Importance of Data Richness in Identification of Ship Maneuvering Hydrodynamic Models

Vallabh Deogaonkar[1], Abhilash Somayajula[1,2]

[1]Marine Autonomous Vessels Laboratory, Department of Ocean Engineering
Indian Institute of Technology Madras, Chennai, India
[2]Center for Maritime Experiments to Maritime Experience (ME2ME), IIT Madras

Corresponding author: oe23d002@smail.iitm.ac.in

Abstract

For autonomous surface vessels (ASVs), accurate modeling of vessel dynamics is critical for the safe and reliable design of guidance and control systems. Unlike conventional vessels with human oversight, ASVs must independently handle challenging scenarios such as dynamic collision avoidance, pursuit–evasion, interception, and close-quarter maneuvering. These tasks inherently involve sharp maneuvers, where the vessel exhibits strongly nonlinear behavior and higher-order hydrodynamic effects play a significant role in determining motion. These nonlinear effects are modeled in simulations through a set of coefficients known as hydrodynamic coefficients. Estimating hydrodynamic coefficients is challenging and has traditionally been performed using experimental Planar Motion Mechanism (PMM) tests or computational fluid dynamics (CFD). However PMM tests require specialized and expensive facilities while CFD methods are time consuming. In the recent years a new approach of estimating hydrodynamic coefficients from free running data is gaining popularity. However most studies exploring this approach rely on only standard maneuvers such as turning circles or zig-zag maneuvers. This study demonstrates that datasets created by such standard maneuvers is not sufficiently rich enough to capture the nonlinear interactions and hence fall short on estimating higher order hydrodynamic coefficients. This study addresses a structured way to quantify data richness and proposes a new maneuver that can provide a rich dataset that can estimate linear and higher order hydrodynamic coefficients accurately.

Keywords: Ship maneuvering, system identification, hydrodynamic coefficients, data richness, model selection, entropy

Nomenclature

Symbol	Description
u, v, r	Surge velocity, sway velocity, and yaw rate in body-fixed frame
x_0, y_0	Position of vessel in global coordinate system
ψ	Heading angle of the vessel
δ	Rudder angle
δ_c	Commanded rudder angle
$\dot{u}, \dot{v}, \dot{r}$	Accelerations in surge, sway, and yaw directions
Δu	Deviation from nominal surge velocity
U	Nominal or design speed of the vessel
X, Y, N	Hydrodynamic surge force, sway force, and yaw moment
X', Y', N'	Non-dimensional hydrodynamic forces and moment
m	Vessel mass
I_{zz}	Mass moment of inertia about vertical axis
x_G	Longitudinal center of gravity in body frame
ρ	Density of water
L	Ship length
$\mathbf{M}$	Mass matrix (rigid body + added mass)
$\mathbf{C}(\nu)$	Coriolis matrix
ν	Velocity vector $[u\ v\ r]^T$
η	Position and orientation vector $[x_0\ y_0\ \psi]^T$
τ	External force vector $[X\ Y\ N]^T$
θ	Vector of hydrodynamic coefficients
λ	Threshold parameter in STLS algorithm
A	Regression matrix used for system identification
RSS	Residual sum of squares
BIC	Bayesian Information Criterion
H	Entropy of dataset
p_i	Probability of occupancy of bin i in entropy calculation

1 Introduction

Ship hydrodynamics is the study of how vessels interact with the surrounding water, encompassing the complex fluid-structure interactions that govern a ship's motion and stability. The effects of hull geometry, propulsion mechanisms and control surfaces and their combined interactions add to the complexity of this study (Birk, 2019). Computational tools such as Computational Fluid Dynamics (CFD) have emerged as powerful techniques to simulate a vessel's hydrodynamic behavior and substantial efforts have been devoted to reduce the computational costs associated with CFD simulations (Huang et al., 2023). Despite these advancements CFD-based approaches remain impractical for real-time applications such as iterative controller design or onboard simulation for autonomous surface vessels (ASVs) (Deraj et al., 2023; Mohan and Somayajula, 2023). These scenarios demand models that are both computationally efficient and capable of accurately capturing complex dynamics. Accurate modeling of vessel dynamics is particularly crucial for ASVs, which must autonomously navigate challenging scenarios such as dynamic obstacle avoidance (Jadhav et al., 2023; Yu and Roh, 2024; Yang et al., 2019), pursuit–evasion, or close-quarter docking without human intervention (Vijayakumar et al., 2024). These tasks typically involve sharp maneuvers and nonlinear motion regimes, where higher-order hydrodynamic effects significantly influence behavior.

Various mathematical models have been proposed to represent the maneuvering behavior of surface vessels, prioritizing computational speed over high-fidelity accuracy (Xu and Guedes Soares, 2025). These models are often categorized according to the amount of physical knowledge they incorporate. A black box model is one that learns the relationship between inputs and outputs purely from data without explicitly representing the underlying physics. Artificial neural networks and neural ordinary differential equation models are a common examples because they can approximate complex nonlinear patterns but do not reveal how individual physical effects contribute to the motion. At the opposite end are white box models which rely entirely on first principles. In such models every term in the equations is derived from physical laws and measured geometric properties. High-fidelity com-

putational fluid dynamics methods combined with rigid-body ship dynamics fall into this category because they resolve the flow field directly from the Navier–Stokes equations and compute the resulting forces on the hull, propeller, and rudder without requiring empirical tuning. Classical approaches based on potential flow theory, added mass formulations, and actuator disk representations for propeller forces also belong to this category. These models are fully interpretable because each term corresponds to a real physical quantity and its influence on the vessel is physically meaningful.

A gray box model combines physical insight with data driven components. In this category the model structure is derived from known principles of fluid structure interaction and ship hydrodynamics, while the unknown parameters are estimated from data. Such models are easier to interpret because each parameter corresponds to a physical effect. Two gray box models that are widely used in ship maneuvering research are the Abkowitz model and the modular Mathematical Model Group (MMG) model. The Abkowitz model was developed in the 1960s during the early efforts to describe maneuvering forces using polynomial expansions. It expresses the total hydrodynamic forces and moments as a Taylor series in terms of the vessel states and the rudder angle. This approach became popular because it provided a flexible mathematical structure that could be fitted to experimental data from towing tanks and free running tests. The MMG model emerged later from collaborative research in Japan led by the Manoeuvring Modelling Group. The goal of this group was to represent ship maneuvering in a way that clearly separated the contributions of the hull, propeller and rudder. This modular formulation made it possible to tune or replace individual components without rewriting the entire model. It also helped researchers understand how each subsystem influences the vessel motion, and it became widely adopted in both academic and industrial studies. Together these models illustrate the spectrum of approaches used in ship maneuvering research. Black box models prioritize flexible function approximation, while gray box models such as the Abkowitz and MMG formulations introduce physical structure that improves interpretability and consistency with known hydrodynamic behavior.

Traditionally, identifying the coefficients in such models has relied on extensive and controlled physical experiments. These are usually performed in specialized test facilities using the planar motion mechanism (PMM) or rotating arm tests, which isolate specific motion components to quantify corresponding hydrodynamic forces. However, such facilities are not widely available and are costly to build and maintain. Developments in data-driven and statistical system identification techniques provide an alternative pathway to identify the nonlinear hydrodynamic coefficients from measured motion data. The work by (Bendat et al., 1992) established a rigorous framework for identification of nonlinear systems from random data, providing procedures to analyze a variety of nonlinear behaviors and even reduce complicated single-input/output models to simpler multiple input/single output (MI/SO) equivalents. Concrete demonstrations of such approaches in the marine and offshore contexts have been provided by (Cheng and Falzarano, 2003), (Wang et al., 2020), (Somayajula and Falzarano, 2017) and (Somayajula and Falzarano, 2016). Free-running model tests conducted in large water basins provide a more accessible option compared to the specialized test facilities where measured motion data can be used for identification without restrictions of equipments required in the traditional facilities. However, their potential for system identification has not been fully explored or utilized. Recent studies have demonstrated the potential of using these free-running datasets in data-driven system identification approaches. However, this process presents several challenges. One major issue is the ill-conditioned nature of the regression matrices that arise during the identification process. Ill conditioning refers to a situation where small changes or noise in the data lead to disproportionately large changes in the estimated parameters. This results from high correlations among the columns of the regression matrices. When the regression matrix is ill conditioned, the identification algorithm becomes extremely sensitive to measurement errors, which can cause the estimated parameters to drift or fluctuate unpredictably and its termed as parameter drift. These conditions can lead to unstable or unreliable estimates. Additionally, free-running data may reflect uncontrolled initial conditions or environmental disturbances, further complicating the identification task.

To address these challenges, a wide range of system identification approaches have been applied in the maneuvering domain. Classical nonlinear regression methods have been widely used, where hydrodynamic forces are expressed using polynomial basis function expansions and coefficients are estimated through least squares formulations (Abkowitz, 1964). Kalman filtering methods, such as extended Kalman filter (Abkowitz, 1980) and unscented Kalman filter (Zheng et al., 2022), have been employed to account for measurement noise and time-varying uncertainties while ensuring smooth parameter evolution during identification. More recently, machine learning models such as feedforward neural networks have been investigated for their ability to represent complex nonlinear maneuvering behavior without requiring an explicit physical model structure (Wakita et al., 2022), (Rajesh and Bhattacharyya, 2008). Regularization techniques such as Lasso have been introduced to address multicollinearity and improve the stability of parameter estimation when datasets contain correlated inputs (Deogaonkar et al., 2023). Regression approaches based on support vectors have also been explored in several studies with variations in how the tolerance for model error and sparsity is controlled (Vijay and Somayajula, 2022), (Wang et al., 2019) (Wang et al., 2015). Frequency-domain identification methods have also been used to estimate hydrodynamic derivatives. (Selvam et al., 2005) introduced a Reverse-MISO spectral framework for identifying linear maneuvering coefficients directly from standard maneuvers . Global optimization methods such as particle swarm optimization and genetic algorithms have also been used to tune hydrodynamic parameters in cases where the search landscape is highly nonconvex or poorly conditioned (Sutulo and Soares, 2014).

A summary of representative studies and the datasets they employ is presented in Table 1. The table shows that the majority of existing work relies on a narrow collection of standard maneuvers, with turning circles and zigzag tests dominating both the training and validation stages. This consistent dependence on a limited set of motions suggests a need for exploration of whether the excitation present in these datasets is sufficient to expose the full nonlinear behavior of the vessel. Consequently, although many models report good predictive performance on these familiar maneuvers, their ability to generalize to a broader range of transient or nonlinear conditions remains unclear. This observation reveals an important gap in current research, namely that the influence of dataset composition on the robustness and generalization of identified models has not been systematically examined.

Classical system identification theory has long emphasized the critical role of input excitation in ensuring accurate model identification. As discussed in foundational works such as (Simpkins, 2012), the identifiability of a model depends on whether the input signal drives the system through a sufficiently varied portion of its phase space. The phase space refers to the mathematical space formed by all variables that describe the system at any instant. For a 3 degree of freedom ship maneuvering model these variables include surge velocity, sway velocity, yaw rates and their derivatives and together they define the state of the vessel. In linear systems, the requirement that the input signal must drive the system through a sufficiently rich portion of its phase space is formalized using conditions on the input itself. The input must vary in a diverse manner so that the system experiences many different combinations of its possible motions. When the input behaves this way, often described in the literature as meeting a persistent excitation condition, the resulting data contain enough variation to identify all parameters of the model reliably. (Box et al., 2015) similarly noted that the quality of an identified model is limited by the richness of the input. (Willems et al., 2005) formalized this idea by showing that persistent excitation is equivalent to requiring the Hankel matrix constructed from input–output time series to have full rank, guaranteeing that the collected data contain sufficient independent information for unique parameter estimation.

Table 1: Summary of representative ship maneuvering identification studies highlighting training and validation maneuvers.

