantially fewer data samples, making them suitable for practical implementation on existing ships with limited operational data.

Another promising line of research leverages Dynamic Mode Decomposition (DMD) for reduced-order, interpretable representations of ship motions. Diez et al. (2021) pioneered the use of DMD to analyse course-keeping performance in irregular waves, later extending this work to manoeuvring in waves. Chen et al. (2023a, 2023b) further developed DMD and higher-order DMD variants for time-series prediction of manoeuvring trajectories and course-keeping under wave disturbance. These techniques highlight the potential of reduced-order models to capture essential dynamics while maintaining computational efficiency and physical interpretability.

Structured information can be incorporated into data-driven models at the levels of data representation, model formulation, and parameter identification. In a broader engineering context, Jiang et al. (2026) used hierarchical clustering to differentiate the contributions of low-fidelity samples within a Gaussian-process-based aerodynamic prediction framework, illustrating how the internal organisation of training data can influence model performance. In ship manoeuvring applications, Atasayan et al. (2024) applied an offline grey-box procedure to calibrate the parameters of conventional manoeuvring models, retaining an explicit connection between system identification and established motion equations. Zhou et al. (2025), in turn, combined physical and data-derived information for time-series prediction of ship manoeuvring responses in regular waves. Related developments include the incorporation of prior hydrodynamic knowledge into least-squares support vector machines (Chen et al., 2022) and the generation of physically consistent synthetic data through virtual captive tests (Alexandersson et al., 2025). These studies show that predictive learning can benefit from structure introduced through the data, the governing model, or the identification procedure. Nevertheless, learning a probabilistic graphical structure that reveals temporal relationships among continuous ship-motion, acceleration, and control variables has received comparatively limited attention.

Overall, the evolution of data-driven manoeuvring models shows a clear trend toward greater realism, robustness, and adaptability. The field is moving from static, calm-water identification toward continuous learning systems capable of handling complex sea states and full-scale operations. Non-parametric models such as neural networks and GPR have proven capable of capturing highly nonlinear and time-dependent relationships, while hybrid and grey-box approaches ensure physical coherence and interpretability. Uncertainty-aware and reduced-order formulations add further value by supporting real-time decision-making in autonomous navigation systems.

Dynamic Bayesian Networks (DBN) extend conventional Bayesian Networks (BN) by explicitly representing temporal dependencies among variables, enabling probabilistic reasoning over dynamic systems. In

maritime and offshore domains, DBN have been widely adopted in risk assessment and decision-support contexts, including offshore operational risk analysis (Dong et al., 2026), Arctic navigational safety (Gu et al., 2026; Liu et al., 2024a, 2024b), emergency management in deep-water drilling (Wu et al., 2026), port state control dynamics (Fan et al., 2025), dynamic resilience assessment for port systems (Wang et al., 2025a, 2025b, 2025c), shipping risk performance (Zou et al., 2026), intention modelling for autonomous collision avoidance (Rothmund et al., 2022), ship machinery fault diagnosis (Han et al., 2024), and reliability assessment for marine mechanical equipment (Li et al., 2026). These studies, DBN in particular, primarily employ DBN to capture system-level risk structures and their temporal evolution, typically using a limited number of discrete time slices to represent operational dynamics. In parallel, data-driven DBN applications have gained increasing attention in maritime research, particularly in studies leveraging historical observational data. Representative examples include ship-to-ship collision risk modelling (Celik et al., 2025; Çelik et al., 2026) and the temporal evolution of emergency scenarios in Arctic maritime accidents (Wang et al., 2025a, 2025b, 2025c). Most of these data-driven studies adopt a discrete DBN formulation, in which continuous variables are first discretized and system dynamics are governed by a predefined Transition Probability Table (TPT). While effective for coarse-grained risk analysis, such discretisation may lead to information loss and limit the model's ability to capture fine-scale temporal variability inherent in high-frequency maritime operational data.

By contrast, continuous DBN formulations model inter-temporal dependencies using parametric probabilistic relationships, commonly Gaussian conditional distributions, rather than discrete TPT, offering greater fidelity for continuous time-series modelling. However, applications of continuous DBN in maritime contexts remain relatively sparse. Recent methodological advances have begun to address this gap. For example, Scutari et al. (2024) extended conventional BN learning to time-series data within the bnlearn framework, enabling DBN inference with continuous variables, and demonstrated its applicability using infodemiology data. Similarly, Zhang et al. (2024) employed the same library for solar power generation forecasting using a DBN with a one-step time lag. Alternative approaches have also been proposed, such as a model tree DBN for piecewise forecasting of nonlinear time series (Quesada et al., 2022). Despite these advances, the potential of DBN for modelling and predicting ship manoeuvring dynamics, characterised by strong temporal dependence, nonlinearity, and continuous state evolution, has not yet been systematically explored. For instance, Gil et al. (2025) recently applied a BN-based framework to model ship responses during evasive manoeuvres, in which the response variables are discretized and predicted in terms of probabilistic outcome intervals rather than continuous temporal trajectories. This gap motivates further investigation into continuous, data-driven DBN frameworks tailored to high-resolution ship motion and operational datasets.

In this context, the present study develops a data-driven manoeuvring model of the KCS based on a DBN. The method advances recent efforts to combine probabilistic representations with dynamic learning by modelling not only the evolution of the motion states but also the conditional relationships that govern their interaction with control inputs. To the authors' knowledge, this study represents the first application of a continuous DBN framework to recursive ship manoeuvring motion prediction. The proposed linear-Gaussian DBN employs a time-lagged structure in which the motion states, control input, and acceleration components from previous time slices are used to predict the acceleration components at the subsequent time step. Unlike deterministic neural formulations, the DBN model preserves an explicit and interpretable structure in which temporal dependencies and influencing factors are represented directly. A physics-guided surge augmentation is further incorporated to improve the recursive prediction of surge dynamics. The turning and zig-zag manoeuvre analyses are used to evaluate the capability of the resulting model to reconstruct the vessel trajectory and associated motion states.

The main contributions of this paper are listed below.

(1) A linear-Gaussian DBN framework is developed for data-driven modelling and recursive prediction of ship manoeuvring motions, providing a probabilistic and interpretable alternative to conventional black-box learning approaches.
(2) The DBN identifies interpretable temporal dependencies among the surge, sway, and yaw motions and the rudder input within a lagged structure, while a physics-guided augmentation is incorporated to improve the prediction of surge dynamics.
(3) A time-lagged prediction mechanism is introduced in which the motion states, control input, and acceleration components from the three preceding time slices are used to predict the acceleration components at the subsequent time step. The predicted accelerations are then recursively integrated to reconstruct the vessel's motion response.
(4) Validation using KCS zig-zag and turning manoeuvres demonstrates that the DBN provides competitive recursive prediction performance compared with the LSTM and GPR benchmark models, with the relative accuracy varying across the investigated motion variables. The DBN requires substantially lower training time than the LSTM, while its mean prediction time per step remains below the experimental sampling interval.

The remainder of the paper is organised as follows. Section 2 outlines the proposed methodology, beginning with the problem formulation and proceeding through the DBN-based identification framework to its application in manoeuvring motion prediction. Section 3 describes the DBN method in detail, including the dataset used and the accuracy analyses conducted to determine the final model structure. Section 4 presents the prediction results obtained from the DBN and compares them with those from LSTM and GPR models and the experimental data, offering further discussion and insights. Finally, Section 5 summarises the key findings and highlights the study's limitations as well as potential directions for future work.

2. Methodology

Initially, the typical ship manoeuvring motions are described and reformulated for data-driven modelling in Section 2.1. The principles of Bayesian reasoning, along with their extension to the temporal domain through DBN modelling, are introduced in Section 2.2 as the underlying system identification framework. Section 2.3 details the LSTM and GPR methods employed in this study as comparative data-driven approaches. Finally, Section 2.4 presents the methodology of DBN model identification and its implementation in recursive motion prediction.

2.1. Problem description

In calm-water conditions, ship manoeuvring motions are typically modelled using 3-DoF, corresponding to surge, sway, and yaw motions. Fig. 1 illustrates the two coordinate systems employed: the earth-fixed frame (x_0, y_0, z_0) and the body-fixed frame (x, y, z). In the body-fixed frame at the midship, u and v denote the surge (forward) and sway (lateral) speeds, respectively, and r is the yaw rate. The control input δ is the rudder angle; β is the ship's drift angle; and ψ is the ship's heading angle. G denotes the ship's centre of gravity.

The ship's manoeuvring dynamics are formulated in the body-fixed coordinate frame, as given by Eq. (1) (Abkowitz, 1964).

$$\begin{cases} m(\dot{u} - vr - x_G r^2) = X \\ m(\dot{v} + ur + x_G \dot{r}) = Y \\ I_{ZZ}\dot{r} + mx_G(\dot{v} + ur) = N \end{cases} \tag{1}$$

Here, m denotes the ship's mass, I_{ZZ} is the moment of inertia about the z-axis, and x_G expresses the longitudinal position of the centre of gravity

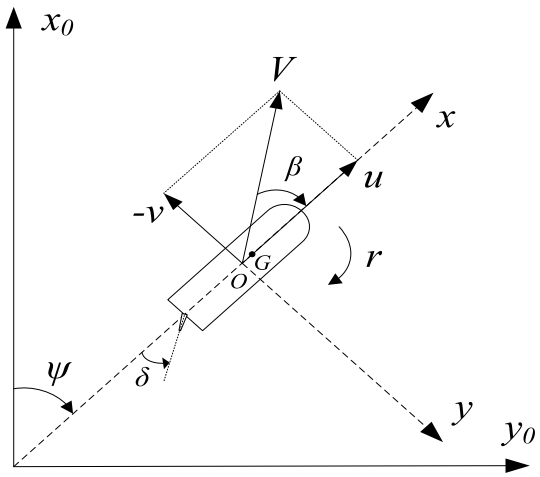


Fig. 1. Description of coordinate systems for ship manoeuvring motions.

in the body-fixed frame. X and Y are the longitudinal and lateral forces acting on the hull, respectively, N indicates the yaw moment. Following Abkowitz (1964), and as formulated in Eq. (2), the hydrodynamic forces and moments acting on the ship are represented as nonlinear functions of the velocity components and the rudder angle.

$$\begin{bmatrix} m - X_{\dot{u}} & 0 & 0 \\ 0 & m - Y_{\dot{v}} & mx_G - Y_{\dot{r}} \\ 0 & mx_G - N_{\dot{v}} & I_{ZZ} - N_{\dot{r}} \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} f_x(u,v,r,\delta) \\ f_y(u,v,r,\delta) \\ f_z(u,v,r,\delta) \end{bmatrix} \tag{2}$$

Here, $X_{\dot{u}}$, $Y_{\dot{v}}$, $Y_{\dot{r}}$, $N_{\dot{v}}$, $N_{\dot{r}}$ represent the inertial hydrodynamic derivatives, while f_x, f_y, and f_z are nonlinear functions of u, v, r, and δ.

After rearranging the equation so that the acceleration terms remain on the left-hand side, the ship-dependent matrices are collected on the right-hand side. The right-hand side can then be written as nonlinear functions of the motion and control variables to be regressed against the acceleration terms as in Eq. (3).

$$\begin{cases} \dot{u} = s_x(u,v,r,\delta) \\ \dot{v} = s_y(u,v,r,\delta) \\ \dot{r} = s_z(u,v,r,\delta) \end{cases} \tag{3}$$

In conventional black-box modelling, the motion and control variables are commonly treated as inputs, $\boldsymbol{X_C} = [u,v,r,\delta]^T$, while the acceleration components are treated as outputs, $\boldsymbol{Y_C} = [\dot{u},\dot{v},\dot{r}]^T$. As indicated by Eq. (3), the functions S_x, S_y, and S_z describe the coupled dependence of the acceleration components on the surge velocity, sway velocity, yaw rate, and rudder input. Therefore, an appropriate data-driven modelling framework should account not only for the input-output relationships but also for the temporal dependencies among the motion, control, and acceleration variables. This requirement motivates the time-lagged dependency-modelling framework introduced in the following section.