Sr. No.	Author and Year	Data type	Identification Maneuvers (Training)	Method used for identification	Validation Maneuvers
1	Abkowitz (1980)	Experiments	10°/10° zig-zag	Extended Kalman Filter	Same 10°/10° zig-zag
2	Källström and Åström (1981)	Experiments	Open loop and closed loop pseudo random binary sequences, Zigzag 5/5, Zigzag 20/20	Model-based (Prediction Error, Maximum Likelihood)	PRBS inputs, Zig-Zag maneuvers, Autopilot perturbations.
3	Sutulo and Soares (2014)	Simulations	20°/20° zig-zag	Genetic algorithm based	35 turning circle and Dieudonne spiral
4	Araki et al. (2012)	Experiments, CFD, System Based	Turning circle 25, 35 ZigZag 10/10, 20/20, 35/35	extended Kalman filtering and constrained least square method using generalized reduced gradient algorithm	Tested on part of dataset not used for training
5	Wakita et al. (2022)	Experiments	Turning circles, ZigZags, Berthing (low speed), manual random	Neural Network	Low-speed tests, random maneuvers
6	Wang et al. (2015)	Simulations	20/20 zig-zag	ϵ-Support Vector Regression	35 turning circle, 10/10 ZigZag, part of 20/20 ZigZag
7	Diez et al. (2021)	Experiments	Course keeping and turning circles in waves	Dynamic Mode Decompositon	Tested on part of dataset not used for training
8	Woo et al. (2018)	Experiments	36 zigzag test data set and six 3-2-1-1 input test data set	LSTM	zigzag
9	Xue et al. (2021)	Experiments	10°/5°, 20°/5° and 30°/5° ZigZag	Gaussian Process Regression	15°/5°, 35°/5° and 10°/10° ZigZag
10	Costa et al. (2021)	Simulations and Experiments	Turning circle 25, 30/30 and 20/20 ZigZag	Singular value decomposition	20/20 ZigZag
11	Mitsuyuki et al. (2024)	Simulation	Turning circle 35	Bayesian inference using the Markov Chain Monte Carlo (MCMC) method	Turning cicle 35, ZigZag 20/20

In nonlinear systems such as ship maneuvering models the situation becomes more intricate. Although the phase space describes all theoretically possible combinations of system variables, physical and operational constraints mean that not all points in this space are actually reachable. Nonlinearity can further partition the phase space into regions that a system will remain confined to unless the excitation becomes sufficiently strong or diverse. The initial conditions and the applied control inputs below certain thresholds restrict the motion to a limited region of phase space and the resulting trajectory explores only a subset of the phase space. Data gathered from such confined motion reveal only the behaviors present in that restricted region. Identification based on these limited trajectories may produce models that fit the observed motions but fail to capture dynamics that occur in other reachable regions of the system's phase space. These considerations highlight the need for datasets that actively explore multiple regions of the vessel's feasible phase space in order to reliably identify its underlying nonlinear behavior.

Outside the marine-vessel community, there is growing recognition of the importance of data richness and experiment design in nonlinear system identification. For example, (Kiss et al., 2024) propose a space-filling input design to uniformly cover the operating range of a nonlinear state-space model; (Jegorova et al., 2020) generate diverse excitation trajectories using a generative-adversarial framework; (Heinz and Nelles, 2018) demonstrate a data-distribution-based approach to produce homogeneous coverage in systems with multiple inputs; and (Deshpande and Rivera, 2013) discuss optimal input signal design for data-centric estimation methods. These underlie the motivation for applying a similar data-design mindset in the context of ship maneuvering dataset generation.

Recent studies in ship maneuvering have begun to recognize that the variability and excitation content of datasets fundamentally influence the reliability of identified ship models. (Li et al., 2023) directly address this issue by proposing a model-structure optimization framework that enables identification under weak excitation. Their method first performs a Spearman-based correlation analysis to identify groups of hydrodynamic coefficients that cannot be independently estimated when the dataset is insufficiently varied. A subsequent sensitivity analysis quantifies the influence of each parameter on the predicted trajectories, allowing the authors to eliminate low-sensitivity terms and derive a reduced Abkowitz model with far fewer coefficients. This simplified structure is shown to yield more stable identification results for weakly excited datasets such as mild zigzags and random-rudder trials. In contrast,(Wakita et al., 2025) focus on improving generalization when only small or low-variability harbor maneuvering datasets are available. They introduce a data augmentation framework that generates synthetic variations of existing trajectories through controlled perturbations, noise injection, and resampling, thereby expanding the effective dataset without additional experiments. A complementary strategy is proposed by (Kanazawa et al., 2023), who mitigate the effects of weak excitation by transferring hydrodynamic knowledge from a well-characterized vessel to a target vessel with sparse data. Their approach constructs a mapping between similar ships and uses the source vessel's parameters as a regularizing prior, improving estimation robustness when the target dataset does not sufficiently excite all dynamic modes. While these approaches help alleviate the consequences of weak excitation, they ultimately address the symptoms rather than the underlying cause that the original datasets lack the variability required to explore the full nonlinear behavior of the vessel.

In the context of ASVs, the accuracy of maneuvering models directly impacts tasks such as trajectory planning, collision avoidance (Gil, 2021), and closed-loop control. These tasks require models that remain reliable under aggressive and nonlinear maneuvers, which are commonly encountered in real-world operational scenarios. Hence, the ability to identify hydrodynamic models from datasets that sufficiently capture these nonlinear behaviors is critical. This motivates the present study, where the focus is on understanding how the design of excitation maneuvers influences the reliability of the resulting models in autonomy applications. Several recent studies have emphasized the importance of sufficient data or simulation coverage to ensure reliable prediction of vessel behavior under uncertain environmental conditions. Simulation-driven analyses have highlighted the need for a sufficient number of realizations to capture stochastic effects such as irregular wave-induced responses in maneu-

vering models (Krata et al., 2024; Gil et al., 2024). These studies demonstrate that inadequate sampling of the operating space can lead to inaccurate or unreliable predictions.

While prior work focuses on the richness of environmental sampling and simulation coverage, the role of excitation richness in maneuvering datasets for system identification has received comparatively less attention. There remains no systematic investigation into how the richness of maneuvering datasets influences the accuracy and reliability of nonlinear ship model identification. This study addresses this gap by using simulated free-running trials generated from the Abkowitz model to systematically compare identification outcomes under varying levels of dataset richness. Here, richness refers to the variability and coverage of the dataset in the vessel's dynamic state space. This is distinct from data quality which is related to presence of noise due to measurements and is assumed ideal in simulation.

By analyzing how identification accuracy changes with dataset variability across two representative vessels, this study aims to highlight the limitations of standard maneuvers and motivate a data-design perspective in future identification studies. The literature also lacks a general, quantitative metric for the richness or coverage of maneuvering datasets. In practice, standard maneuvers such as turning circles and zig-zags are assumed to provide sufficient excitation. However, this assumption is not tested and no prior work has systematically quantified how well such maneuvers capture the complete variability of the dynamics of the vessel. The present study proposes and evaluates an information-theoretic metric as a quantitative measure of dataset richness. By applying this metric and showing how trajectories with different levels of excitation lead to different identification accuracy and generalization, this work fills the gap of defining a richness measure, revealing the limitations of standard maneuvers, and motivating a shift toward input-data design in system identification.

The rest of the paper is organized as follows: section 2 describes the maneuvering model used for simulation and the structure to be identified; section 3 presents the identification framework; section 4 examines the variability of different maneuver datasets; section 5 presents the metrics for quantifying the richness of different datasets and section 6 provides a comparative analysis of model identification outcomes using each dataset.

2 Vessel Dynamics

The dynamics of a ship moving through water is governed by Newton's laws of motion expressed by a set of differential equations with linear and nonlinear terms, which are collectively referred to as the maneuvering equations of motion. The hydrodynamic forces on a surface vessel are a result of the interaction of the hull with the surrounding water, the propeller thrust force and the forces and moments due to presence of various control surfaces such as the rudder.

In general, two frames of reference are considered in the mathematical models. A global coordinate system (GCS) frame $O_0 - X, Y, Z$ with Z-axis vertically downwards and a body-fixed coordinate system (BCS) frame $O - x_b, y_b, z_b$ with its origin at the intersection of midship, centerline and waterplane. The GCS and BCS are shown in Figure 1. The x-axis of the body fixed frame is pointed along the bow, and the y-axis is towards the starboard. In maneuvering studies, the ship motion is typically modeled in three degrees of freedom (3-DOF) namely the surge, sway, and yaw, under the assumption that roll, pitch, and heave motion are negligible compared to the maneuvering modes in calm-water conditions. In 3-DOF, the state of the vessel is expressed in terms of its x-coordinate (x_0) and y-coordinate (y_0) in GCS frame and heading angle (ψ) given by the angle between the x-axes of the BCS and GCS. The maneuvering equations of motion of the vessel are expressed in terms of the surge (u), sway (v), and yaw (r) velocities and they are measured in the body-fixed frame. The kinematics of such 3-DOF motion are given by (1).

$$\begin{aligned} \{\dot{\boldsymbol{\eta}}\} &= [R(\psi)]\{\boldsymbol{\nu}\} \\ R(\psi) &= \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{1}$$

where, $\{\boldsymbol{\nu}\} = \begin{bmatrix} u\ v\ r \end{bmatrix}^T$ denotes the vector of velocities expressed in BCS and $\{\boldsymbol{\eta}\} = \begin{bmatrix} x_0\ y_0\ \psi \end{bmatrix}^T$ denotes the vector of states expressed in GCS. $R(\psi)$ represents the rotation matrix such that any vector

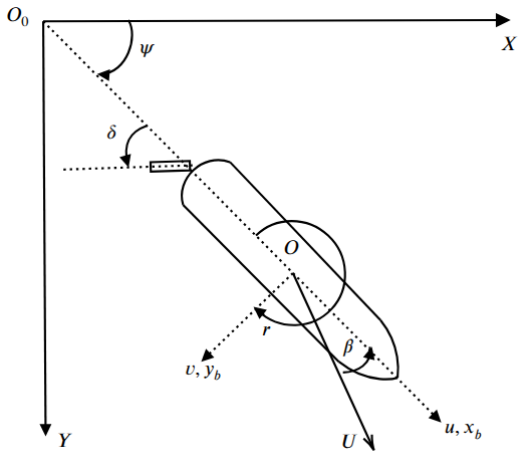


Figure 1: Ship coordinate frames and kinematic variables

in BCS when pre-multiplied by this rotation matrix results in the same vector expressed in GCS.

In this work, the (Abkowitz, 1980) model is adopted as the representative maneuvering model. It expresses the hydrodynamic forces and moments as a Taylor series in terms of the vessel's motion variables and control inputs about a nominal operating condition. The coefficients of the Taylor series are known as the hydrodynamic coefficients and encode the nonlinear dependence of hullform on the hydrodynamic forces experienced by the vessel. This makes it a general and versatile framework applicable to a wide range of surface vessels. Unlike the MMG model that requires explicit separation of the hull, rudder, and propeller contributions, the Abkowitz model encapsulates the overall hydrodynamic interaction implicitly through its coefficients, reducing the need for detailed subsystem modeling. The detailed mathematical formulation of the Abkowitz model is presented in subsection 2.1.

As shown in (Fossen, 2011) the nonlinear dynamics of a surface vessel in the 3 DOF of surge, sway and yaw can be expressed using (2)

$$[M]\{\dot{\boldsymbol{\nu}}\} + [C(\boldsymbol{\nu})]\{\boldsymbol{\nu}\} = \{\tau\} \tag{2}$$

where $[M]$ is the total mass matrix comprising of summation of rigid body and added mass matrix, $[C(\boldsymbol{\nu})]$ is the total coriolis matrix which is summation of rigid body and added mass coriolis matrix and $\{\boldsymbol{\tau}\}$ is the vector of external forces and moments due to surrounding water and control inputs. Here $\tau = \begin{bmatrix} X & Y & N \end{bmatrix}^T$ whose components are surge hydrodynamic force (X), the sway hydrodynamic force (Y) and the yaw hydrodynamic moment (N).

Maneuvering equations are usually expressed in non-dimensionalized forms to remove unit-dependent scaling, improve numerical conditioning and allow ship-to-ship comparison for identification and prediction. The superscript $(.)'$ is used to represent dimensionless variables. The total mass matrix and coriolis matrix with non dimensional terms are given by (3) and (4) respectively.

$$[M'] = \begin{bmatrix} m' - X'_{\dot{u}} & 0 & 0 \\ 0 & m' - Y'_{\dot{v}} & m'x'_G - Y'_{\dot{r}} \\ 0 & m'x'_G - N'_{\dot{v}} & I'_{zz} - N'_{\dot{r}} \end{bmatrix} \tag{3}$$

$$[C'(\boldsymbol{\nu})] = \begin{bmatrix} 0 & -m'r' & -m'x'_G r' + Y'_{\dot{v}} v' + Y'_{\dot{r}} r' \\ m'r' & 0 & -X'_{\dot{u}} u' \\ m'x'_G r' - Y'_{\dot{v}} v' - Y'_{\dot{r}} r' & X'_{\dot{u}} u' & 0 \end{bmatrix} \tag{4}$$

where m' and I'_{zz} denote the non dimensional mass and mass moment of inertia. x'_G is the position of centre of gravity in the longitudinal direction of body coordinate system (BCS). $X'_{\dot{u}}, Y'_{\dot{v}}, Y'_{\dot{r}}, N'_{\dot{v}}, N'_{\dot{r}}$ are the non dimensional added mass and added mass moment of inertia. This study follows the non-dimensionalization system as defined

$$\begin{aligned}
m' &= \frac{m}{\frac{1}{2}\rho L^3} & x'_G &= \frac{x_G}{L} & \delta' &= \delta \\
\dot{u}' &= \frac{\dot{u}L}{U^2} & \dot{v}' &= \frac{\dot{v}L}{U^2} & \dot{r}' &= \frac{\dot{r}L^2}{U^2} \\
u' &= \frac{u}{U} & v' &= \frac{v}{U} & r' &= \frac{rL}{U} \\
X'_{\dot{u}} &= \frac{X_{\dot{u}}}{\frac{1}{2}\rho L^3} & Y'_{\dot{v}} &= \frac{Y_{\dot{v}}}{\frac{1}{2}\rho L^3} & Y'_{\dot{r}} &= \frac{Y_{\dot{r}}}{\frac{1}{2}\rho L^4} \\
N'_{\dot{v}} &= \frac{N_{\dot{v}}}{\frac{1}{2}\rho L^4} & N'_{\dot{r}} &= \frac{N_{\dot{r}}}{\frac{1}{2}\rho L^5} & I'_{zz} &= \frac{I_{zz}}{\frac{1}{2}\rho L^5}
\end{aligned}$$

where ρ is the density of water, L is the ship length, U is the design speed of the ship.

2.1 Abkowitz model

The key differentiating factor between the various maneuvering models lies in how the vector $\{\boldsymbol{\tau}\}$ of external hydrodynamic forces and moments on the

ship is assumed. The Abkowitz model assumes that the components of vector $\{\boldsymbol{\tau}\}$ can be expressed as polynomial functions in surge, sway and yaw velocities and the instantaneous rudder angle (δ) and propeller speed. For the present analysis, only the rudder angle is considered as the control input that varies during the maneuvers, while the propeller revolutions per minute (RPM) are kept constant. This simplification isolates the effect of rudder-induced dynamics and enables a focused investigation of rudder excitation on the identifiability and generalization of the learned dynamic model. The components of $\{\boldsymbol{\tau}\}$ are then expressed as (5)

$$\begin{aligned} X &= f_X(u, v, r, \delta) \\ Y &= f_Y(u, v, r, \delta) \\ N &= f_N(u, v, r, \delta) \end{aligned} \tag{5}$$

where f_X, f_Y and f_N are nonlinear functions. A Taylor series approximation of these functions at a nominal operating point is used as the mathematical model for maneuvering motion. The nominal point about which the expansion is taken is given by $u = U$, $v = 0$, $r = 0$ and $\delta = 0$ where U is the design speed of the vessel. Expanding f_X about the nominal point yields the second–order approximation (6)

$$\begin{aligned} X(u, v, r, \delta) = {} & f_X(U, 0, 0, 0) + \left.\frac{\partial f_X}{\partial u}\right|_{(U,0,0,0)} \Delta u \\ & + \left.\frac{\partial f_X}{\partial v}\right|_{(U,0,0,0)} v + \left.\frac{\partial f_X}{\partial r}\right|_{(U,0,0,0)} r + \left.\frac{\partial f_X}{\partial \delta}\right|_{(U,0,0,0)} \delta \\ & + \tfrac{1}{2}\left(\left.\frac{\partial^2 f_X}{\partial u^2}\right|_{(U,0,0,0)} \Delta u^2 + \left.\frac{\partial^2 f_X}{\partial v^2}\right|_{(U,0,0,0)} v^2 \right) \\ & + \tfrac{1}{2}\left(\left.\frac{\partial^2 f_X}{\partial r^2}\right|_{(U,0,0,0)} r^2 + \left.\frac{\partial^2 f_X}{\partial \delta^2}\right|_{(U,0,0,0)} \delta^2 \right) \\ & + \left.\frac{\partial^2 f_X}{\partial u\, \partial v}\right|_{(U,0,0,0)} \Delta u\, v + \left.\frac{\partial^2 f_X}{\partial u\, \partial r}\right|_{(U,0,0,0)} \Delta u\, r \\ & + \left.\frac{\partial^2 f_X}{\partial u\, \partial \delta}\right|_{(U,0,0,0)} \Delta u\, \delta + \left.\frac{\partial^2 f_X}{\partial v\, \partial r}\right|_{(U,0,0,0)} v\, r \\ & + \left.\frac{\partial^2 f_X}{\partial v\, \partial \delta}\right|_{(U,0,0,0)} v\, \delta + \left.\frac{\partial^2 f_X}{\partial r\, \partial \delta}\right|_{(U,0,0,0)} r\, \delta + \cdots \end{aligned} \tag{6}$$

Here, all partial derivatives are evaluated at $(u, v, r, \delta) = (U, 0, 0, 0)$, and $\Delta u = u - U$ denotes the deviation from the nominal surge speed. Corresponding Taylor expansions for $Y(u, v, r, \delta)$ and $N(u, v, r, \delta)$ follow the same structure and are omitted for brevity. In maneuvering theory, it is customary to introduce shorthand notation for the Taylor-series terms so that the hydrodynamic forces can be written in the standard polynomial Abkowitz form. Each coefficient in this model corresponds directly to a partial derivative of the underlying nonlinear function evaluated at the nominal operating point $(u, v, r, \delta) = (U, 0, 0, 0)$.

The zero-order terms are defined as

$$\begin{aligned} X_0 &= f_X(U, 0, 0, 0) \\ Y_0 &= f_Y(U, 0, 0, 0) \\ N_0 &= f_N(U, 0, 0, 0), \end{aligned} \tag{7}$$

representing the steady hydrodynamic force or moment when the vessel travels at the design speed with no sway, no yaw, and zero rudder angle.

The first-order derivatives give rise to the linear hydrodynamic coefficients. For example,

$$\begin{aligned} X_u &= \left.\frac{\partial f_X}{\partial u}\right|_{(U,0,0,0)} \\ X_v &= \left.\frac{\partial f_X}{\partial v}\right|_{(U,0,0,0)} \\ X_\delta &= \left.\frac{\partial f_X}{\partial \delta}\right|_{(U,0,0,0)} \end{aligned} \tag{8}$$

and similarly for $Y_{(\cdot)}$ and $N_{(\cdot)}$. These coefficients quantify how sensitively the corresponding hydrodynamic force or moment changes with a small perturbation in one of the state variables or the rudder input.

The second-order derivatives define the quadratic (nonlinear) hydrodynamic coefficients. For instance,

$$\begin{aligned} X_{uu} &= \frac{1}{2!} \left.\frac{\partial^2 f_X}{\partial u^2}\right|_{(U,0,0,0)} \\ X_{uv} &= \frac{1}{2!} \left.\frac{\partial^2 f_X}{\partial u\, \partial v}\right|_{(U,0,0,0)} \\ X_{v\delta} &= \frac{1}{2!} \left.\frac{\partial^2 f_X}{\partial v\, \partial \delta}\right|_{(U,0,0,0)} \end{aligned} \tag{9}$$

where unmixed second-order terms such as X_{uu}, X_{vv}, $X_{\delta\delta}$ capture curvature with respect to a single variable, and mixed terms such as X_{uv}, $X_{u\delta}$ represent nonlinear interactions between different

motion components or between motion and rudder input.

Using this shorthand notation, the long Taylor-series expansion can be written compactly in the standard Abkowitz polynomial form, where each hydrodynamic coefficient appears as a constant multiplying monomials such as u, vr, r^2, $v\delta$, and so on. This provides a finite, interpretable parameterization of the maneuvering dynamics suitable for data-driven identification. The 3 DOF non dimensional Abkowitz model in matrix form is given as follows :

$$[M']\{\dot{\boldsymbol{\nu}}'\} + [C'(\boldsymbol{\nu})]\{\boldsymbol{\nu}'\} = \{\tau'\} = \begin{bmatrix} X' \\ Y' \\ N' \end{bmatrix} \quad (10)$$

where $\dot{\boldsymbol{\nu}}' = \begin{bmatrix} \dot{u}' & \dot{v}' & \dot{r}' \end{bmatrix}^T$ and $\dot{u}', \dot{v}', \dot{r}'$ denote the non-dimensional accelerations in the surge, sway and yaw directions respectively. The non dimensional external forces and moments acting on the vessel in surge, sway and yaw modes respectively are given by X', Y' and N'. The taylor series expansion in $\Delta u', v', r'$ and δ of the nonlinear functions of these non dimensional external forces and moments up to third order is given in (**??**), where $\Delta u' = u' - 1$. The coefficients in this expression denote the non dimensional hydrodynamic coefficients. These coefficients are functions of the hull geometry and Reynolds number.

$$\begin{aligned} X' = {} & X_0 + X'_u\Delta u' + X'_v v' + X'_r r' + X'_\delta \delta \\ & + X'_{uu}\Delta u'^2 + X'_{vv}v'^2 + X'_{rr}r'^2 + X'_{\delta\delta}\delta^2 \\ & + X'_{uv}\Delta u'v' + X'_{ur}\Delta u'r' + X'_{u\delta}\Delta u'\delta \\ & + X'_{vr}v'r' + X'_{v\delta}v'\delta + X'_{r\delta}r'\delta + X'_{uuu}\Delta u'^3 \\ & + X'_{vvv}v'^3 + X'_{rrr}r'^3 + X'_{\delta\delta\delta}\delta'^3 + X'_{uuv}\Delta u'v' \\ & + X'_{uur}\Delta u'r' + X'_{uu\delta}\Delta u'\delta + X'_{vvu}v'^2\Delta u' \\ & + X'_{vvr}v'^2r' + X'_{vv\delta}v'^2\delta + X'_{rru}r'^2\Delta u' \\ & + X'_{rrv}r'^2v' + X'_{rr\delta}r'^2\delta + X'_{\delta\delta u}\delta'^2\Delta u' \\ & + X'_{\delta\delta v}\delta'^2v' + X'_{\delta\delta r}\delta'^2r' + X'_{uvr}\Delta u'v'r' \\ & + X'_{uv\delta}\Delta u'v'\delta' + X'_{ur\delta}\Delta u'r'\delta + X'_{vr\delta}v'r'\delta \end{aligned} \quad (11)$$

$$\begin{aligned} Y' = {} & Y_0 + Y'_u\Delta u' + Y'_v v' + Y'_r r' + Y'_\delta \delta \\ & + Y'_{uu}\Delta u'^2 + Y'_{vv}v'^2 + Y'_{rr}r'^2 + Y'_{\delta\delta}\delta^2 \\ & + Y'_{uv}\Delta u'v' + Y'_{ur}\Delta u'r' + Y'_{u\delta}\Delta u'\delta \\ & + Y'_{vr}v'r' + Y'_{v\delta}v'\delta + Y'_{r\delta}r'\delta + Y'_{uuu}\Delta u'^3 \\ & + Y'_{vvv}v'^3 + Y'_{rrr}r'^3 + Y'_{\delta\delta\delta}\delta'^3 + Y'_{uuv}\Delta u'v' \\ & + Y'_{uur}\Delta u'r' + Y'_{uu\delta}\Delta u'\delta + Y'_{vvu}v'^2\Delta u' \\ & + Y'_{vvr}v'^2r' + Y'_{vv\delta}v'^2\delta + Y'_{rru}r'^2\Delta u' \\ & + Y'_{rrv}r'^2v' + Y'_{rr\delta}r'^2\delta + Y'_{\delta\delta u}\delta'^2\Delta u' \\ & + Y'_{\delta\delta v}\delta'^2v' + Y'_{\delta\delta r}\delta'^2r' + Y'_{uvr}\Delta u'v'r' \\ & + Y'_{uv\delta}\Delta u'v'\delta' + Y'_{ur\delta}\Delta u'r'\delta + Y'_{vr\delta}v'r'\delta \end{aligned} \quad (12)$$