2.2. System identification algorithm

A Bayesian Network is represented by a Directed Acyclic Graph ($\mathcal{G}$) which describes the conditional dependencies between a set of variables $X = \{X_1,\ldots,X_N\}$. The joint distribution can be written as Eq. (4) below (Scutari, 2010).

$$P(X) = \prod_{i=1}^{N} P(X_i|\pi(X_i)) \tag{4}$$

where $\pi(X_i)$ represents the parent set of X_i. Each node in the graph $\mathcal{G}$ represents a variable in X, while the arcs in $\mathcal{G}$ specify the dependencies among these variables. When all variables are continuous, as in the case of acceleration terms, velocity components in the dataset, it is common to assume that X follows a multivariate normal distribution. Consequently, the local distributions $P(X_i|\pi(X_i))$ take the form of linear regression models, where the node X_i serves as the response variable and its parent nodes $\pi(X_i)$ act as predictors (Scutari et al., 2024).

$$X_i = \mu_i + \pi(X_i)\beta_i + \varepsilon_i \tag{5}$$

Here, the μ_i denotes the intercept term in the local linear regression, capturing the mean of X_i in the absence of parental influence, the coefficient β_i represents the regression parameter associated with the parent nodes $\pi(X_i)$, while the error terms ε_i are assumed to be independent and normally distributed with variance σ_i^2, specific to each variable X_i. In cases where a node has no parents, the corresponding equation reduces to a model containing only the intercept and the error term. This static formulation of BN can be expanded in temporal space to obtain the dynamic interpretation as a linear-Gaussian DBN, following the strategies below.

(1) To model temporal dynamics, augment $\mathcal{G}$ with two nodes for each variable X_i, one representing the observation at the current time t, denoted X_i^t, and one representing the observation at the next time step $t+1$, denoted X_i^{t+1} or the previous time step as X_i^{t-1}.

Eq. (3) expresses the underlying continuous-time physical relationship, according to which the instantaneous accelerations depend on the vessel motion states and rudder input. The DBN implementation represents this relationship through a discrete-time one-step-ahead formulation. Specifically, the acceleration associated with the transition from $(t-1)$ to (t) is estimated using the motion, control and acceleration history available up to $(t-1)$:

$$\widehat{\boldsymbol{\nu}}_t = F(\boldsymbol{z}_{t-1}, \boldsymbol{z}_{t-2}, \boldsymbol{z}_{t-3}) \tag{6}$$

where $\nu_t = [u_t, v_t, r_t]^{\mathrm{T}}$ and $z_t = [u_t, v_t, r_t, \delta_t, \dot{u}_t, \dot{v}_t, \dot{r}_t]^{\mathrm{T}}$. The resulting acceleration estimate is subsequently used to update the vessel motion state through first-order time integration;

$$\widehat{\boldsymbol{\nu}}_t = \widehat{\boldsymbol{\nu}_{t-1}} + \Delta t\, \widehat{\dot{\boldsymbol{\nu}}}_t \tag{7}$$

Accordingly, the DBN does not disregard the physical dependence described by Eq. (3). Rather, the dependence is approximated through time-lagged statistical relationships obtained from the recent motion and control history.

(2) To maintain an unambiguous temporal interpretation, the DBN was formulated using directed inter-slice dependencies, while connections among variables within the same time slice were excluded. This restriction was adopted as a modelling choice to preserve temporal precedence, avoid imposing an arbitrary causal ordering among contemporaneous variables, and maintain a Granger-style (Bressler and Seth, 2011) predictive interpretation in which past information is used to estimate a subsequent response. It should not be interpreted as a general requirement of the Granger-causality framework, which does not categorically prohibit contemporaneous statistical dependencies.

Unlike static modelling, DBN can model feedback mechanisms by permitting mutual dependencies between variables at different time points. The learning process of a DBN involves two main stages: structure learning and parameter estimation. Given a dataset D and a parameter set Θ, the Bayesian formulation of DBN learning can be expressed as (Zhang et al., 2024);

$$P(\mathcal{G}, \Theta|D) = P(\mathcal{G}|D) \cdot P(\Theta|\mathcal{G}, D) \tag{8}$$

where $P(\mathcal{G}|D)$ corresponds to structure learning and $P(\Theta|\mathcal{G}, D)$ represents parameter learning. The first step identifies the optimal directed acyclic graph structure that best captures the dependencies among

variables, while the second estimates the parameters of the conditional distributions defined by that structure.

Score-based search methods are commonly used for learning DBN structures. In *bnlearn* library (Scutari, 2010), two main options are available for time-series models: tabu search and hill-climbing. Tabu search reduces the risk of getting trapped in local optima but is computationally more demanding. Hill-climbing, in contrast, is simpler, more efficient, and widely adopted in continuous-variable DBN applications. For these reasons, hill-climbing (Cooper and Herskovits, 1992) is employed with the scoring function given in Eq. (9).

$$PL(D,\omega) = log\, L(D) - \omega\, log(n)p \tag{9}$$

where $PL(D,\omega)$ denotes the penalised log-likelihood score used to evaluate a candidate DBN structure, and $L(D)$ is the likelihood of the observed dataset D under that candidate structure. Furthermore, p is the number of parameters in the DBN, n indicates the number of samples in the training dataset, and ω denotes a penalty coefficient that controls model sparsity. The intercepts (μ_i) and regression coefficients (β_i) for each node are then estimated using maximum likelihood. For reference, $\omega = 1$ corresponds to the Bayesian Information Criterion (BIC). Larger penalty values discourage the inclusion of unnecessary arcs, favouring simpler and more interpretable models without compromising predictive accuracy.

2.3. Alternative models for comparison

In addition to the proposed DBN, two established data-driven approaches, LSTM and GPR, are employed as benchmark models. These methods represent two complementary modelling paradigms. LSTM provides a deep recurrent architecture capable of learning nonlinear temporal patterns from sequential data, whereas GPR offers a kernel-based probabilistic formulation that is particularly suitable for regression using relatively limited datasets. Their inclusion therefore enables the proposed DBN to be evaluated against both a recurrent neural-network model and a probabilistic regression model.

2.3.1. LSTM model

In this study, alongside the DBN model, the LSTM neural network, a well-established machine learning algorithm introduced by Hochreiter and Schmidhuber (1997), is employed as an alternative system identification approach for comparison and validation purposes. The LSTM architecture is chosen for its demonstrated ability to capture nonlinear temporal dependencies and long-term correlations in dynamic systems, which are essential for accurately representing vessel motion behaviour. Previously, Jiang et al. (2022) applied the LSTM model to predict ship manoeuvring motions based on free-running tests of the KVLCC2 vessel. Following a similar modelling framework, Guo et al. (2024) implemented the LSTM approach to develop a manoeuvring model for an Unmanned Surface Vehicle (USV) and successfully predicted its dynamic responses. More recently, hybrid LSTM approaches have also emerged. For example, Zhou et al. (2025) integrated a physics-based model with the LSTM algorithm to predict the motions of the ONRT benchmark ship in regular waves. Although LSTM networks have demonstrated strong predictive capability in ship manoeuvring problems, this study primarily aims at interpretable system identification and explicit representation of time-lagged dependencies in vessel dynamics, for which the DBN framework is more suitable. Therefore, LSTM is employed as a benchmark model for comparative validation rather than as the main modelling approach.

The LSTM layer differs from a standard Recurrent Neural Network (RNN) through the inclusion of a cell state and three types of gates, input, forget, and output gates, that regulate information flow. Each unit computes a weighted combination of the current input and the previous hidden state, determines what portion of information should be retained or discarded, and updates its internal state accordingly. This gating mechanism allows the network to learn how to retain long-term dependencies while suppressing redundant or noisy information. Mathematically, the internal operations of an LSTM cell are demonstrated in Fig. 2 and can be expressed as (del Águila Ferrandis et al., 2021);

$$f_t = \sigma\left(\boldsymbol{W_f}[h_{t-1}, x_t] + \boldsymbol{b_f}\right),$$

$$i_t = \sigma(\boldsymbol{W_i}[h_{t-1}, x_t] + \boldsymbol{b_i}),$$

$$\widetilde{c_t} = tan\, h(\boldsymbol{W_c}[h_{t-1}, x_t] + \boldsymbol{b_c}),$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c_t},$$

$$o_t = \sigma(\boldsymbol{W_o}[h_{t-1}, x_t] + \boldsymbol{b_o}), h_t = o_t \odot tan\, h(c_t) \tag{10}$$

where x_t represents the input vector at time t, h_t is the hidden state, c_t indicates the cell state, $\widetilde{c_t}$ denotes the candidate cell state, and σ expresses the sigmoid activation function. The gates (f_t, i_t, o_t) control which components of the memory are updated or exposed, ensuring the model can selectively remember or forget temporal information as needed. The operator, $\odot$, denotes the element-wise multiplication. The $\boldsymbol{W}$ and $\boldsymbol{b}$ represent the weight matrix and bias vector, respectively. This internal regulation is particularly advantageous for capturing manoeuvring dynamics that exhibit delayed hydrodynamic responses or coupling effects among surge, sway, and yaw motions (Sun et al., 2022).

The dataset used in this study consisted of four time-varying motion and control variables collectively represented as the input vector $\boldsymbol{X_C}$. The corresponding target outputs are the derivatives of these quantities, $\boldsymbol{Y_C}$, which reflect the system's dynamic responses. Each input and output variable is independently normalised using a MaxAbsScaler fitted exclusively to the training data to ensure numerical stability and to prevent variables with larger magnitudes from dominating the learning process. A sliding-window approach with a sequence length of ten time steps is then employed to transform the continuous time series into overlapping temporal segments. This formulation enables the LSTM model to observe both the historical evolution and the intrinsic dependencies among the motion and control variables when predicting the next-step accelerations.

The LSTM model is implemented using PyTorch and comprises two stacked LSTM layers followed by a fully connected layer that maps the hidden representation to the three output variables. The input size is four, while the output size is three. Each LSTM layer contained 200 hidden units, a configuration chosen to provide adequate representational power for learning the nonlinear relationships while maintaining manageable model complexity. Stacking two recurrent layers enables the network to form hierarchical temporal abstractions, where the first layer captures local variations and the second layer integrates broader dynamic trends, enhancing its ability to model manoeuvring sequences of varying duration and amplitude (Guo et al., 2024). The network is trained using the Mean-Squared Error (MSE) loss function, optimised via the AdamW algorithm with a learning rate of 1×10^{-4}. The AdamW

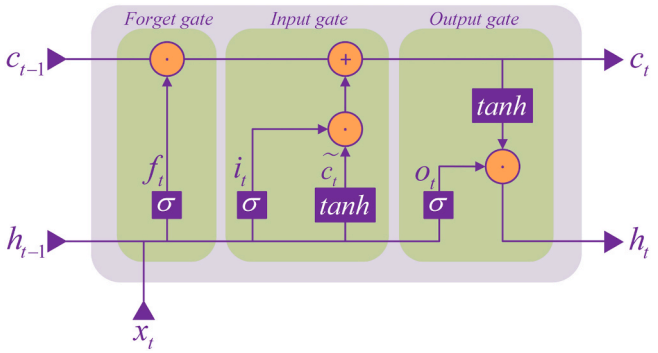


Fig. 2. Cell structure of LSTM.

optimiser is preferred for its decoupled weight-decay regularisation, which improves generalisation in recurrent networks. A batch size of 64 and a total of 50 epochs are adopted after initial experiments demonstrated stable convergence without oscillations or premature stagnation of the training loss. The model weights and fitted scaling parameters are stored after training to enable consistent inference during deployment.

2.3.2. GPR model

GPR was included to complement the recurrent LSTM benchmark with a probabilistic, kernel-based regression method. This provides a more comprehensive assessment of the proposed DBN by allowing it to be compared with both a deep-learning model and an alternative probabilistic approach. The use of GPR for modelling ship manoeuvring dynamics has previously been investigated by Ouyang and Zou (2021). In the present study, however, GPR is implemented as an independent benchmark under the same recursive test conditions used for the DBN and LSTM models.

For a scalar dynamic response (y), the regression problem is expressed as

$$y = f(x) + \varepsilon, \varepsilon \sim \mathcal{N}(0, \sigma_n^2) \tag{11}$$

where $\boldsymbol{x}$ is the model input, $\boldsymbol{f}(\boldsymbol{x})$ is an unknown function, and $\boldsymbol{\varepsilon}$ represents Gaussian observation noise. A Gaussian process prior is assigned to the unknown function as

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \tag{12}$$

where $\boldsymbol{m}(\boldsymbol{x})$ and $\boldsymbol{k}(\boldsymbol{x}, \boldsymbol{x}')$ denote the mean and covariance functions, respectively (Williams and Rasmussen, 2006). Following data standardisation, the prior mean is taken as zero.

The GPR input at time t corresponds to the motion and control vector defined in Section 2.1:

$$X_{C,t} = [u_t, v_t, r_t, \delta_t]^T \tag{13}$$

Since the adopted GPR implementation produces a scalar output, three separate models are trained to estimate the components of the acceleration output vector $Y_{C,t}$:

$$\hat{u}_t = f_u(X_{C,t}), \hat{v}_t = f_v(X_{C,t}), \hat{r}_t = f_r(X_{C,t}) \tag{14}$$

This formulation allows the three acceleration components to be represented by independently optimised covariance functions. The input and output variables are standardised separately using parameters calculated exclusively from the training data, and the same fitted transformations are subsequently applied during testing.

For each output model, the covariance function combines a constant-scaled anisotropic radial basis function (RBF) kernel with a white-noise component. Separate characteristic length scales are assigned to surge velocity, sway velocity, yaw rate, and rudder angle, allowing the covariance function to adapt to the contribution of each input variable. The signal variance, input-specific length scales, and noise level are independently estimated for each acceleration model by maximising the log marginal likelihood of the training observations.

For a new input (x_*), GPR produces a Gaussian predictive distribution:

$$p(f_*|x_*, \boldsymbol{D}, \boldsymbol{y}) = \mathcal{N}(\mu_*, \sigma_*^2) \tag{15}$$

where the posterior mean and variance are

$$\mu_* = k_*^T \boldsymbol{K}_y^{-1} \boldsymbol{y} \tag{16}$$

$$\sigma_*^2 = k(x_*, x_*) - \boldsymbol{k}_*^T \boldsymbol{K}_y^{-1} \boldsymbol{y} \boldsymbol{k}_* \tag{17}$$

Here, $\boldsymbol{D}$ and $\boldsymbol{y}$ are the training inputs and corresponding target observations, $\boldsymbol{k}_*$ contains the covariances between $\boldsymbol{x}_*$ and the training inputs, and $\boldsymbol{K}_y$ is the covariance matrix of the noisy training observations. Although both the posterior mean and variance are available from GPR, the posterior mean is used to conduct the deterministic trajectory comparison with the DBN and LSTM models.

During recursive testing, the predicted acceleration components are converted back to their physical scales and used to advance the vessel states:

$$\widehat{\boldsymbol{\nu}_{t+1}} = \hat{\boldsymbol{\nu}}_t + \Delta t \hat{\dot{\boldsymbol{\nu}}}_t \tag{18}$$

where, $\hat{\dot{\boldsymbol{\nu}}}_t = [\hat{\dot{u}}_t, \hat{\dot{v}}_t, \hat{\dot{r}}_t]^T$.

The resulting motion state and rudder angle form the input for the following prediction step. The heading and position are updated using the common recursive simulation procedure described in Section 2.4. The GPR simulation uses the same initial conditions, sampling interval, numerical integration scheme, rudder-control logic, and test manoeuvre as the DBN and LSTM simulations.

2.4. Application of the identification algorithm in motion prediction

The objective of the DBN-based identification framework developed in this study is to predict future manoeuvring behaviour using past motion-state, control, and acceleration variables. As summarised in Fig. 3, the complete methodology consists of two main stages: identification and prediction.

Stage 1. (Identification/Training): In the identification stage, the DBN is trained as a data-driven predictor using measured time-series signals. The procedure begins with a normalisation step, where all kinematic and control variables and target accelerations are scaled using min–max normalisation with parameters determined exclusively from the training data to ensure numerical stability. Next, a time-lag setup is applied. Each variable is augmented with its history from the previous three time steps, resulting in the conditional formulation $P(\dot{\boldsymbol{v}}_t|\boldsymbol{z}_{t-1}, \boldsymbol{z}_{t-2}, \boldsymbol{z}_{t-3})$, which allows the DBN to encode short-term temporal dependencies characteristic of ship dynamics. Structure learning is then carried out to determine the directed graph, $\mathcal{G}$, that captures the time-lagged dependencies among the motion, control, and acceleration variables. To enhance stability and reduce the sensitivity of the learned structure to measurement noise or local optima, bagging and model averaging are incorporated. In this process, multiple bootstrap subsets of the training data are generated, and a DBN structure is learned independently from each subset. Bagging reduces variance by ensuring that individual anomalous samples do not dominate the learning process, while model averaging produces a consensus structure by aggregating the most consistently occurring edges across all bootstrap models. Following the approach proposed by Scutari et al. (2024), this ensemble-based strategy is employed to improve the stability of the learned network structure and reduce its sensitivity to individual training samples. With the final consensus structure identified, parameter learning is performed to estimate $P(\Theta|\mathcal{G}, D)$, resulting in a DBN model that compactly encodes the conditional dependencies governing the motion-state transitions and acceleration responses.

Stage 2. (Prediction/Inference): with the identified DBN structure and parameters, the network performs recursive inference to predict the ship's manoeuvring response. At each time step, t, the previous three samples of the augmented vector $\boldsymbol{z}$, defined in Section 2.2, are used to predict the acceleration vector $\hat{\dot{\boldsymbol{\nu}}}_t$. These accelerations are propagated forward using the first-order time-integration scheme defined in Eq. (7) to obtain the predicted motion state at the next time step. The corresponding heading angle and position are then updated using the kinematic relationships given in Eq. (19).

$$\begin{aligned} u &= \dot{x}\cos\psi + \dot{y}\sin\psi \\ v &= -\dot{x}\sin\psi + \dot{y}\cos\psi \\ r &= \dot{\psi} \end{aligned} \tag{19}$$

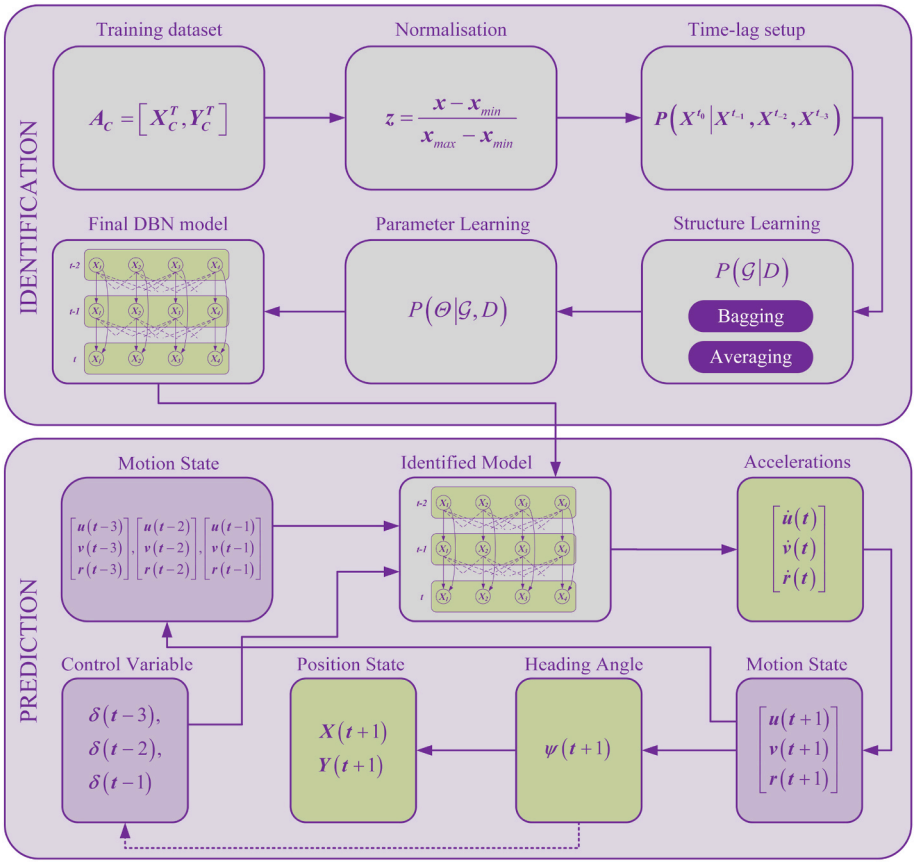


Fig. 3. Flowchart of recursive prediction for ship manoeuvring motions using the DBN model.

For zig-zag test simulations, the control variable, δ, is updated based on the instantaneous heading deviation: the commanded rudder switches between $\pm\delta_0$ when the heading crosses the prescribed deviation thresholds (e.g., $\pm 20^\circ$). Finally, the recursively obtained variables at time t are used to update the augmented vector $\boldsymbol{z}_t$, which, together with its lagged values, provides the input for the next prediction step. Overall, Fig. 3 summarises the DBN identification and recursive prediction workflow for ship manoeuvring motions.

3. Case study: DBN-based system identification of the KCS

The DBN model is implemented using a well-established benchmark vessel, the KRISO Container Ship (KCS), from the SIMMAN workshop series (SIMMAN, 2020). Section 3.1 presents the principal particulars of the ship and describes the training, validation, and testing datasets, including the manoeuvres performed and the sampling interval. Section 3.2 outlines the final DBN configuration, detailing the accuracy analyses and the identification of the final model structure based on these results. Section 3.3 interprets the resulting DBN model through mutual-information analysis of the learned dependencies and discusses the application of physics-guided augmentation to the final DBN.

3.1. Dataset of KCS free-running tests

Free-running tests of the KCS model were conducted by MARIN in deep, calm water, and the resulting datasets were disseminated through the SIMMAN 2020 workshop in South Korea (SIMMAN, 2020). The physical model has a length of 6.07 m, corresponding to 230 m at full scale (scale ratio = 1/37.89). Tests were performed at a model speed of 2.005 m/s, which corresponds to approximately 24 knots at full scale under Froude scaling.