$$\begin{aligned} N' = {} & N_0 + N'_u\Delta u' + N'_v v' + N'_r r' + N'_\delta \delta \\ & + N'_{uu}\Delta u'^2 + N'_{vv}v'^2 + N'_{rr}r'^2 + N'_{\delta\delta}\delta^2 \\ & + N'_{uv}\Delta u'v' + N'_{ur}\Delta u'r' + N'_{u\delta}\Delta u'\delta \\ & + N'_{vr}v'r' + N'_{v\delta}v'\delta + N'_{r\delta}r'\delta + N'_{uuu}\Delta u'^3 \\ & + N'_{vvv}v'^3 + N'_{rrr}r'^3 + N'_{\delta\delta\delta}\delta'^3 + N'_{uuv}\Delta u'v' \\ & + N'_{uur}\Delta u'r' + N'_{uu\delta}\Delta u'\delta + N'_{vvu}v'^2\Delta u' \\ & + N'_{vvr}v'^2r' + N'_{vv\delta}v'^2\delta + N'_{rru}r'^2\Delta u' \\ & + N'_{rrv}r'^2v' + N'_{rr\delta}r'^2\delta + N'_{\delta\delta u}\delta'^2\Delta u' \\ & + N'_{\delta\delta v}\delta'^2v' + N'_{\delta\delta r}\delta'^2r' + N'_{uvr}\Delta u'v'r' \\ & + N'_{uv\delta}\Delta u'v'\delta' + N'_{ur\delta}\Delta u'r'\delta + N'_{vr\delta}v'r'\delta \end{aligned} \quad (13)$$

2.2 Rudder angle variation

In this study, to account for rudder rate constraints in order to match practical scenarios, the variation of actual rudder angle δ is governed by a first order equation as shown in (14).

$$T'_R\dot{\delta} + \delta = \delta_c \quad (14)$$

Here δ_c denotes the commanded rudder angle. The actual rudder angle δ evolves slowly for a step change in commanded rudder angle δ_c. A lower value of T'_R leads to a faster response time. In addition, the rudder rate $\dot{\delta}$ is also saturated at $\dot{\delta}_{max}$. So the rudder-rate is given by:

$$\dot{\delta} = \begin{cases} \frac{\delta_c - \delta}{T'_R} & \text{if } |\frac{\delta_c - \delta}{T'_R}| \leq \dot{\delta}_{max} \\ \dot{\delta}_{max} & \text{if } \frac{\delta_c - \delta}{T'_R} > \dot{\delta}_{max} \\ -\dot{\delta}_{max} & \text{if } \frac{\delta_c - \delta}{T'_R} < -\dot{\delta}_{max} \end{cases} \quad (15)$$

In this study, the non-dimensional rudder time-constant T'_R is taken as 1 and $\dot{\delta}_{max}$ is chosen as 5° per second for the full scale ship. The maximum commanded rudder angle value (δ_{max}) is chosen to be 35° for both starboard and port side rudder command.

2.3 Vessels used for analysis

Two benchmark vessels are considered in this study : the Mariner class vessel and the KCS (KRISO Container Ship) model. These vessels are widely used in maneuvering studies as benchmark hulls and their hydrodynamic coefficients are publicly available. Thus the identified coefficients from system identification can be easily compared with the true values.

The Mariner model represents a conventional full-form cargo vessel with moderate block coefficient and low slenderness ratio. Such hull forms are known to exhibit stronger nonlinear coupling effects in sway and yaw during sharp maneuvers. The Mariner-class vessel coefficients are taken from the Mariner model implemented in the MSS toolbox (Perez et al., 2006). The KCS model, in contrast, is a fine-form container ship characterized by a higher length-to-beam ratio and lower block coefficient, leading to distinct hydrodynamic responses. The hydrodynamic coefficients used in the simulations for Mariner and KCS vessel are summarized in Table 2

3 System identification methodology

3.1 Problem setup

In this study the simulated trajectory data is generated using the Abkowitz model described in subsection 2.1. The simulation is carried out using Runge Kutta 4^{th} order explicit scheme with 0.05 non dimensional timestep. The resulting trajectories include surge velocity u, sway velocity v, yaw rate r, and commanded rudder angle δ. Distinct trajectory data is generated with various commanded rudder angle inputs. These datasets are then used to identify the hydrodynamic coefficients through a regression and compared with their ground truth values shown in Table 2.

3.2 Formulation for regression

In order to setup a regression for hydrodynamic coefficient, the Abkowitz model in (10) is rewritten as shown in (16).

$$[A] \cdot \{X_c\} = (m' - X'_{\dot{u}})\{\dot{u'}\} - \{X'_{Coriolis}\} \tag{16a}$$

$$[A] \cdot \{Y_c\} = (m' - Y'_{\dot{v}})\{\dot{v'}\} + (m'x'_G - Y'_{\dot{r}})\{\dot{r'}\} - \{Y'_{Coriolis}\} \tag{16b}$$

$$[A] \cdot \{N_c\} = (m'x'_G - N'_{\dot{v}})\{\dot{v'}\} + (I'_{zz} - N'_{\dot{r}})\{\dot{r'}\} - \{N'_{Coriolis}\} \tag{16c}$$

where $[A]$ denotes the regressor matrix for surge, sway, yaw equation of motion and is defined as shown in (17a) where n is the number of data points. $\{X_c\}$, $\{Y_c\}$, $\{N_c\}$ are the coefficient vectors to be identified using regression. The coefficient vector $\{X_c\}$ is shown in (17b) and $\{Y_c\}$, $\{N_c\}$ follow the same structure and are omitted for brevity.

$$\begin{aligned}[A] = [&1, \Delta u', v', r', \delta, \Delta u'^2, v'^2, r'^2, \delta'^2,\\ &\Delta u'v', \Delta u'r', \Delta u'\delta, v'r', v'\delta, r'\delta\\ &\Delta u'^3, v'^3, r'^3, \delta^3, \Delta u'^2v', \Delta u'^2r',\\ &\Delta u'^2\delta, v'^2\Delta u', v'^2r', v'^2\delta, r'^2\Delta u', r'^2v', r'^2\delta,\\ &\delta^2\Delta u', \delta^2v', \delta^2r', \Delta u'v'r', \Delta u'v'\delta\\ &\Delta u'r'\delta, v'r'\delta]_{n\times 35}\end{aligned} \tag{17a}$$

$$\begin{aligned}X_c = [&X_0, X'_u, X'_v, X'_r, X'_\delta, X'_{uu}, X'_{vv}, X'_{rr}, X'_{\delta\delta},\\ &X'_{uv}, X'_{ur}, X'_{u\delta}, X'_{vr}, X'_{v\delta}, X'_{r\delta},\\ &X'_{uuu}, X'_{vvv}, X'_{rrr}, X'_{\delta\delta\delta}, X'_{uuv}, X'_{uur}, X'_{uu\delta},\\ &X'_{vvu}, X'_{vvr}, X'_{vv\delta}, X'_{rru}, X'_{rrv}, X'_{rr\delta}, X'_{\delta\delta u},\\ &X'_{\delta\delta v}, X'_{\delta\delta r}, X'_{uvr}, X'_{uv\delta}, X'_{ur\delta}, X'_{vr\delta}]^T_{35\times 1}\end{aligned} \tag{17b}$$

$\{X'_{Coriolis}\}$, $\{Y'_{Coriolis}\}$, $\{N'_{Coriolis}\}$ denote the surge, sway and yaw component of non dimensional Coriolis force and are given by (18).

$$\begin{aligned}X_{Coriolis} &= -m'r'v' - m'x'_Gr'^2 + Y'_{\dot{v}}v'r' + Y'_{\dot{r}}r'^2\\ Y_{Coriolis} &= m'r'u' - X'_{\dot{u}}u'r'\\ N_{Coriolis} &= m'x'_Gr'u' - Y'_{\dot{v}}u'v' - Y'_{\dot{r}}u'r' + X'_{\dot{u}}u'v'\end{aligned} \tag{18}$$

Table 2: Mariner and KCS non dimensional Hydrodynamic coefficients (all values to be multiplied by 10^{-5})

Surge	Mariner	KCS	Sway	Mariner	KCS	Yaw	Mariner	KCS
X'_0	0.0	0.0	Y'_0	-4.0	0.0	N'_0	3.0	0.0
X'_u	-184.0	-232.5	Y'_u	-8.0	0.0	N'_u	6.0	0.0
X'_v	0.0	0.0	Y'_v	-1160.0	-1203.2	N'_v	-264.0	221.2
X'_r	0.0	0.0	Y'_r	-499.0	303.4	N'_r	-166.0	-257.3
X'_δ	0.0	0.0	Y'_δ	278.0	-268.9	N'_δ	-139.0	136.0
X'_{uu}	-110.0	-75.6	Y'_{uu}	-4.0	0.0	N'_{uu}	3.0	0.0
X'_{vv}	-899.0	666.4	Y'_{vv}	0.0	1873.9	N'_{vv}	0.0	-1026.6
X'_{rr}	18.0	0.0	Y'_{rr}	0.0	110.2	N'_{rr}	0.0	-340.7
$X'_{\delta\delta}$	-95.0	-146.9	$Y'_{\delta\delta}$	0.0	0.0	$N'_{\delta\delta}$	0.0	0.0
X'_{uv}	0.0	0.0	Y'_{uv}	-1160.0	4278.4	N'_{uv}	-264.0	638.4
X'_{ur}	0.0	0.0	Y'_{ur}	-499.0	1838.5	N'_{ur}	-166.0	-367.8
$X'_{u\delta}$	0.0	0.0	$Y'_{u\delta}$	556.0	-586.3	$N'_{u\delta}$	-278.0	180.2
X'_{vr}	798.0	420.1	Y'_{vr}	0.0	760.6	N'_{vr}	0.0	-948.9
$X'_{v\delta}$	93.0	-269.0	$Y'_{v\delta}$	0.0	0.0	$N'_{v\delta}$	0.0	-83.7
$X'_{r\delta}$	0.0	0.0	$Y'_{r\delta}$	0.0	0.0	$N'_{r\delta}$	0.0	36.8
X'_{uuu}	-215.0	0.0	Y'_{uuu}	0.0	0.0	N'_{uuu}	0.0	0.0
X'_{vvv}	0.0	0.0	Y'_{vvv}	-8078.0	17100.6	N'_{vvv}	1636.0	-13610.5
X'_{rrr}	0.0	0.0	Y'_{rrr}	0.0	793.9	N'_{rrr}	0.0	-911.8
$X'_{\delta\delta\delta}$	0.0	0.0	$Y'_{\delta\delta\delta}$	-90.0	106.6	$N'_{\delta\delta\delta}$	45.0	-66.4
X'_{uuv}	0.0	0.0	Y'_{uuv}	0.0	-931.3	N'_{uuv}	0.0	0.0
X'_{uur}	0.0	0.0	Y'_{uur}	0.0	2861.9	N'_{uur}	0.0	-352.5
$X'_{uu\delta}$	0.0	0.0	$Y'_{uu\delta}$	278.0	-1002.9	$N'_{uu\delta}$	-139.0	239.1
X'_{vvu}	0.0	0.0	Y'_{vvu}	0.0	0.0	N'_{vvu}	0.0	0.0
X'_{vvr}	0.0	0.0	Y'_{vvr}	15356.0	15339.7	N'_{vvr}	-5483.0	-15406.1
$X'_{vv\delta}$	0.0	0.0	$Y'_{vv\delta}$	1190.0	0.0	$N'_{vv\delta}$	-489.0	1532.2
X'_{rru}	0.0	263.0	Y'_{rru}	0.0	0.0	N'_{rru}	0.0	-352.0
X'_{rrv}	0.0	0.0	Y'_{rrv}	0.0	0.0	N'_{rrv}	0.0	-4530.2
$X'_{rr\delta}$	0.0	0.0	$Y'_{rr\delta}$	0.0	163.7	$N'_{rr\delta}$	0.0	265.6
$X'_{\delta\delta u}$	-190.0	-92.8	$Y'_{\delta\delta u}$	0.0	0.0	$N'_{\delta\delta u}$	0.0	0.0
$X'_{\delta\delta v}$	0.0	0.0	$Y'_{\delta\delta v}$	-4.0	0.0	$N'_{\delta\delta v}$	13.0	0.0
$X'_{\delta\delta r}$	0.0	0.0	$Y'_{\delta\delta r}$	0.0	0.0	$N'_{\delta\delta r}$	0.0	0.0
X'_{uvr}	0.0	0.0	Y'_{uvr}	0.0	0.0	N'_{uvr}	0.0	0.0
$X'_{uv\delta}$	93.0	0.0	$Y'_{uv\delta}$	0.0	0.0	$N'_{uv\delta}$	0.0	-365.1
$X'_{ur\delta}$	0.0	0.0	$Y'_{ur\delta}$	0.0	0.0	$N'_{ur\delta}$	0.0	0.0
$X'_{vr\delta}$	0.0	0.0	$Y'_{vr\delta}$	0.0	1743.5	$N'_{vr\delta}$	0.0	746.1