The manoeuvring dataset comprises $10^\circ/10^\circ$ and $20^\circ/20^\circ$ zigzag tests initiated towards the Port Side (PS), a $20^\circ/20^\circ$ zigzag test initiated towards the Starboard Side (SB), and a 35° turning-circle test initiated towards the starboard side. To establish a clear separation between model development and final evaluation, the complete $10^\circ/10^\circ$ PS zigzag dataset was allocated to training. The $20^\circ/20^\circ$ SB zigzag dataset was divided chronologically into training and validation segments. Given the limited amount of available manoeuvring data, the split point was selected to preserve as much of the dataset as possible for training while retaining a distinct final segment for validation. Rather than applying an arbitrary sample-wise split, the validation segment was separated at a point where the rudder returned to its neutral position ($\delta = 0^\circ$) during the latter part of the zigzag test, resulting in 72.7% of the data being allocated to training and the remaining 27.3% to validation. The validation segment was used exclusively for selecting the lag order and penalty coefficient. The $20^\circ/20^\circ$ PS zigzag and 35° SB turning-circle datasets were excluded from all model-selection decisions and reserved exclusively for final testing. The resulting allocation of the KCS free-running manoeuvring datasets is summarised in Table 1. All datasets were sampled at $\Delta t = 0.08123$ s.

Table 1
Allocation of the KCS free-running manoeuvring datasets.

Manoeuvre	Dataset Role
$10^\circ/10^\circ$ zigzag test, starting to port side	Training
$20^\circ/20^\circ$ zigzag test, starting to starboard side	Training (72.7%) and Validation (27.3%)
$20^\circ/20^\circ$ zigzag test, starting to port side	Test
35°turning circle test, starting to starboard side	Test

3.2. Accuracy analysis and structure identification of DBN model

As described in Section 2.2, DBN modelling proceeds in two stages: structure learning followed by parameter learning. In the structure-learning stage, the network topology is inferred by identifying dependencies among variables across multiple time lags. The lag order and penalty coefficient are treated as model hyperparameters and selected using the independent validation segment described in Section 3.1. Because both hyperparameters affect the dimensionality, structural complexity, and predictive performance of the DBN, candidate lag orders from 1 to 6 were evaluated together with the predefined candidate penalty coefficients. A score-based hill-climbing search is used to construct the graph, yielding the set of inter-slice connections that best explains the data under the chosen scoring criterion. In this work, the Bayesian Information Criterion for Gaussian networks (BIC-g) is employed as the scoring function. This criterion assesses each candidate structure by considering both how well it fits the observed data and how complex it is, under the assumption that the variables are normally distributed. To enhance the robustness and stability of the learned structure, a bagging-averaging procedure is incorporated into the structure-learning stage. Specifically, 500 subsampled training datasets are generated, each constructed by randomly sampling 75% of the training dataset without replacement. A separate hill-climbing search is performed for each sample, yielding an ensemble of candidate network structures. The DBN topology is then obtained through model averaging, in which edges that appear consistently across the ensemble are retained according to their empirical frequency. This ensemble-based approach mitigates sensitivity to noise, reduces variance in the learned graph, and provides a more reliable estimate of the underlying dependency structure than a single structure learned from the full dataset. In the parameter-learning stage, using the training dataset as in structure learning, the coefficients of the local conditional distributions are estimated for each node. Given the linear-Gaussian DBN formulation, parameters of the local conditional probability distributions are obtained.

To evaluate the predictive performance of the candidate DBN configurations, a validation-based lag-penalty analysis was conducted using the experimental heading trajectory. Rather than relying exclusively on single-step prediction errors, each candidate model was used to recursively predict the complete heading trajectory over the independent validation segment. This multi-step evaluation accounts for error accumulation and numerical integration effects during recursive prediction. Given the heading sequence $\widehat{\psi}(t)$ predicted by the DBN and the corresponding experimental measurements $\psi(t)$, prediction accuracy was quantified using the adjusted coefficient of determination defined in Eq. (20) (Montgomery et al., 2021).

$$Accuracy = adj.R^2 = 1 - \left(\frac{SS_{res}/(n-p-1)}{SS_{tot}/(n-1)}\right) \tag{20}$$

where $SS_{res} = \sum_{t=1}^{n} (\psi(t) - \widehat{\psi}(t))^2$ is the residual sum of squares, $SS_{tot} = \sum_{t=1}^{n} (\psi(t) - \overline{\psi}(t))^2$ denotes the total variance of the experimental heading, $\overline{\psi}(t)$ is the mean of all observed $\psi(t)$ values, n indicates the number of time steps, p expresses the number of predictors in the regression model. This formulation is identical to the standard adjusted R^2 used in regression analysis, but here it is applied to compare entire simulated trajectories rather than single-step predictions. This trajectory-based measure captures the cumulative prediction errors and numerical integration effects arising during recursive multi-step simulation. To investigate the effects of temporal depth and structural regularisation, the DBN was re-learned for each candidate combination of lag order and penalty coefficient. The following procedure was applied:

1) A DBN structure was learned exclusively from the normalised training dataset for each candidate lag-penalty combination.
2) The corresponding model parameters were estimated using the same training dataset.
3) The total number of arcs and structure-learning time were recorded to quantify structural and computational complexity.
4) The vessel response was recursively predicted over the independent validation segment using the measured rudder angle as an exogenous input.
5) The predicted heading trajectory was compared with the experimental validation trajectory using the adjusted R^2 accuracy measure introduced in Eq. (10).

Table 2 presents the validation heading accuracy, DBN learning time, selected penalty coefficient, and number of retained arcs for candidate lag orders from 1 to 6. For each lag order, the reported accuracy corresponds to the penalty coefficient that produced the highest adjusted R^2 on the independent validation segment. The reported learning time corresponds to the structure- and parameter-learning processes for the same selected configuration. The one-lag model produced an unstable recursive response and therefore did not yield a valid trajectory-level accuracy. The validation heading accuracy increased from 0.977 for the two-lag model to 0.997 for the three-lag model and remained effectively unchanged for higher lag orders. In contrast, the computational cost and network complexity generally increased with temporal depth: the learning time increased from 0.78 s for the three-lag model to 2.22 s for the six-lag model, while the number of retained arcs increased from 150 to 241. Therefore, the three-lag model was selected as the preferred configuration among those achieving the maximum reported validation accuracy, since increasing the lag order beyond three provided no appreciable improvement in heading prediction while increasing the structural and computational complexity. All computations were performed using an Intel Core i7-7700HQ CPU at 2.80 GHz and 16 GB of RAM.

Fig. 4 further illustrates the effect of the penalty coefficient on validation heading accuracy and the number of retained arcs for the selected three-lag configuration. At low penalty values, the learned networks contained a relatively large number of arcs, indicating greater structural complexity. Increasing the penalty progressively reduced the number of arcs by discouraging the inclusion of weak or redundant dependencies. However, excessive regularisation resulted in the removal of important temporal relationships and a corresponding reduction in validation accuracy. The highest validation heading accuracy, together with a relatively compact network structure, was obtained at $\omega = 0.075$.

Based on this validation analysis, a lag order of three and a penalty coefficient of $\omega = 0.075$ were retained for the final DBN configuration. Both hyperparameters were fixed before the reserved test manoeuvres were used for performance evaluation. The resulting final DBN structure is presented in Fig. 5.

3.3. Model interpretation and physics-guided augmentation

In order to understand the internal structure learned by the DBN and to assess whether the inferred dependencies align with the known ship dynamics, an arc-strength analysis has been conducted using Gaussian Mutual Information (MI). This procedure quantifies the statistical rele-

Table 2
Effect of lag order on the validation accuracy, computational cost, and structural complexity of the DBN.

Lag Order	Penalty [ω]	Accuracy [*adj.* R^2]	Structure and Parameter Learning Time [*s*]	Number of Arcs
1	0.001	N/A (diverged)	0.10	49
2	0.020	0.977	0.34	97
3	**0.075**	**0.997**	**0.78**	**150**
4	0.3	0.997	0.79	160
5	0.4	0.997	1.23	193
6	0.4	0.997	2.22	241

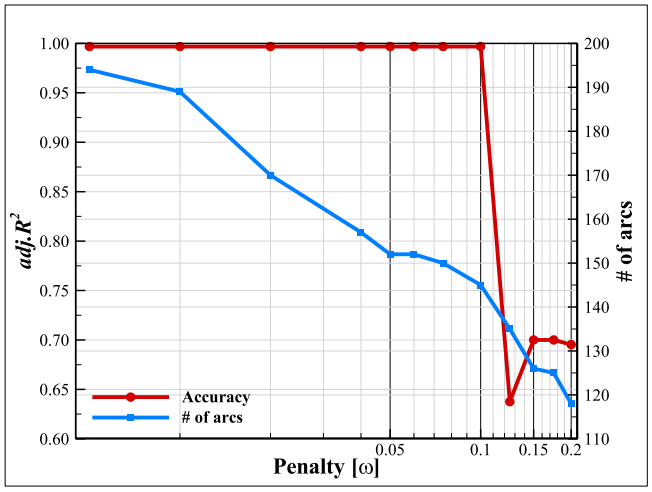


Fig. 4. Effect of the penalty coefficient on validation heading accuracy and network complexity for the selected lag order of three.

vance of each directed edge (arc) in the network and therefore enables a principled interpretation of which variables substantially influence the system at different time lags. Table 3 shows the mutual information values between each acceleration component at time t and its respective parent variables. The strongest dependency among the parent-child relationships associated with the current-time acceleration nodes is $\dot{v}_{t-2} \rightarrow \dot{r}_t$ with MI = 0.9877, highlighting the tight interaction between sway and yaw accelerations, consistent with classical 3-DoF manoeuvring theory. Multiple additional arcs with substantial information gain involve higher-order lags of sway and yaw variables, indicating that the learned DBN represents the inherent coupling in lateral dynamics. The influence of the rudder angle (δ) appears primarily in arcs associated with sway and yaw states, both as direct and indirect contributors, but rarely with dominant strength. This observation agrees with the physical understanding that rudder forces act mainly in the lateral plane.

The mutual information values indicate: (1) strong sway influence on yaw (lateral speed and acceleration terms), (2) detectable, though less dominant, rudder contributions, (3) a multi-lag dependency structure consistent with the temporal characteristics of manoeuvring dynamics.

In contrast, the most striking outcome concerns the surge acceleration node, $\dot{u}_t$. The arcs entering this node exhibit extremely low mutual information values. The strongest parent relationships identified were $\dot{r}_{t-3} \rightarrow \dot{u}_t$, $\dot{r}_{t-2} \rightarrow \dot{u}_t$, and $\dot{r}_{t-1} \rightarrow \dot{u}_t$ with MI values 0.091, 0.071, and 0.067, respectively. Surge acceleration exhibits only weak statistical dependencies, with all MI values falling below 0.1. All remaining parent influences, including surge states, exhibit very low MI values. Notably, there is no direct arc from the rudder angle to $\dot{u}_t$ at any time lag; nor are there strong indirect pathways statistically linking the rudder angle to surge acceleration. From a hydrodynamic perspective, this is physically reasonable. Surge acceleration in low-speed manoeuvres is governed predominantly by thrust and hull resistance, with minimal rudder-induced longitudinal effects. Therefore, the learned DBN structure represents the surge component through comparatively weak statistical dependencies with the other motion and control variables.

While Table 3 focuses specifically on the parent-child relationships associated with the current-time acceleration nodes $\dot{u}_t$, $\dot{v}_t$, and $\dot{r}_t$, Table 4 provides a complementary global ranking of the strongest MI-based

Table 3
Mutual information between the current-time acceleration nodes $\dot{u}_t$, $\dot{v}_t$, $\dot{r}_t$ and their respective parent variables.

Surge ($\dot{u}_t$)		Sway ($\dot{v}_t$)		Yaw ($\dot{r}_t$)	
Parents	MI	Parents	MI	Parents	MI
$\dot{r}_{t-3}$	0.0910	$\dot{v}_{t-1}$	0.3400	$\dot{v}_{t-2}$	0.9877
$\dot{r}_{t-2}$	0.0710	$\dot{u}_{t-1}$	0.2230	v_{t-3}	0.0750
$\dot{r}_{t-1}$	0.0670	$\dot{u}_{t-3}$	0.1850	v_{t-1}	0.0410
		r_{t-2}	0.0220	r_{t-3}	0.0055
		v_{t-3}	0.0190	u_{t-3}	0.0030
		δ_{t-1}	0.0039	$\dot{v}_{t-3}$	0.0014

Table 4
Mutual information ranking of the strongest dependencies in the DBN.

From	To	MI	Interaction Type
$\dot{v}_{t-2}$	$\dot{r}_t$	0.9877	Sway-yaw coupling
v_{t-3}	$\dot{r}_{t-1}$	0.8789	Sway-yaw coupling
v_{t-2}	$\dot{r}_{t-1}$	0.6933	Sway-yaw coupling
$\dot{u}_{t-2}$	$\dot{r}_{t-1}$	0.5593	Surge-yaw cross-coupling
u_{t-3}	$\dot{v}_{t-2}$	0.5128	Surge-sway cross-coupling
$\dot{v}_{t-1}$	$\dot{v}_t$	0.3400	Temporal continuity in sway acceleration
u_{t-2}	$\dot{v}_{t-1}$	0.3327	Surge-sway cross-coupling
$\dot{u}_{t-3}$	$\dot{v}_{t-2}$	0.2998	Surge-sway cross-coupling
$\dot{u}_{t-3}$	$\dot{v}_{t-1}$	0.2705	Surge-sway cross-coupling
$\dot{u}_{t-1}$	$\dot{v}_t$	0.2230	Surge-sway cross-coupling
r_{t-3}	$\dot{r}_{t-1}$	0.2188	Temporal coupling within the yaw dynamics
$\dot{u}_{t-3}$	$\dot{v}_t$	0.1850	Surge-sway cross-coupling
$\dot{r}_{t-3}$	$\dot{u}_t$	0.0910	Yaw-surge cross-coupling
v_{t-3}	$\dot{r}_t$	0.0750	Sway-yaw coupling
$\dot{r}_{t-2}$	$\dot{u}_t$	0.0710	Yaw-surge cross-coupling
$\dot{r}_{t-1}$	$\dot{u}_t$	0.0670	Yaw-surge cross-coupling

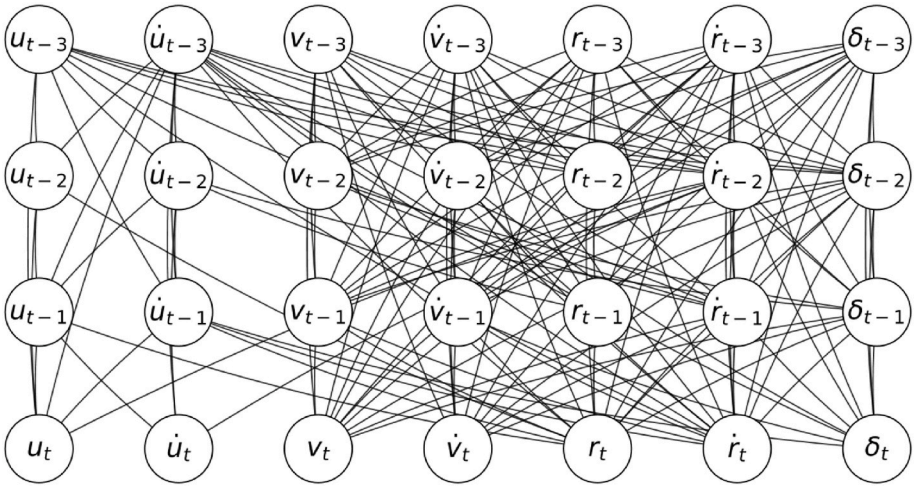


Fig. 5. Network structure of the final DBN model.

dependencies identified across the entire learned DBN structure.

The global MI ranking reveals a markedly non-uniform dependency structure across the three motion components. The yaw-acceleration dynamics ($\dot{r}$) display some of the strongest dependencies, most notably the pronounced sway-yaw coupling represented by $\dot{v}_{t-2} \rightarrow \dot{r}_t$ (MI ≈ 0.99). This reflects the well-known lateral-yaw interaction in ship manoeuvring, where sway-induced forces strongly influence yaw acceleration. In contrast, the surge-acceleration dynamics ($\dot{u}$) exhibit comparatively weaker parental contributions to the current-time surge acceleration, primarily through yaw-to-surge interactions such as $\dot{r}_{t-3} \rightarrow \dot{u}_t$ and $\dot{r}_{t-2} \rightarrow \dot{u}_t$, both with MI values below 0.10. The sway-acceleration component ($\dot{v}$) occupies an intermediate position, governed by a mixture of temporal continuity (e.g., $\dot{v}_{t-1} \rightarrow \dot{v}_t$) and cross-couplings from surge and yaw.

Given the weak statistical dependencies identified for surge acceleration, the DBN alone shows limited capability in reproducing the surge response during recursive simulations. The surge speed does not diverge, but instead tends towards an overly smoothed response. As a result, the DBN captures the general trend of the surge dynamics but does not fully reproduce the variability observed in the experimental data.

To address this, an additional surge term based on a physics-guided surge surrogate is incorporated and combined with the DBN prediction through a convex blending mechanism. Since the augmentation is introduced specifically to address the limited DBN representation of surge dynamics, it is applied only to the surge channel, while the DBN predictions for sway and yaw remain unchanged. A simple, interpretable regression formulation is used to approximate the surge dynamics at time t, as given in Eq. (21).

$$\widehat{\dot{u}_t^{reg}} = f(u_{t-1}, v_{t-1}, r_{t-1}, \delta_{t-1}) + \varepsilon_t \tag{21}$$

The surrogate model uses the following set of physically meaningful explanatory variables: longitudinal speed (u_{t-1}), nonlinear drag term ($u_{t-1}|u_{t-1}|$), sway-yaw coupling term ($v_{t-1}r_{t-1}$), and rudder-induced contribution (δ_{t-1}^2). These terms are consistent with standard low-speed nonlinear manoeuvring models for marine vehicles (Fossen, 2011), but here they serve only as data-calibrated structural guidance. The regression model is therefore formulated as:

$$\widehat{\dot{u}_t^{reg}} = \alpha_0 + \alpha_1 u_{t-1} + \alpha_2 u_{t-1}|u_{t-1}| + \alpha_3 v_{t-1} r_{t-1} + \alpha_4 \delta_{t-1}^2 \tag{22}$$

where $\alpha_0, \ldots, \alpha_4$ are estimated from the training set via least squares. This model provides physics-guided structural information by explicitly representing surge-related terms that exhibit weak statistical dependencies in the learned DBN. The final surge prediction is computed as a convex combination of the surrogate and the DBN prediction, as given in Eq. (23).

$$\hat{u}_t = (1-\eta)\widehat{\dot{u}_t^{reg}} + \eta \widehat{\dot{u}_t^{DBN}} \tag{23}$$

The blending factor was determined through a sensitivity analysis conducted exclusively on the validation segment. The DBN structure-learning penalty was fixed at $\omega = 0.075$, and η was varied from 0 to 1 in increments of 0.1. For each candidate value, a complete recursive validation simulation was performed. Since the augmentation acts directly on the surge channel, the adjusted R^2 calculated between the experimental and predicted surge-velocity trajectories was adopted as the primary selection criterion. The corresponding heading accuracy was also recorded to assess whether changes in the surge formulation affected the overall recursive manoeuvring response. According to Eq. (23), $\eta = 0$ represents the surge-surrogate-only formulation, whereas $\eta = 1$ represents the DBN-only formulation.

As shown in Table 5, the surge-surrogate-only formulation $\eta = 0$ achieved an adjusted R^2 of 0.936, whereas the DBN-only formulation ($\eta = 1$) produced a substantially lower surge accuracy of 0.319. This finding indicates that the physics-guided surrogate accounts for a considerable proportion of the improvement in surge prediction and that the time-lagged linear-Gaussian DBN alone has limited capability to reproduce the surge-velocity variation during recursive simulation. Nevertheless, the highest surge accuracy, 0.958, was obtained using the combined formulation at $\eta = 0.5$, exceeding the accuracy of both individual components. The DBN prediction therefore provides complementary information when combined with the physics-guided surrogate, despite being insufficient as a stand-alone representation of the surge response.

Table 5
Sensitivity of the surge blending factor evaluated on the validation segment.

Blending Factor [η]	Surge Accuracy [$adj.R^2$]	Heading Accuracy [$adj.R^2$]
0.0	0.936	0.988
0.1	0.939	0.988
0.2	0.943	0.988
0.3	0.948	0.988
0.4	0.952	0.988
0.5	**0.958**	**0.989**
0.6	0.956	0.989
0.7	0.947	0.990
0.8	0.890	0.991
0.9	0.636	0.993
1.0	0.319	0.997

The surge accuracy remained comparatively high around the selected configuration, particularly over $\eta = 0.4 - 0.6$, after which it decreased as the DBN contribution became dominant. The selected value of $\eta = 0.5$ lies within this high-accuracy region and assigns equal weights to the two components. Although the heading accuracy increased gradually with η, it remained high for every candidate value and reached 0.989 at the selected configuration. Since the blending mechanism was introduced specifically to improve the surge channel, η was selected according to surge accuracy rather than heading accuracy. Following this validation-based selection, $\eta = 0.5$ was fixed before evaluation on the test manoeuvres, and no further adjustment was performed using the test data.

4. Results and discussions

As highlighted in Section 3.1, the 20 °/20 ° PS zig-zag manoeuvre was reserved exclusively for independent testing and was not used during training, validation, or hyperparameter selection, providing an independent evaluation of the DBN-based system-identification model. At each time step t, the model predicts the acceleration components in surge, sway, and yaw; these are then integrated using the first-order time-integration scheme defined in Eq. (7) to obtain the predicted motion states. The corresponding heading angle and position are updated using the kinematic relationships given in Eq. (19). For the zig-zag protocol, the rudder command follows the standard switching rule based on heading deviations, and the recursively obtained variables are used to update the augmented vector $\boldsymbol{z}_t$, which, together with its lagged values, provides the input for the next prediction step, allowing recursive prediction over the full zig-zag manoeuvre.

To evaluate the accuracy of the prediction models, the Normalised Root Mean Squared Error (NRMSE) is used. NRMSE provides a scale-independent measure of how closely the predicted values follow the reference measurements. The deviation between each predicted value $\hat{y}_i$ and its corresponding measurement y_i is first evaluated, and these squared differences are averaged over the entire sequence. The square root of this average yields the Root Mean Squared Error (RMSE), which represents the characteristic magnitude of the prediction error.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{24}$$

Because the physical scales of the variables differ, this error is then normalised by the numerical span of the measured data, $y_{max} - y_{min}$, yielding the NRMSE as

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}} \tag{25}$$

This normalisation allows direct comparison across different variables such as u, v, r, and ψ, regardless of their units or magnitudes. Lower NRMSE values indicate that the predicted response tracks the measured motion more closely.

The computational performance of the models was evaluated by distinguishing offline training from recursive simulation. Each complete simulation was repeated 30 times on the same computer, equipped with an Intel Core i7-7700HQ processor at 2.80 GHz and 16 GB of RAM, and the mean total simulation time was reported. The mean computation time per step was obtained by dividing the total simulation time by the number of prediction steps.