The non dimensional accelerations $\{\dot{u'}\}$, $\{\dot{v'}\}$ and $\{\dot{r'}\}$ are discretized and computed for each datapoint using 7 point finite-difference stencil. The added mass coefficients in the (16) are assumed to be known since these can be computed using a 3 dimensional potential theory based panel method code given the geometry of the vessel. The hydrodynamic derivatives in (17b) are obtained by solving (16) using regression as described further.

3.3 Sequential Thresholded Least Squares estimation

In the identification of maneuvering models, each experiment typically yields hundreds to thousands of time-series samples, whereas the number of hydrodynamic coefficients to be estimated is comparatively small. As a result, the regression problem contains far more equations than unknowns, forming what is known as an overdetermined system. Although such systems offer the advantage of abundant data, they can also introduce difficulties when the measurements contain noise, when input channels are correlated, or when the excitation is weak in certain directions. Under these conditions, a standard least-squares solution may become ill-conditioned, producing coefficients with large magnitudes that compensate for noisy or redundant information rather than reflecting true physical effects. Regularization is therefore essential to obtain stable and interpretable parameter estimates. Sparse regularization is particularly valuable in maneuvering dynamics, where only a subset of hydrodynamic terms are expected to contribute significantly and the remaining terms should be suppressed for physical plausibility.

To address these challenges, the hydrodynamic coefficients in this study are estimated using the Sequential Thresholded Least Squares (STLS) algorithm, which iteratively applies least-squares fitting and coefficient thresholding to produce a stable, sparse, and physically interpretable model. This implementation is described in the Sparse Identification of Nonlinear Dynamics (SINDy) framework (Brunton et al., 2016) and summarized here for brevity.

The regression problem can be written as:

$$[A]\{\theta\} = \{b\}, \tag{19}$$

where A is the regression matrix composed of candidate nonlinear terms derived from the measured states and control inputs as in (17a), b represents the measured accelerations which are the derivatives of the motion states, and θ is the vector of hydrodynamic coefficients to be identified as shown in (17b).

The STLS algorithm proceeds as follows:

1. Initial least-squares estimate:
   An initial estimate of the coefficients is obtained by solving
   $$\theta^{(0)} = \arg\min_{\theta} \|[A]\{\theta\} - \{b\}\|_2^2 \tag{20}$$
2. Thresholding step:
   Coefficients with small magnitudes are set to zero based on a predefined threshold λ:
   $$\theta_i^{(k+1)} = \begin{cases} \theta_i^{(k)}, & \text{if } |\theta_i^{(k)}| > \lambda, \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$
3. Refitting on the active set:
   The least-squares regression is repeated using only the active (nonzero) coefficients to obtain an updated estimate of θ.

This sequential process of fitting, thresholding, and refitting continues until the active coefficient set converges. Compared to conventional least-squares regression, the STLS method yields a sparser and more interpretable model, eliminating weakly contributing terms that do not significantly affect the system dynamics. This approach is particularly suitable for identifying nonlinear maneuvering models, where the number of potential interaction terms is large but only a subset significantly influences the vessel's motion.

3.4 Model selection using BIC

To evaluate the parsimony of the identified models, the Bayesian Information Criterion (BIC) is employed. BIC score is defined as follows:

$$\text{BIC} = \ln(n)k - 2\ln(\mathcal{L}(\hat{\theta}|\{\Delta u\}, \{v\}, \{r\}, \{\delta\})) \tag{22}$$

where k is the number of estimated parameters, n is the number of data samples, and $\mathcal{L}(\hat{\theta} \,|\, \{\Delta u\}, \{v\}, \{r\}, \{\delta\})$ denotes the likelihood of

the model. For a given dataset changing the value of the thresholding parameter (λ) in the STLS algorithm changes the number of parameters being estimated and the likelihood. To compare different models with different values of the (λ) the BIC score is used. A model with a lower BIC score is preferred. The BIC score therefore penalizes models with a larger number of parameters while rewarding those that achieve a higher likelihood, providing a principled criterion for selecting parsimonious models. The likelihood function represents the joint probability of observing the data $\{\Delta u\}, \{v\}, \{r\}, \{\delta\}$ given a particular choice of model parameters. The likelihood is expressed as follows:

$$\mathcal{L}(\theta \mid \mathcal{D}) = P(d_1, d_2, ..., d_n; \theta), \tag{23}$$

where $d_{i=1,2,...,n}$ denotes the i^{th} data sample, $\mathcal{D}$ denotes the complete dataset ($\{\Delta u\}, \{v\}, \{r\}, \{\delta\}$) and P denotes the joint probability density. For the ship maneuvering dynamical system, the consecutive data samples in a trajectory are not statistically independent, since the vessel's motion evolves according to a time-dependent differential equation. As a result, the velocities and accelerations at time t_{i+1} depend on those at time t_i. However, following standard practice in regression-based identification, the common approximation is to assume that the samples are conditionally independent given the model parameters. Under this assumption, the joint likelihood can be written as a product of individual probability densities,

$$\mathcal{L}(\theta \mid \mathcal{D}) = \prod_{i=1}^{n} p(d_i \mid \theta) \tag{24}$$

which allows the log-likelihood to be expressed as

$$\ln \mathcal{L}(\theta \mid \mathcal{D}) = \sum_{i=1}^{n} \ln p(d_i \mid \theta) \tag{25}$$

By treating the samples as independent, the identification process ignores the transition structure, as well as any correlation in the process noise or model errors across time. Consequently, the regression captures only the instantaneous mapping between velocities, control inputs, and their derivatives, but it does not enforce consistency with the underlying time evolution of the dynamics. As a result, the identified model may predict accelerations accurately at individual instants, yet still drift when its predictions are integrated forward in time to reconstruct a full trajectory. This occurs because the temporal dependencies inherent to the vessel's equations of motion are not explicitly enforced in the regression formulation. This approximation is common in regression-based identification, but it should be recognized.

To compute the likelihood consider the regression for surge equation of motion, written in matrix form as

$$\{b_1\} = [A]\{X_c\} + \{\varepsilon\}, \tag{26}$$

where $\{b_1\} \in R^n$ is the vector of measured outputs as shown in the RHS of (16a), $[A] \in R^{n\times 35}$ is the regressor matrix corresponding to the monomials in the Abkowitz expansion , $X_c \in R^{35}$ is the vector of unknown coefficients, and $\{\varepsilon\}$ represents the error between the measured outputs and the model prediction. The residual vector $\{\varepsilon\}$ in (26) assumed to consist of independent, zero-mean Gaussian random variables with variance σ^2, i.e. $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. This assumption is commonly adopted because it leads to a mathematically tractable likelihood expression. Under Gaussian errors, the log-likelihood reduces to a quadratic form in the residuals, allowing the regression problem to be solved and compared analytically. Additionally, the aggregate measurement and modelling errors at each sample typically arise from many small independent sources, and are therefore well approximated by a Gaussian distribution as a result of the central limit theorem.

Under the independence and Gaussian assumptions the likelihood of the coefficient vector $\{X_c\}$ conditional on the measured dataset is given by the product of scalar Gaussian densities:

$$\begin{aligned}&\mathcal{L}(\{X_c\} \mid \mathcal{D}) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(b_{1,i} - A_i^\top \{X_c\})^2}{2\sigma^2}\right),\end{aligned} \tag{27}$$

where $A_i^\top$ denotes the i-th row of the regressor matrix $[A]$ and $b_{1,i}$ is the i-th entry of $\{b_1\}$. The

log-likelihood is therefore

$$\ln \mathcal{L}(\{X_c\} \mid \mathcal{D}) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(b_{1,i} - A_i^\top\{X_c\}\right)^2. \tag{28}$$

The residual sum of squares (RSS) associated with coefficient vector $\{X_c\}$ is defined as

$$\begin{aligned}\mathrm{RSS}(\{X_c\}) &= \sum_{i=1}^{n}\left(b_{1,i} - A_i^\top\{X_c\}\right)^2 \\ &= \|\{b_1\} - [A]\{X_c\}\|_2^2. \end{aligned} \tag{29}$$

The parameter estimate $\hat{\theta}$ used in the BIC computation is the maximum likelihood estimate (MLE), obtained by maximizing the likelihood function, or equivalently minimizing the negative log-likelihood:

$$\begin{aligned}\hat{\theta} &= \arg\max_{\theta}\ \mathcal{L}(\theta \mid \mathcal{D}) \quad \text{or} \\ \hat{\theta} &= \arg\min_{\theta}\ \left[-\ln \mathcal{L}(\theta \mid \mathcal{D})\right]. \end{aligned} \tag{30}$$

For the surge equation of motion the maximum likelihood estimate $\{\hat{X}_c\}$ is the minimizer of $\mathrm{RSS}(\{X_c\})$. Maximizing (28) with respect to σ^2 yields the well-known estimator

$$\widehat{\sigma}^2 = \frac{1}{n}\,\mathrm{RSS}\big(\{\hat{X}_c\}\big). \tag{31}$$

Evaluating the log-likelihood (28) at the MLE $\{X_c\} = \{\hat{X}_c\}$ and $\sigma^2 = \hat{\sigma}^2$ gives

$$\begin{aligned}\ln \mathcal{L}\big(\{\hat{X}_c\} \mid \mathcal{D}\big) &= -\frac{n}{2}\ln\big(2\pi\hat{\sigma}^2\big) - \frac{1}{2\hat{\sigma}^2}\,\mathrm{RSS}\big(\{\hat{X}_c\}\big) \\ &= -\frac{n}{2}\ln\big(2\pi\hat{\sigma}^2\big) - \frac{n}{2}, \end{aligned} \tag{32}$$

where the second line follows from (31).

Substituting (32) into the BIC definition (22) yields the following closed form

$$\begin{aligned}\mathrm{BIC} &= \ln(n)\,k - 2\ln \mathcal{L}\big(\{\hat{X}_c\} \mid \mathcal{D}\big) \\ &= \ln(n)\,k + n\ln\big(2\pi\widehat{\sigma}^2\big) + n \\ &= \ln(n)\,k + n\ln\left(\frac{2\pi\mathrm{RSS}\big(\{\hat{X}_c\}\big)}{n}\right) + n. \end{aligned} \tag{33}$$

For a given dataset, among the models identified with varying threshold levels λ, the one with the lowest BIC value is chosen as the most compact yet informative representation of the underlying dynamics.