To provide a broader assessment of the predictive performance of the proposed DBN, the 20°/20° PS zig-zag manoeuvre was simulated using the DBN, LSTM, and GPR models. As shown in Table 6, all three models produce relatively close NRMSE values across the investigated motion variables, indicating generally comparable prediction performance for this test case. The DBN reproduces the investigated motion variables with relatively low normalised errors. A notable result is obtained for the heading response, for which the DBN achieves the lowest NRMSE of 0.0685, compared with 0.0713 for the LSTM and 0.0777 for the GPR. This indicates that the DBN maintains an accurate representation of the heading evolution throughout the recursive zig-zag prediction. Although the relative performance varies across the individual motion variables, the results demonstrate that the DBN provides competitive overall prediction performance while retaining its interpretable time-lagged network structure.

Regarding computational performance, the DBN required 0.78 s for structure and parameter learning, substantially less than the training times of 38.77 s for the LSTM and 10.82 s for the GPR. The mean total simulation times were 5.963 s, 2.488 s, and 2.300 s for the DBN, LSTM, and GPR, respectively, corresponding to mean computation times of 5.315 ms, 2.213 ms, and 2.047 ms per prediction step. Although the DBN required a longer computation time per step than the benchmark models, its mean prediction time remained substantially below the experimental sampling interval of 81.228 ms. The results therefore demonstrate that the DBN can complete each recursive prediction step within the available experimental sampling interval under the tested computational conditions, while its principal computational advantage lies in the substantially lower model-learning time.

Fig. 6 presents a comparison between the predicted and measured time histories of the heading angle during the 20°/20° PS zig-zag manoeuvre. All three models follow the experimental trajectory closely, with generally consistent phase behaviour throughout the oscillatory motion. The DBN prediction remains particularly well aligned with the measured response, maintaining the timing and overall shape of the turning cycles with only small deviations in amplitude. This observation is consistent with the quantitative results in Table 6, where the DBN achieves the lowest heading NRMSE among the three models. Overall, the figure demonstrates that the DBN accurately reproduces the heading evolution of the zig-zag manoeuvre, while the LSTM and GPR predictions also remain close to the experimental response.

Fig. 7 shows the evolution of the surge speed during the zig-zag manoeuvre, comparing the experimental measurements with the predictions obtained from the LSTM, GPR, and DBN models. All three models reproduce the overall evolution of the surge velocity, including the initial reduction in forward speed and the subsequent oscillatory variations throughout the manoeuvre. The predicted responses remain close to one another and generally follow the measured trajectory, although local differences in amplitude are observed at several stages of the manoeuvre. Consistent with the NRMSE values reported in Table 6, the LSTM and GPR exhibit slightly lower surge-velocity errors than the DBN, while the DBN nevertheless maintains close agreement with the experimental response throughout the simulation. These results indicate that the physics-guided surge augmentation enables the DBN to reproduce the principal surge-velocity characteristics during the zig-zag manoeuvre.

Fig. 8 illustrates the predicted and measured sway-velocity histories throughout the zig-zag manoeuvre. All three models reproduce the dominant periodic behaviour of the lateral motion and generally follow the phase and amplitude evolution of the experimental response. The largest discrepancy for the DBN occurs during the first sway-velocity peak, where the predicted amplitude is lower than the measured value. After this initial transient, the DBN follows the subsequent oscillatory response more closely, with its predicted peaks and troughs remaining generally consistent with the experimental measurements. The GPR achieves the lowest sway-velocity NRMSE of 0.0771, while the DBN and LSTM provide comparable values of 0.0940 and 0.0948, respectively. Overall, the results show that the DBN reproduces the principal sway-motion characteristics of the zig-zag manoeuvre, despite the larger deviation observed during the initial response.

Fig. 9 presents the evolution of the yaw rate during the same manoeuvre, comparing the model predictions with the experimental measurements. All three models reproduce the dominant oscillatory behaviour of the measured yaw-rate response, with relatively small differences in phase and amplitude. The largest discrepancies occur during the initial turning cycle, whereas the predictions become more closely aligned with the experimental response over the subsequent cycles. The GPR achieves the lowest yaw-rate NRMSE of 0.0865, followed closely by the DBN and LSTM with values of 0.0914 and 0.0929, respectively. The DBN maintains good agreement with the measured response throughout the manoeuvre, particularly after the initial transient, reproducing both the timing and magnitude of the subsequent yaw-rate oscillations. Overall, the close NRMSE values and trajectory responses indicate comparable yaw-rate prediction performance among the three models for the zig-zag manoeuvre.

Fig. 10 compares the measured and predicted trajectories and motion-state histories for the 35° turning-circle manoeuvre. Both models reproduce the sustained turning response, although differences emerge in their predicted trajectories and individual velocity components. The DBN trajectory remains closer to the experimental path and reproduces the principal turning-circle dimensions more accurately, whereas the LSTM trajectory progressively develops a larger positional deviation. The motion-state histories indicate a mixed comparison: the DBN follows the measured surge velocity and yaw rate more closely, while the LSTM provides a more accurate representation of the sway velocity. These observations are evaluated quantitatively following Fig. 10.

The GPR model was also examined for the turning-circle case but did not provide a stable and sufficiently accurate recursive prediction for this manoeuvre and was therefore not included in the quantitative comparison in Table 7. This behaviour may be associated with the limited manoeuvring diversity of the development data and the challenge of extrapolating from the available zig-zag data to a sustained

Table 6
Prediction errors and computational performance of the LSTM, GPR, and DBN models for the 20°/20° PS zig-zag manoeuvre.

Model	u	v	r	sum	ψ	Training Time [s]	Total Simulation Time [s]	Simulation Time per Step [ms]
LSTM	0.0536	0.0948	0.0929	0.2413	0.0713	38.77	2.488	2.213
GPR	0.0571	0.0771	0.0865	0.2207	0.0777	10.82	2.300	2.047
DBN	0.0636	0.0940	0.0914	0.2490	0.0685	0.78	5.963	5.315

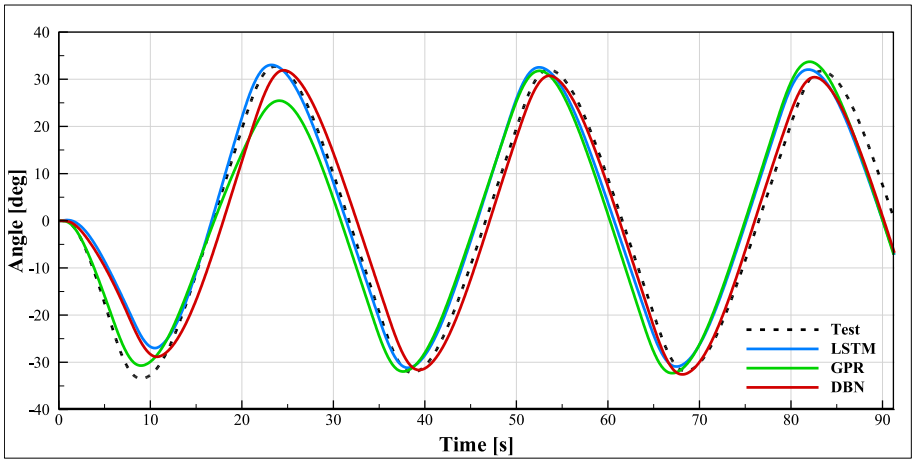


Fig. 6. Comparison of predicted and measured heading responses during the 20°/20° zig-zag manoeuvre initiated towards the port side.

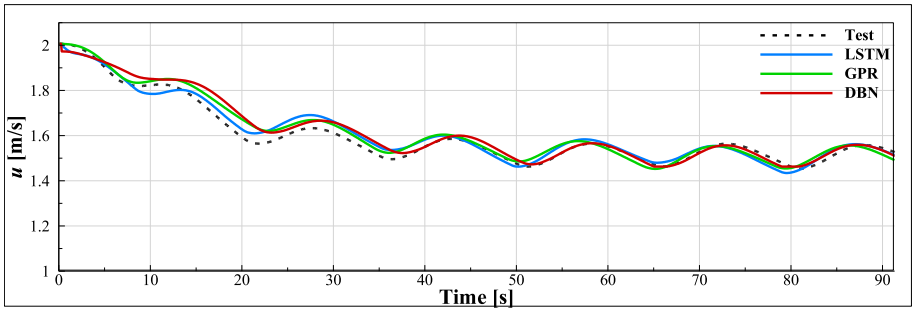


Fig. 7. Comparison of predicted and measured surge-speed responses during the 20°/20° PS zig-zag manoeuvre.

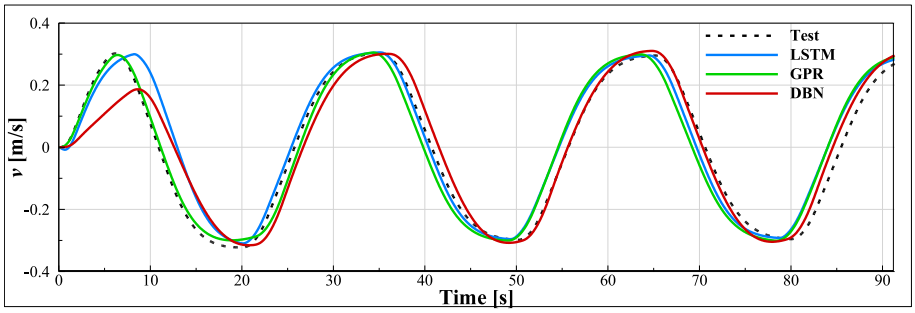


Fig. 8. Comparison of predicted and measured sway-speed responses during the 20°/20° PS zig-zag manoeuvre.

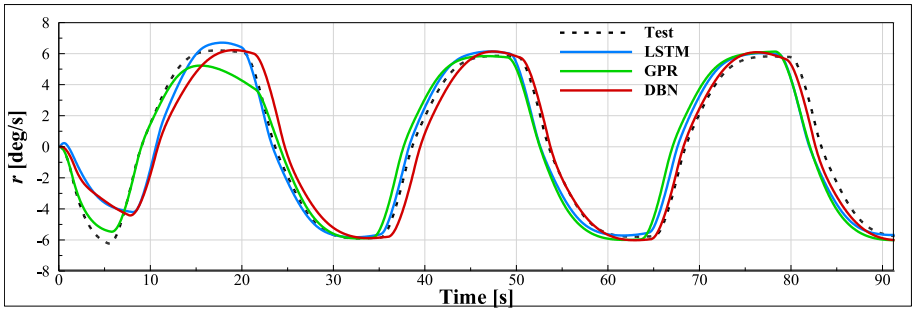


Fig. 9. Comparison of predicted and measured yaw-rate responses during the 20°/20° PS zig-zag manoeuvre.

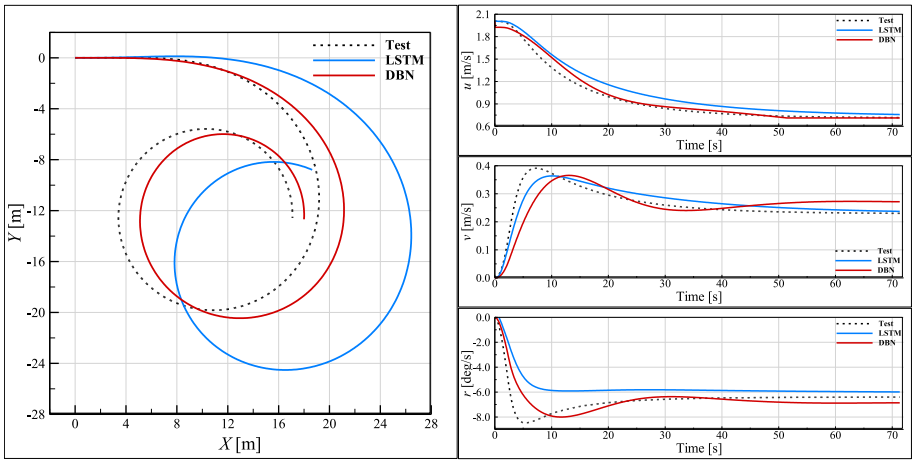


Fig. 10. Comparison of predicted and measured trajectories and motion states during the 35° turning manoeuvre.