3.5 Trajectory reconstruction and validation

The identified hydrodynamic coefficients are substituted back into the Abkowitz equations (10) to simulate the vessel trajectories. These simulated trajectories are compared against the reference trajectories generated using the original coefficients to assess the fidelity of the identified models.

Model performance is quantified using both trajectory-level errors and coefficient-level deviations. The trajectory-level error is computed as the mean square error (MSE) between the simulated and reference positions in the x-y plane where as the coefficient-level deviations quantifies the difference between the identified and original hydrodynamic derivatives. This dual evaluation ensures that the identified model not only reproduces the observed motion accurately over time, but also captures the underlying physical parameter relationships.

4 Data content analysis

This section examines the characteristics of the datasets used for model estimation, with emphasis on their excitation content and coverage in the input space. From an autonomy perspective, the richness of these datasets determines whether the identified model can generalize to diverse operating conditions encountered during real-world missions. Autonomous vessels frequently operate outside the limited regimes represented by standard maneuvers, making it essential to evaluate whether such datasets sufficiently explore the system dynamics. This section analyzes the composition and diversity of standard maneuver datasets and introduces a new excitation input, termed the *Dynamic Excitation Maneuver* (DEM), designed to provide richer dynamic variations.

In the context of this study, the system identification task aims to approximate the mapping between the input state–control vector $[\Delta u', v', r', \delta]$

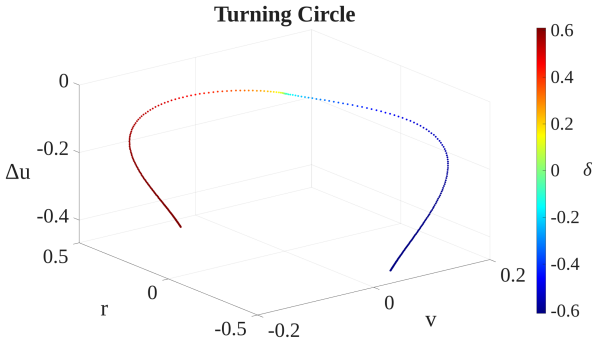


Figure 2: Turning circle maneuver trajectory in input space

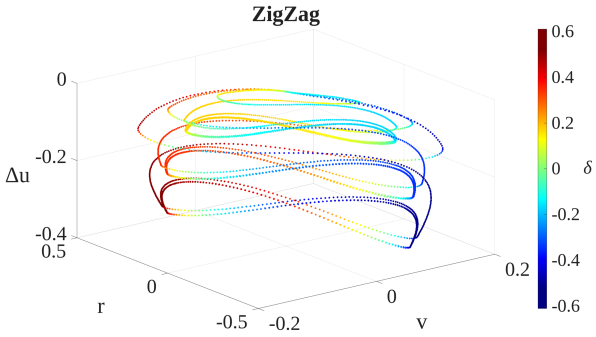


Figure 3: Zig Zag maneuver trajectory in input space

and the corresponding hydrodynamic forces and moment. Sufficient excitation is essential for reliable parameter estimation and function approximation. Most prior studies shown in Table 1, rely on a limited set of standard free-running maneuvers such as turning circles, zig-zags, and spirals. To evaluate the richness of such datasets, these maneuvers are simulated using the Abkowitz model for the KCS vessel. The trajectories of these maneuvers in the input space ($\Delta u', v', r', \delta$) is shown in Figure 2, Figure 3 and Figure 4. The plots here are shown for the KCS vessel and they are similar for the Mariner vessel. In these plots $\Delta u', v', r'$ are the non-dimensionalized velocities and they span along the coordinate axes as shown, while δ (in radians) is represented by the colour bar. The composition of the trajectories shown is as follows :

1. Figure 2 contains 2 turning circle maneuvers, $\delta = -35°$ and $\delta = 35°$. Each trajectory consists of 1500 datapoints.

2. Figure 3 contains 6 zig-zag maneuvers of the form zig-zag ψ/δ and zig-zag $-\psi/-\delta$. Each trajectory consists of 1500 datapoints.

3. Figure 4 contains 2 spiral maneuvers. One increasing and one decreasing in the range $[-35°, 35°]$ at every $5°$ spacing. Each commanded rudder angle is kept constant for 100 non dimensional time steps and each trajectory consists of 1500 datapoints.

Figures 2, 3, and 4 show that the turning circle, zig-zag, and spiral maneuvers occupy distinct and localized regions of the input space. Under

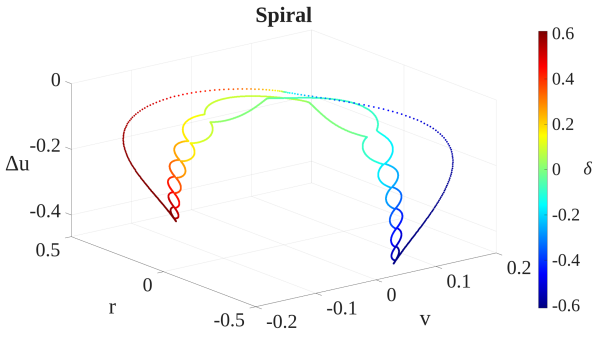


Figure 4: Spiral maneuver trajectory in input space

the given operating conditions there is a bounded set of velocity combinations $(\Delta u', v', r')$ that are physically attainable by the vessel and this set is defined as the admissible input space. Each maneuver excites the vessel over a narrow and structured subset within this admissible input space. For example, the turning circle produces a localized arc trajectory, with very little spread and the trajectory remains confined to a thin, nearly one-dimensional curve in the input space. The zig-zag maneuver covers a larger region of the input space compared to the turning circle, but its coverage remains highly structured. Each rudder reversal generates a loop that lies close to the previous ones, producing a repeated family of thin looped arcs. The trajectories exhibit a wider spread in the v'–r' plane while the variation along the $\Delta u'$ direction remains limited. Each zig–zag run produces a loop that occupies only a narrow band of $\Delta u'$, and although different zig–zag maneuvers lie in slightly different $\Delta u'$ ranges, none of them traverse the broader span of admissible surge variation. Furthermore, the $\Delta u'$ variability that does occur is concentrated near the boundaries of the admissible input space rather than within its central region, leaving large portions of the admissible domain unexplored. Thus, although the zig-zag explores more variability, it still lacks broad variability across the wider input domain. The spiral maneuver occupies roughly the same boundary region of the admissible input space as the turning circle, but exhibits additional feature. As the rudder angle is ramped gradually during the spiral, the trajectory shows a smooth variation in surge deviation $\Delta u'$ along the arc, which can be seen in Figure 4 as a continuous change in colour across the curve. This gradual $\Delta u'$ modulation provides transitions in the surge direction that are absent in the turning circle and zig–zag maneuvers. But the spiral remains confined to a narrow region near the admissible-space boundary and does not sweep through the central portions. Thus, although the spiral contributes continuous $\Delta u'$ variation that complements the looped and oscillatory patterns of the other maneuvers, it alone is insufficient to provide broad coverage of the admissible input space.

The trajectories generated by the standard maneuvers show limited connectivity within the admissible input space, in the sense that they do not provide transitions that span broadly across all three velocity dimensions. For a nonlinear maneuvering system, a transition between two regions of the input space corresponds to how the vessel's attainable $(\dot{u}', \dot{v}', \dot{r}')$ vary as (u', v', r') and the rudder angle δ change. At a given state (u', v', r'), only a finite set of accelerations is physically realizable, and different parts of the state space admit different ranges of achievable accelerations. A rich dataset would therefore require data that traces trajectories through regions where these attainable accelerations also exhibit substantial variability along with the variability in input space. The standard maneuvers considered here do not provide such variability. The turning circle remains confined to a thin arc with little variation in any velocity component. The zig–zag maneuver introduces wider variation in the v'–r' plane but remains narrow in the $\Delta u'$ direction, thereby exploring only a limited subset of the accelerations the system can produce. The spiral maneuver adds smooth variation in $\Delta u'$ but still remains close to the boundary of the admissible input space. Across all three maneuvers, large portions of the interior of the admissible domain remain unexplored, and the data does not include transitions that move simultaneously across broad ranges of u', v', and r'.

Such limitations in the variability of the data can influence the quality of the learned model. Since the identification uses only the data collected within the regions of the input space that the maneuvers explore, the estimated model is shaped by the dynamics exhibited in those regions and remains largely unconstrained outside them. As a result, the model is expected to reproduce the motions contained in the training maneuvers, but its behavior outside these regions depends on how well the local structure of the data reflects the broader dynamics of the vessel. When the data does not include transitions that span multiple directions in the input space, it becomes difficult to ascertain whether the identified model generalizes reliably to operating conditions that were not represented in the dataset. These considerations become relevant in practical applications. For example, reinforcement-learning agents trained in simulation depend on the fidelity of the underlying dynamic model to experience realistic vessel responses. An additional advantage of obtaining a more accurate maneuvering model is that it enables the use of dynamics–based state estimators, such as extended

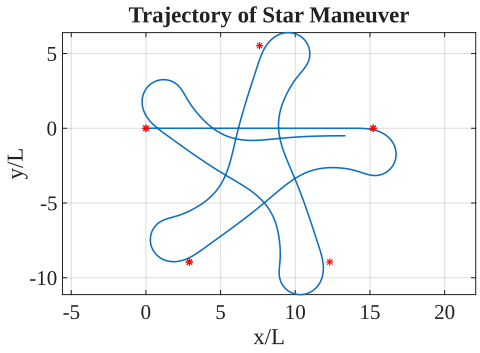


Figure 5: Trajectory of the star maneuver in x-y plane.

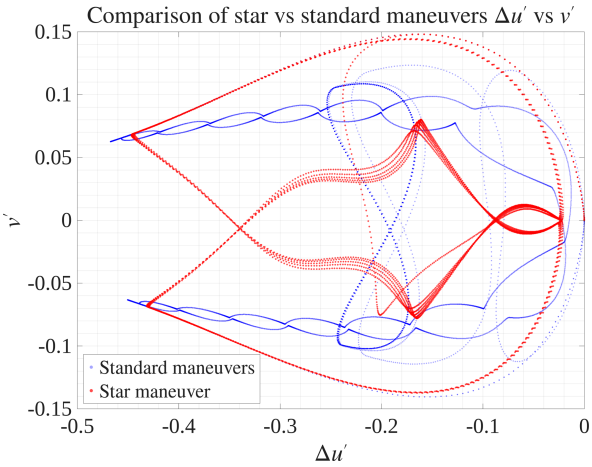


Figure 6: Comparison of star vs standard maneuvers : $\Delta u'$ vs v'

Kalman filters, even when only low–cost or comparatively noisy sensors are available. A reliable process model allows the filter to compensate for reduced measurement quality, thereby maintaining consistent state estimation despite sensor limitations. In both cases, the utility of the identified model is linked to how well the training data captures the diversity of dynamics encountered in practice, although establishing this connection rigorously would require further analysis beyond the present scope.

To demonstrate that trajectories can exist which generate transition patterns not present in the standard maneuver set, a non–standard excitation is considered in the form of a "star" trajectory as shown in Figure 5. The relevant behavior is captured in the input space $(\Delta u', v', r', \delta)$ and projected onto two–dimensional subspaces of the input space. The resulting transitions differ qualitatively from those observed in the turning circle, zigzag, and spiral maneuvers as shown in Figure 6 and Figure 7. In particular, the star trajectory contains segments where the vessel simultaneously exhibits changes in all three velocity components together. This behavior is different from the trajectories of standard maneuvers as zigzags trajectory shows more variation in v' and r' compared to $\Delta u'$, whereas the spiral maneuver shows more variation along $\Delta u'$ compared to v' or r'. The trajectory of star maneuver also shows transitions which are structurally different from standard maneuvers. This difference is due to the variation in the rudder angle excitation in each of the maneuvers as shown in Figure 8. Turning circles maintain a near-constant rudder deflection, spirals slowly vary amplitude of rudder in steps, and zig-zags exhibit abrupt but relatively symmetric reversals. On the other hand, star maneuvers includes partial reversals of δ that do not complete a full switch to the opposite extreme. This results in a trajectory that allows exploration of the interior of the admissible domain. Such patterns are not generated by the standard maneuvers, whose transitions arise from their fixed structural protocols. These observations are not intended as an evaluation of the standard maneuver set, but rather as a constructive indication that alternative maneuver designs may access regions of the admissible input space which are not produced by the conventional protocols. The variation in the excitation motivates the development of a systematic maneuver that combines the established transition modes of the standard set with these additional switching behaviors.