Table 7
Quantitative comparison of the DBN and LSTM models for the KCS 35° turning-circle test.

	Test	LSTM	DBN	LSTM Error [%]	DBN Error [%]
***u* [NRMSE]**	-	0.0865	0.0385	-	-
***v* [NRMSE]**	-	0.0651	0.1207	-	-
***r* [NRMSE]**	-	0.1533	0.0809	-	-
sum [NRMSE]	-	0.3049	0.2401	-	-
Advance [m]	18.76	26.01	20.77	38.68	10.72
Transfer [m]	8.16	10.82	9.31	32.52	14.09
Tactical Diameter [m]	19.44	24.14	20.15	24.19	3.68
Final Position Error [m]	-	4.09	0.92	-	-

turning response. Further investigation using a broader training dataset would be required to establish the underlying cause.

The 35° turning-circle dataset was obtained from the KCS free-running model tests and was used exclusively as an independent final test case; it was not included in model training, validation, or hyper-parameter selection. In addition to the visual comparison in Fig. 10, the turning performance was evaluated using the NRMSE values of surge velocity, sway velocity, and yaw rate. The geometric characteristics of the turning circle were quantified using advance and transfer at a heading change of 90° and tactical diameter at a heading change of 180°. Linear interpolation between adjacent time steps was used to determine the positions corresponding to these heading changes. The final position error was calculated as the Euclidean distance between the measured and predicted positions at the end of the common evaluation period. The resulting quantitative comparison is presented in Table 7.

As presented in Table 7, the relative performance of the two models varies among the individual motion components. The DBN achieves lower NRMSE values for surge velocity and yaw rate, decreasing the respective errors from 0.0865 to 0.0385 and from 0.1533 to 0.0809 compared with the LSTM. Conversely, the LSTM produces a lower sway-velocity NRMSE of 0.0651, compared with 0.1207 for the DBN. When the three component-specific NRMSE values are summed, the DBN yields a lower cumulative value of 0.2401, compared with 0.3049 for the LSTM. This cumulative value provides a concise summary of the component-wise normalised errors but is not treated as a separate multivariate error metric.

The differences between the two models are more pronounced in the geometric turning-circle characteristics. The experimental advance, transfer, and tactical diameter are 18.76 m, 8.16 m, and 19.44 m, respectively. The corresponding DBN predictions are 20.77 m, 9.31 m, and 20.15 m, resulting in relative errors of 10.72%, 14.09%, and 3.68%. The LSTM produces larger values of 26.01 m, 10.82 m, and 24.14 m, with respective errors of 38.68%, 32.52%, and 24.19%. The final position error is also lower for the DBN, at 0.92 m, compared with 4.09 m for the LSTM. These results indicate that the DBN provides a more accurate reconstruction of the overall turning-circle geometry, despite the superior LSTM performance for the sway-velocity component.

The generalisation of the present results should be considered within the scope of the available operating conditions. The model was developed and evaluated using KCS free-running tests conducted in deep and calm water at a single speed. Although the zig-zag and turning tests demonstrate predictive capability across different manoeuvres within this domain, further validation is required under environmental disturbances, measurement noise, different speeds and loading conditions, and more diverse control inputs such as random rudder excitation. Validation using additional ship types and full-scale operational data is also required before broader applicability can be established. Future work will therefore examine the robustness and transferability of the proposed DBN using broader model-scale and full-scale datasets under more diverse operating and environmental conditions.

5. Conclusions

This study introduced a DBN-based system-identification framework for predicting ship manoeuvring motions. The approach models time-lagged dependencies among the motion-state, control, and acceleration variables and recursively predicts the acceleration outputs used to reconstruct the vessel motion. The proposed framework employs a data-driven linear-Gaussian DBN with time-lagged dependency learning and physics-guided surge augmentation. Its graphical structure provides an interpretable representation of the learned temporal dependencies among the manoeuvring variables, while the mutual-information analysis enables the relative strength of these dependencies to be examined quantitatively.

The methodology was evaluated using free-running model tests of the KCS, and its predictive performance was compared against LSTM and GPR benchmark models. The three models showed generally comparable prediction performance for the unseen zig-zag manoeuvre, with the relative accuracy varying across the individual motion variables. A notable result was obtained for the heading response, for which the DBN achieved the lowest prediction error among the three models. The DBN also reproduced the surge, sway, and yaw-rate responses with relatively low normalised errors, although the benchmark models achieved lower errors for some individual motion components. From a computational perspective, the DBN required substantially less model-learning time

than both benchmark models, while its mean recursive prediction time per step remained below the experimental sampling interval.

The turning-circle analysis provided further evaluation of the DBN on an unseen manoeuvre with dynamics distinct from the zig-zag test. The relative prediction accuracy again varied across the individual motion components, with the DBN providing lower errors for surge velocity and yaw rate, while the LSTM produced a lower error for sway velocity. More pronounced differences were observed in the reconstructed turning trajectory, where the DBN reproduced the advance, transfer, tactical diameter, and final position more accurately than the LSTM. These results demonstrate the capability of the DBN to reproduce both motion-state responses and the overall turning-circle geometry within the investigated operating conditions.

A further advantage of the DBN formulation lies in its interpretable dependency structure. The mutual-information analysis revealed a clear hierarchy among the learned temporal dependencies, with yaw acceleration exhibiting strong cross-couplings, particularly with sway-related variables, while surge acceleration showed comparatively weak statistical dependencies. This analysis provides insight into the relative importance of the learned relationships and helps identify motion components that are less strongly represented by the data-driven DBN structure. In particular, the weak dependencies associated with surge acceleration motivated the incorporation of the physics-guided surge augmentation used in the final model.

The current implementation is limited to model-scale KCS manoeuvres conducted in deep and calm water at a single speed and using a limited set of manoeuvring cases. Consequently, the present findings should be interpreted within this investigated operating domain. Future work will extend the validation to environmental disturbances, measurement noise, different speeds and loading conditions, and more diverse control inputs, including random rudder excitation. Validation using additional ship types and full-scale operational data will also be required to assess the robustness and transferability of the proposed framework under broader operating conditions.

In summary, the proposed framework combines recursive manoeuvring prediction with an interpretable representation of time-lagged dependencies among ship motion and control variables. The results demonstrate competitive predictive performance across the investigated manoeuvres, substantially lower model-learning time than the benchmark models, and the ability to identify weakly represented dynamics that can be addressed through physics-guided augmentation. These findings demonstrate the potential of the proposed DBN framework as an interpretable data-driven approach for ship manoeuvring modelling and prediction.

CRediT authorship contribution statement

Cihad Çelik: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Huanhuan Li:** Writing – review & editing, Visualization, Validation, Supervision, Project administration, Methodology, Formal analysis, Conceptualization, Resources. **Lu Zou:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Data curation. **Musa Bashir:** Writing – review & editing, Validation, Supervision, Formal analysis. **Zaili Yang:** Writing – review & editing, Visualization, Supervision, Resources, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

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

Acknowledgements

This work is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 864724).

References

Abkowitz, M.A., 1964. *Lectures on ship hydrodynamics-steering and manoeuvrability* (HY-5). https://trid.trb.org/View/159100.

Alexandersson, M., Mao, W., Ringsberg, J.W., Kjellberg, M., 2025. Identification of manoeuvring models for wind-assisted ships with large rudders using virtual captive tests. Int. J. Nav. Archit. Ocean Eng., 100664

Araki, M., Sadat-Hosseini, H., Sanada, Y., Tanimoto, K., Umeda, N., Stern, F., 2012. Estimating maneuvering coefficients using system identification methods with experimental, system-based, and CFD free-running trial data. Ocean. Eng. 51, 63–84. https://doi.org/10.1016/j.oceaneng.2012.05.001.

Atasayan, E., Milanov, E., Alkan, A.D., 2024. Application of an offline grey box method for predicting the manoeuvring performance. Brodogradnja : An International Journal of Naval Architecture and Ocean Engineering for Research and Development 75 (3), 1–36. https://doi.org/10.21278/brod75304.

Bressler, S.L., Seth, A.K., 2011. Wiener–Granger Causality: a well established methodology. Neuroimage 58 (2), 323–329. https://doi.org/10.1016/j.neuroimage.2010.02.059.

Celik, C., Li, H., Liu, J., Bashir, M., Zou, L., Yang, Z., 2025. Data-driven BN and DBN models prediction performance for ship collision risk assessment. In: 2025 8th International Conference on Transportation Information and Safety (ICTIS), pp. 1945–1952. https://doi.org/10.1109/ICTIS68762.2025.11214913.

Çelik, C., Li, H., Liu, J., Bashir, M., Zou, L., Yang, Z., 2026. Integrating geometric and causation probability approaches into Dynamic Bayesian Networks for real-time collision risk prediction. Transport. Res. E Logist. Transport. Rev. 205, 104520. https://doi.org/10.1016/j.tre.2025.104520.

Chen, C.-Z., Liu, S.-Y., Zou, Z.-J., Zou, L., Liu, J.-Z., 2023a. Time series prediction of ship maneuvering motion based on dynamic mode decomposition. Ocean. Eng. 286, 115446.

Chen, C.-Z., Zou, Z.-J., Zou, L., Zou, M., Kou, J.-Q., 2023b. Time series prediction of ship course keeping in waves using higher order dynamic mode decomposition. Phys. Fluids 35 (9).

Chen, L., Yang, P., Li, S., Tian, Y., Liu, G., Hao, G., 2022. Grey-box identification modeling of ship maneuvering motion based on LS-SVM. Ocean. Eng. 266, 112957.

Cooper, G.F., Herskovits, E., 1992. A Bayesian method for the induction of probabilistic networks from data. Mach. Learn. 9 (4), 309–347.

del Águila Ferrandis, J., Triantafyllou, M.S., Chryssostomidis, C., Karniadakis, G.E., 2021. Learning functionals via LSTM neural networks for predicting vessel dynamics in extreme sea states. Proceedings of the Royal Society A 477 (2245), 20190897.

Diez, M., Serani, A., Campana, E.F., Stern, F., 2021. Data-driven modelling of ship maneuvers in waves via dynamic mode decomposition (arXiv:2105.13062). arXiv. https://doi.org/10.48550/arXiv.2105.13062.

Dong, J., Wang, J., Zhang, F., 2026. Dynamic evolutionary Bayesian network model for offshore risk assessment. Ocean. Eng. 343, 123593. https://doi.org/10.1016/j.oceaneng.2025.123593.

Fan, L., Huang, S., Wang, M., Li, K.X., 2025. Dynamics of port state control and impact on ship safety: case studies of Tokyo and Indian Ocean Memorandum of understanding. Case Stud. Transp. Policy 21, 101518. https://doi.org/10.1016/j.cstp.2025.101518.

Fossen, T.I., 2011. Handbook of Marine Craft Hydrodynamics and Motion Control. John Wiley & Sons.

Gil, M., Montewka, J., Krata, P., 2025. Predicting a passenger ship's response during evasive maneuvers using Bayesian learning. Reliab. Eng. Syst. Saf. 256, 110765. https://doi.org/10.1016/j.ress.2024.110765.