To address the limited exploration of the input space with standard maneuvers, a new maneuver is designed — referred to as the *Dynamic Excitation Maneuver* (DEM). The DEM is designed with the objective of producing structured yet diverse transitions in the input space, thereby inducing variations in $(\Delta u', v', r')$ that are not obtained from standard maneuvers. This maneuver combines multiple desirable characteristics. It features alternating sign changes in rudder input, which are characteristic of zig-zag maneuvers, and it incorporates monotonous step amplitude variation, similar

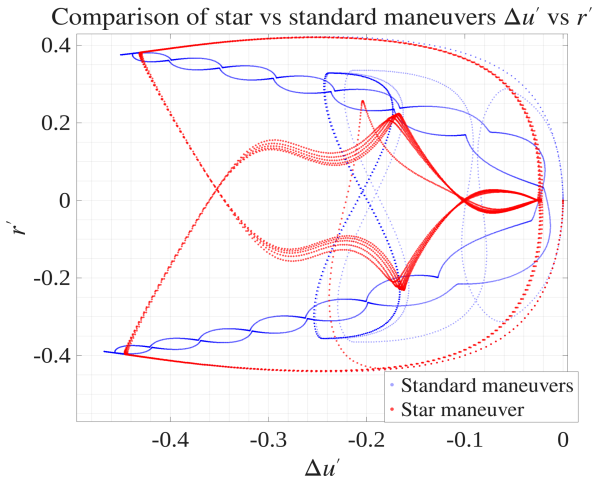


Figure 7: Comparison of star vs standard maneuvers : $\Delta u'$ vs r'

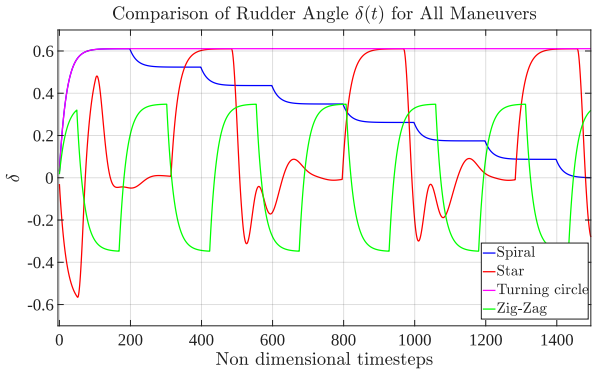


Figure 8: Comparison of rudder angle excitation of different maneuvers

to spiral maneuvers. Additionally, it includes intervals of constant rudder deflection, as seen in turning circle maneuvers. The maneuver also includes transitions with rudder deflection switch from extremes to midway on opposite sign and back, a feature similar of the star maneuver.

4.1 Construction of the DEM signal.

The DEM rudder sequence is generated by defining a set of increasing or decreasing rudder amplitudes. The construction is explained for the increasing rudder variation and follows similar process for the decreasing variation. First a sequence of increasing rudder amplitudes is created as shown

$$L_\Delta = \{\, \delta_{\min},\, \delta_{\min} + \Delta\delta,\, \delta_{\min} + 2\Delta\delta, \ldots, \delta_{\max} \,\},$$

where $\delta_{\min} = -35°$ and $\delta_{\max} = 35°$ and $\Delta\delta = 5°$. Next, an alternating sign pattern on successive elements of the sequence is applied. This yields a list of target angles

$$\delta_i = (-1)^{\,i-1}\, |L_{\Delta_i}|, \qquad i = 1, \ldots, n,$$

where n is the number of elements in L_Δ which introduces left–right switching analogous to a zig–zag maneuver but with systematically increasing amplitude, similar to a spiral maneuver. To incorporate both full–deflection and intermediate transitions, each δ_i is expanded into a four-element block,

$$\{\, \delta_i,\, -\tfrac{1}{2}\delta_i,\, \delta_i,\, 0 \,\},$$

resulting in a refined excitation pattern containing full-amplitude deflections, half-amplitude opposite-sign switches, and repeated short intervals of constant deflection. These segments collectively emulate features of turning circles, spirals, and zig–zags, while introducing additional mid-range switching behavior that is absent in standard maneuvers. The rudder signal for the DEM is depicted in Figure 9, and the corresponding input space coverage is shown in Figure 10. Variants of DEM can be generated by changing the number of non dimensional timesteps for which the rudder angle is kept constant in the input signal. For the rudder signal in Figure 9 the rudder angle is kept constant for 50 non dimensional timesteps.

The DEM structure is motivated by the insight that different patterns of rudder variation produce distinct regions and transitions in the input space. Standard maneuvers tend to excite

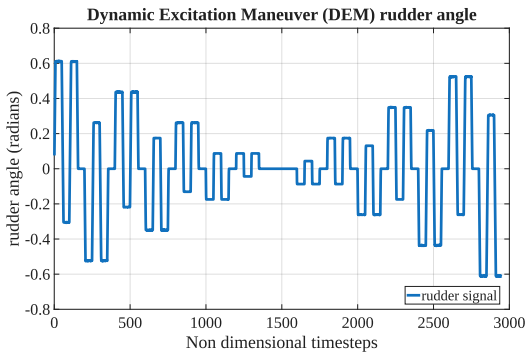


Figure 9: Dynamic Excitation Maneuver (DEM) rudder angle signal

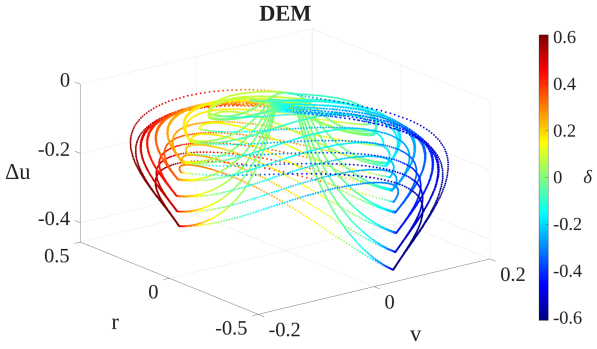


Figure 10: Dynamic Excitation Maneuver (DEM) span in input space

only a subset of these transitions. Zig–zag maneuvers introduce large variations in v' and r' but yield limited surge variation, spiral maneuvers generate monotonic changes in $\Delta u'$ with comparatively little sway–yaw coupling, and turning circles produce nearly one-dimensional input trajectories. The DEM is constructed specifically to merge these behaviors while incorporating additional mid-range switches. These mid-range switches cause the vessel to traverse directions in the $(\Delta u', v', r')$ space that are not accessed by standard maneuvers, thereby producing more diverse local transitions between neighboring velocity states. The introduction of both full-deflection and partial-deflection switches encourages the vessel to repeatedly move between different dynamic regimes, leading to a varied input–output data for model identification. Overall, the DEM is generated in this manner to ensure that the resulting dataset contains a broader variety of velocity transitions than those obtainable from standard maneuvers, while retaining a structured design that is simple to implement and compatible with safe experimental operation. The subsequent section analyzes how differences in data richness influence the accuracy and generalization of the identified maneuvering models.

5 Quantification of Dataset Richness

This study aims to demonstrate that incremental improvements in dataset richness lead to more accurate and stable system identification. For this purpose, three datasets of varying levels of excitation and diversity are considered. The exact list of maneuvers used in each of the thee datasets is given in Table 3 but can be broadly described as follows :

1. *Standard Maneuvers (Literature)*: The set of conventional standard maneuvers typically used in ship maneuvering identification studies.

2. *Standard Maneuvers (Extended)*: An extended collection of standard maneuvers considering additional zig-zags and turning circles combining all the standard maneuvers used in training in prior studies in Table 1

3. *Dynamic Excitation Maneuver (DEM)*: The newly proposed dynamic excitation maneuver, described in section 4, designed specifically to maximize state-space transitions and input excitation diversity.

All commanded rudder inputs, including those used in the DEM, are passed through the actuator dynamics described in subsection 2.2. As a result, the realized rudder motion is subject to rate saturation and finite response time, ensuring that the effective excitation remains physically realizable. High-frequency components present in the commanded rudder signal are naturally attenuated due to the finite time constant and rate limits of the steering gear. Consequently, the generated maneuvering data are consistent with the steering actuator assumptions adopted in this study and avoid unrealistically high-frequency rudder commands.

As established in the previous section, a dataset can be considered *richer* if it exhibits more variation in state transitions and greater occupancy in the admissible input space. To quantify these properties, two complementary metrics are employed: entropy and the condition number of the regressor matrix.

5.1 Entropy as a Measure of Data Spread

Entropy originates from information theory and quantifies the level of uncertainty or unpredictability in a system. Entropy measures the expected amount of information needed to describe the state of a random variable, considering distribution of probability across all potential states. In the context of dataset distribution, higher entropy implies that the samples are distributed across a larger portion of the available space, indicating a richer exploration of system dynamics. Hence a dataset with higher entropy spans a larger portion of the admissible state space and captures more diverse state transitions. The term "admissible state space" here refers to the common discretized state-space region obtained from the union of all simulated trajectories considered in this study. It should therefore be interpreted as a simulation-based approximation of the explored operating region rather than the complete physically reachable state space of the vessel.

In this study the entropy is computed over the seven-dimensional space formed by the state–derivative pair given by $\left[\Delta u', v', r', \delta, \dot{u}', \dot{v}', \dot{r}'\right]$. The derivatives of velocities are included in the entropy computation to include the representation of state transitions also. The space is discretized into bins to form a multidimensional histogram, and the entropy is calculated as

$$H = -\sum_i p_i \log_{10} p_i, \tag{34}$$

where p_i represents the normalized occupancy of the i^{th} bin, computed as the ratio of the number of samples in that bin to the total number of samples in the dataset. The entropy is evaluated over the discretized state space defined by the common histogram grid, ensuring that entropy values remain directly comparable across datasets with differing levels of state-space coverage. Bins that are unoccupied in a given dataset are assigned zero probability and therefore contribute zero to the entropy calculation, consistent with the limiting relation which follows from an application of L'Hôpital's rule.

$$\lim_{p_i \to 0^+} p_i \log p_i = 0$$

An important case to note is that different maneuvers may occupy different regions of the admissible state space. In some cases, a region that is physically admissible may remain unvisited in one dataset while being explored by another. Excluding such regions separately for each dataset would ignore the fact that these states are dynamically realizable and would make the entropy values dependent on the particular support of each dataset. Consequently, a maneuver with limited but uniformly distributed coverage over its own occupied region could yield an artificially high entropy despite exploring only a small portion of the admissible dynamics. By using a common discretization constructed from the union of regions explored across all datasets, the entropy metric provides a consistent basis for comparison while being physically meaningful.

In this study, 20 bins are used for each of the non dimensional states and their derivatives, while 14 bins are used for the rudder angle δ, corresponding to 5° intervals. This discretization ensures sufficient resolution to capture the variability

Table 3: Summary of datasets used for model identification. Each dataset corresponds to an equal-length simulation record for both vessels.

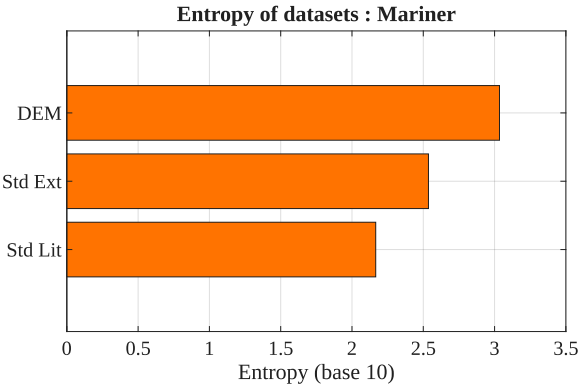
Dataset	Maneuvers Included	Total Datapoints
Standard Maneuvers (Literature)	Turning circle 35° port, Turning circle 35° starboard, Zig-zag 10–10, Zig-zag 20–20	9,000
Standard Maneuvers (Extended)	Turning circle 35° port, Turning circle 35° starboard, Zig-zag 10–10, Zig-zag 20–20, Zig-zag 30–30, Turning circle 25° port, Turning circle 25° starboard, Zig-zag 5–5, Zig-zag 10–5, Zig-zag 20–5, Zig-zag 30–5	27,000
Dynamic Excitation Maneuver (DEM)	Two DEM variants with constant step size of 200, one ascending from $-35°$ to 35° and the other descending from 35° to $-35°$	23,600

of the data without introducing sparsity artifacts. To enable consistent comparison across different datasets, the bin ranges in each dimension are defined using the global minimum and maximum values computed across all datasets considered in this study. This ensures that the discretized state space is identical for all cases and that differences in entropy reflect variations in data distribution rather than differences in domain definition.