Gu, S., Fu, S., Wu, M., Hao, Y., Mao, W., 2026. Risk management of ship independent navigation in ice-covered Arctic waters: an early warning model. Reliab. Eng. Syst. Saf. 267, 111920. https://doi.org/10.1016/j.ress.2025.111920.

Guo, H., Zou, Z., 2017. System-based investigation on 4-DOF ship maneuvering with hydrodynamic derivatives determined by RANS simulation of captive model tests. Appl. Ocean Res. 68, 11–25. https://doi.org/10.1016/j.apor.2017.08.006.

Guo, R., Mao, Y., Xiang, Z., Hao, L., Wu, D., Song, L., 2024. Research on LSTM-based maneuvering motion prediction for USVs. J. Mar. Sci. Eng. 12 (9), 1661.

Guo, S., Zhuang, S., Wang, J., Peng, X., Liu, Y., 2025. Deep learning-based non-parametric system identification and interpretability analysis for improving ship motion prediction. J. Mar. Sci. Eng. 13 (10), 2017.

Han, Z., Zhang, D., Fan, L., Zhang, J., Zhang, M., 2024. A Dynamic Bayesian network model to evaluate the availability of machinery systems in maritime autonomous surface ships. Accid. Anal. Prev. 194, 107342. https://doi.org/10.1016/j.aap.2023.107342.

He, H.-W., Wang, Z.-H., Zou, Z.-J., Liu, Y., 2022. Nonparametric modeling of ship maneuvering motion based on self-designed fully connected neural network. Ocean. Eng. 251, 111113. https://doi.org/10.1016/j.oceaneng.2022.111113.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Jiang, L., Fang, H., Shang, X., Zhang, Z., 2026. Hierarchical clustering weighted gaussian process regression low-fidelity model and application in airfoil drag coefficients prediction. IEEE Trans. Aero. Electron. Syst. 62, 9070–9081. https://doi.org/10.1109/TAES.2026.3678593.

Jiang, Y., Hou, X.-R., Wang, X.-G., Wang, Z.-H., Yang, Z.-L., Zou, Z.-J., 2022. Identification modeling and prediction of ship maneuvering motion based on LSTM deep neural network. J. Mar. Sci. Technol. 27 (1), 125–137.

Li, H., Meng, X., Zhang, W., Zhou, X.-Y., Yang, X., 2026. A real-time reliability assessment framework for marine mechanical equipment integrating machine learning and physical knowledge: toward applications in maritime autonomous surface ships. Reliab. Eng. Syst. Saf. 271, 112233. https://doi.org/10.1016/j.ress.2026.112233.

Liu, S.-Y., Chen, C.-Z., Zou, L., Zou, Z.-J., He, Y., 2024a. Echo state network-based black-box modeling and prediction of ship maneuvering motion. Ocean. Eng. 312, 119163. https://doi.org/10.1016/j.oceaneng.2024.119163.

Liu, S.-Y., Ouyang, Z.-L., Chen, G., Zhou, X., Zou, Z.-J., 2023. Black-box modeling of ship maneuvering motion based on Gaussian process regression with wavelet threshold denoising. Ocean. Eng. 271, 113765.

Liu, Y., Xue, Y., Lu, Y., Yuan, L., Li, F., Li, R., 2024b. A Dynamic Bayesian network model for ship navigation risk in the Arctic Northeast Passage. Ocean. Eng. 312, 119024. https://doi.org/10.1016/j.oceaneng.2024.119024.

Luo, W., Li, X., 2017. Measures to diminish the parameter drift in the modeling of ship manoeuvring using system identification. Appl. Ocean Res. 67, 9–20. https://doi.org/10.1016/j.apor.2017.06.008.

Meng, Y., Zhang, X., Zhang, X., Duan, Y., Xu, H., Guedes Soares, C., 2025. Multi-step online prediction of 4 degrees of freedom motion under various wind scales based on pre-training strategy and ensemble learning framework. Ocean. Eng. 339, 122096. https://doi.org/10.1016/j.oceaneng.2025.122096.

Montgomery, D.C., Peck, E.A., Vining, G.G., 2021. Introduction to Linear Regression Analysis. John Wiley & Sons.

Ouyang, Z.-L., Chen, G., Zou, Z.-J., 2023c. Identification modeling of ship maneuvering motion based on local Gaussian process regression. Ocean. Eng. 267, 113251.

Ouyang, Z.-L., Liu, S.-Y., Zou, Z.-J., 2022. Nonparametric modeling of ship maneuvering motion in waves based on Gaussian process regression. Ocean. Eng. 264, 112100.

Ouyang, Z.-L., Zou, Z.-J., 2021. Nonparametric modeling of ship maneuvering motion based on Gaussian process regression optimized by genetic algorithm. Ocean. Eng. 238, 109699. https://doi.org/10.1016/j.oceaneng.2021.109699.

Ouyang, Z.-L., Zou, Z.-J., Zou, L., 2023a. Adaptive hybrid-kernel function based Gaussian process regression for nonparametric modeling of ship maneuvering motion. Ocean. Eng. 268, 113373.

Ouyang, Z.-L., Zou, Z.-J., Zou, L., 2023b. Nonparametric modeling and control of ship steering motion based on local Gaussian process regression. J. Mar. Sci. Eng. 11 (11), 2161.

Quesada, D., Bielza, C., Fontán, P., Larrañaga, P., 2022. Piecewise forecasting of nonlinear time series with model tree dynamic Bayesian networks. Int. J. Intell. Syst. 37 (11), 9108–9137. https://doi.org/10.1002/int.22982.

Rothmund, S.V., Tengesdal, T., Brekke, E.F., Johansen, T.A., 2022. Intention modeling and inference for autonomous collision avoidance at sea. Ocean. Eng. 266, 113080. https://doi.org/10.1016/j.oceaneng.2022.113080.

Sarigul, D.O., Celik, C., Kinaci, O.K., Sarioz, K., Goren, O., 2025. A system-based approach to estimate manoeuvring performance of naval ships supported by captive tests in regular waves. Ocean. Eng. 321, 120378. https://doi.org/10.1016/j.oceaneng.2025.120378.

Scutari, M., 2010. Learning Bayesian networks with the bnlearn R package. J. Stat. Software 35, 1–22. https://doi.org/10.18637/jss.v035.i03.

Scutari, M., Kerob, D., Salah, S., 2024. Inferring skin–brain–skin connections from infodemiology data using dynamic Bayesian networks. Sci. Rep. 14 (1), 10266. https://doi.org/10.1038/s41598-024-60937-3.

SIMMAN, 2020. In: The Third Workshop on Verification and Validation of Ship Manoeuvring Simulation Methods. https://simman2020.kr/.

Sun, Q., Tang, Z., Gao, J., Zhang, G., 2022. Short-term ship motion attitude prediction based on LSTM and GPR. Appl. Ocean Res. 118, 102927.

Tian, K., Wen, Y.-Q., Zhu, M., Lei, T., Wang, S.-Y., Huang, Y.-M., 2025. Hybrid modeling and online prediction of 4-DOF ship maneuvering motion in disturbed waters. Ocean. Eng. 341, 122566. https://doi.org/10.1016/j.oceaneng.2025.122566.

Wang, N., Wu, M., Yuen, K.F., 2025a. Dynamic enterprise resilience assessment for port systems: a framework integrating Bayesian networks and Dempster-Shafer evidence theory. Reliab. Eng. Syst. Saf. 262, 111105. https://doi.org/10.1016/j.ress.2025.111105.

Wang, T., Ma, L., Ma, X., 2025b. Emergency scenario evolution of Arctic maritime accidents using data-driven dynamic Bayesian network. Transport. Res. Transport Environ. 146, 104845. https://doi.org/10.1016/j.trd.2025.104845.

Wang, Z., Park, S., Milano, C., Sanada, Y., Sarigul, D.O., Stern, F., Gunderson, A., Scherer, J., Yasukawa, H., Diez, M., 2025c. Physical analysis methods for planing hull turning circles. Ocean. Eng. 321, 120348.

Williams, C.K., Rasmussen, C.E., 2006. In: Gaussian Processes for Machine Learning, vol. 2. MIT press Cambridge, MA. Issue 3.

Wu, S., Cui, R., Zhang, L., 2026. Uncertain node-state PI-DBN: a novel framework for predictive modeling of real-time blowout risk in deepwater drilling. Reliab. Eng. Syst. Saf. 267, 111784. https://doi.org/10.1016/j.ress.2025.111784.

Xu, H., Guedes Soares, C., 2025. Review of system identification for manoeuvring modelling of marine surface ships. J. Mar. Sci. Appl. 1–20.

Xue, Y., Liu, Y., Ji, C., Xue, G., 2020. Hydrodynamic parameter identification for ship manoeuvring mathematical models using a Bayesian approach. Ocean. Eng. 195, 106612.

Yang, Y., el Moctar, O., 2025. A four-quadrant model for ship maneuvering in forward and backward motions. Appl. Ocean Res. 164, 104779. https://doi.org/10.1016/j.apor.2025.104779.

Yasukawa, H., Yoshimura, Y., 2015. Introduction of MMG standard method for ship maneuvering predictions. J. Mar. Sci. Technol. 20 (1), 37–52. https://doi.org/10.1007/s00773-014-0293-y.

Yuan, Q., Liu, Z., Wen, X., Peng, J., Dong, F., Zhou, R., Ye, J., 2025. Identification modeling of ship maneuvering motion based on AE-MSVR. J. Mar. Sci. Eng. 13 (10), 1942. https://doi.org/10.3390/jmse13101942.

Zhang, Q., Yan, H., Liu, Y., 2024. Power generation forecasting for solar plants based on Dynamic Bayesian networks by fusing multi-source information. Renew. Sustain. Energy Rev. 202, 114691. https://doi.org/10.1016/j.rser.2024.114691.

Zhang, X., Tian, R., Xu, H., Liu, X., Guedes Soares, C., 2025. Ship manoeuvring model identification using mutual information clustering with stepwise regression structural optimisation under small-sample constraints. Ocean. Eng. 341, 122594. https://doi.org/10.1016/j.oceaneng.2025.122594.

Zhang, Y.-Y., Wang, Z.-H., Zou, Z.-J., 2022. Black-box modeling of ship maneuvering motion based on multi-output nu-support vector regression with random excitation signal. Ocean. Eng. 257, 111279.

Zhou, X., Zou, L., He, H.-W., Wu, Z.-X., Zou, Z.-J., 2024. Real-time prediction of full-scale ship maneuvering motions at sea under random rudder actions based on BiLSTM-SAT hybrid method. Ocean. Eng. 314, 119664. https://doi.org/10.1016/j.oceaneng.2024.119664.

Zhou, X., Zou, L., Ouyang, Z.-L., Liu, S.-Y., Zou, Z.-J., 2023. Nonparametric modeling of ship maneuvering motions in calm water and regular waves based on R-LSTM hybrid method. Ocean. Eng. 285, 115259. https://doi.org/10.1016/j.oceaneng.2023.115259.

Zhou, X., Zou, L., Zou, Z.-J., 2025. A unified physics-data-driven modeling method for predicting time-series ship maneuvering motions in regular waves. Ocean. Eng. 341, 122747. https://doi.org/10.1016/j.oceaneng.2025.122747.

Zou, C., Hu, S., Chen, L., 2026. Environmental data driven dynamic Bayesian network: spatiotemporal evolution of coastal shipping risk performance in China. Reliab. Eng. Syst. Saf. 271, 112247. https://doi.org/10.1016/j.ress.2026.112247.