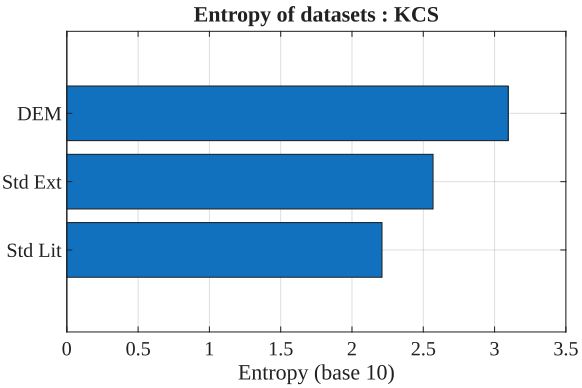
The computed entropy values for each dataset and vessel are shown in Figure 11. It can be seen that the DEM shows the highest entropy for both vessels, confirming that it provides the most diverse set of state transitions and the broadest coverage in the input space.

For entropy calculation, Miller–Madow bias correction (Paninski, 2003) was also applied to account for finite sampling, although its magnitude was below 0.02 in all cases and did not affect the ranking of dataset richness.

To further examine how dataset richness evolves with the amount of available data, an incremental entropy growth analysis is carried out. Instead of computing entropy over the entire dataset, the metric is evaluated as a function of the number of sampled datapoints. For computational tractability in the seven-dimensional state–derivative space, a randomized subsampling strategy is used. The entropy is recomputed after every 500 datapoints up to the first 3000 samples, and subsequently after increments of 3000 datapoints. At each increment, a random subset of the corresponding size is drawn without replacement from the full dataset,

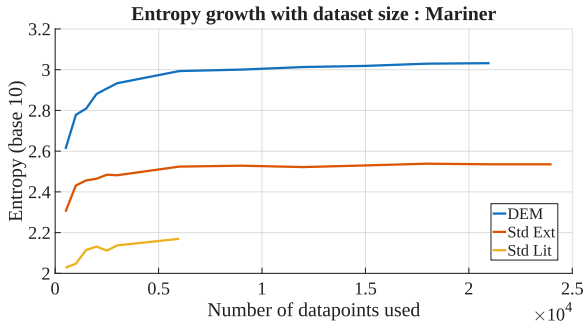


(a) Entropy for datasets : Mariner

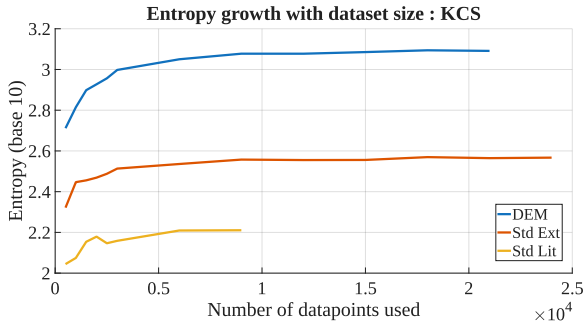


(b) Entropy for datasets : KCS

Figure 11: Entropy calculated for various datasets for both vessels



(a) Entropy growth with data size (randomized sampling): Mariner



(b) Entropy growth with data size (randomized sampling): KCS

Figure 12: Growth of entropy as a function of number of datapoints for all datasets with randomized sampling of datapoints.

preventing bias due to temporal ordering or trajectory grouping. This procedure makes it possible to assess how rapidly each dataset begins to populate the admissible state space and how its exploration expands as more datapoints are introduced.

Figure 12 shows the resulting entropy growth curves for all datasets and for both vessels. In both the Mariner and KCS cases, the DEM dataset exhibits the steepest initial rise in entropy and consistently maintains the highest entropy across all data volumes. This indicates that DEM not only spans a broader region of the feasible state space when considered in full, but also begins to explore this space more rapidly as points are added. In contrast, the Standard Maneuvers Extended and Standard Maneuver Literature datasets, display slower entropy growth and saturate at lower levels, reflecting their constrained coverage of the state–transition space.

Following the random-sampling methodology,

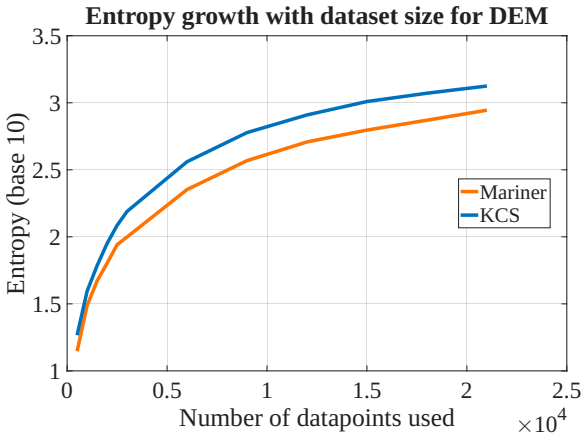


Figure 13: Growth of entropy as a function of the number of datapoints for the DEM dataset using sequential sampling. Results are shown for both the Mariner vessel (orange) and the KCS vessel (blue).

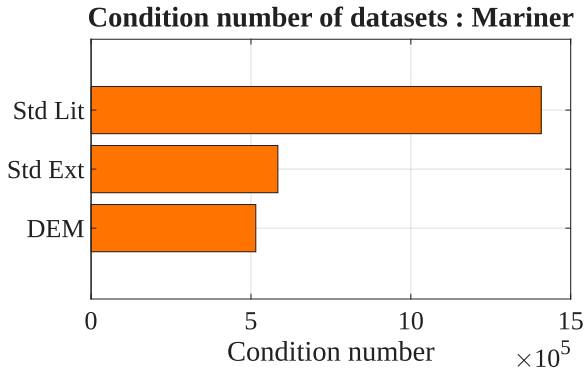
the DEM dataset is examined to analyze how entropy improves as datapoints are added sequentially with the same temporal order of the execution of the maneuver. When datapoints are introduced in chronological order rather than through randomized sampling, the DEM entropy increases smoothly and monotonically with the number of datapoints, demonstrating that the richness of the dataset is not concentrated in isolated segments of the dataset but is consistently present throughout its trajectories.

5.2 Condition Number as a Measure of Regressor Independence

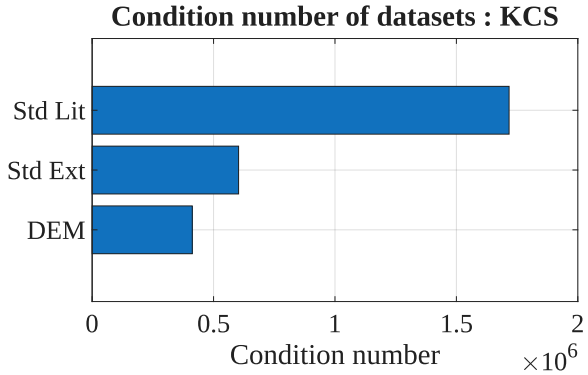
A second indicator of dataset richness is the condition number of the regression matrix $[A]$ used in model identification as shown in (17a). The condition number is defined as the ratio of the largest to the smallest singular value of the matrix:

$$\kappa(X) = \frac{\sigma_{\max}(X)}{\sigma_{\min}(X)}. \tag{35}$$

Traditionally, the condition number serves as a numerical indicator of collinearity among the linearly dependent variables of regression. A lower condition number indicates that the variables are not dependent on each other and that the matrix is better conditioned, allowing for more stable and



(a) Condition number for datasets : Mariner



(b) Condition number for datasets : KCS

Figure 14: Condition number calculated for various datasets for both vessels

reliable parameter estimation. Conversely, a high condition number indicates redundancy or insufficient excitation among the variables used in regression.

The condition numbers for each dataset and vessel are presented in Fig. 14. It is observed that the DEM has the lowest condition number, and hence provides a more balanced and well-excited regression space, with less collinearity among regressors.

5.3 Summary

Based on the above two complementary measures—entropy and condition number—it can be concluded that the DEM dataset is the richest among the three considered. The higher entropy indicates broader excitation coverage, while the lower condition number demonstrates greater independence among regressors. In the following section, the impact of this increased richness on model identifiability and generalization performance is analyzed.

6 Results and discussion

This section presents how varying the richness of excitation data fundamentally alters the outcome of maneuvering model identification. The analysis focuses on examining how the choice of excitation signals and dataset richness influences model accuracy, parameter recovery, and generalization capability to unseen trajectories. Three categories of maneuver datasets are considered in this study as shown in Table 3. The comparative assessment with each dataset is carried out for both KCS and Mariner vessels.

6.1 Overview of Evaluation Framework

To assess the influence of dataset richness on the quality of system identification, three distinct maneuver datasets are used. For each maneuver dataset, the hydrodynamic coefficients are identified using the model structure described in subsection 3.2 and identification procedure using STLS and BIC score described in subsection 3.3 and subsection 3.4 to ensure a consistent basis for comparison. The resulting models are subsequently validated against a common set of unseen test cases, comprising a 10–35 zig-zag maneuver which is an unseen standard maneuver along with a random step rudder signal and a random rudder angle signal sampled from a uniform distribution.

The random step rudder signal is specifically designed to test the identified models' performance in the interior region of the admissible input space. This signal consists of a sequence of constant-amplitude steps with varying durations and magnitudes, constructed using the following procedure: each step has a duration randomly selected between 50 and 150 non dimensional time units to ensure sufficient steady-state excitation at each amplitude level. The step amplitudes are constrained within ± 35 and are generated such that consecutive steps differ by at least 10, thereby enforcing significant variation between adjacent segments. This minimum jump constraint prevents the signal from

dwelling in localized regions of the input space and promotes broad coverage of the admissible rudder angle range. The motivation for employing this step-based structure stems from the observation that such changes in control input, similar to those encountered in star maneuvers, can effectively probe the interior regions of the admissible state space rather than merely tracing its boundaries. These test cases have been chosen to evaluate the models' generalization capability beyond the specific conditions under which they are identified.

The evaluation is conducted using two complementary metrics. First is the trajectory reconstruction accuracy on unseen test cases. This is quantified by the mean squared error (MSE) between the positions of trajectories simulated using original and identified set of coefficients. Second is the comparison between the estimated and original hydrodynamic coefficients. Both these metrics are evaluated for each dataset for both the vessels being analyzed in this study. Together these metrics provide a comprehensive assessment of how the richness of the dataset influences identified model accuracy and robustness. Finally, for each dataset the Bayesian Information Criterion (BIC) is analyzed to examine the trade-off between model fit and complexity under different regularization levels controlled by the threshold parameter λ in the STLS algorithm.

6.2 Performance with Standard Maneuvers Literature

The *Standard Maneuvers Literature* dataset corresponds to the conventional set of maneuvers used in most ship maneuvering identification studies. This dataset consists of 35° port and starboard turning circles, along with zigzag 10–10 and 20–20 maneuvers. It is selected to serve as a baseline for evaluating the influence of dataset richness on system identification. Hydrodynamic coefficients for both for KCS and Mariner vessels are identified using the procedure described in section 3 and the resulting models are subsequently validated against a set of unseen test trajectories to assess their generalization capability. Figure 15 shows the trajectories obtained using predicted coefficients compared against the reference trajectories from the Abkowitz model for three unseen input signals: a 10–35 zigzag maneuver, a random rudder sequence sampled from uniform distribution and random step rudder signal as described in subsection 6.1. The results indicate that the model identified using the Standard Maneuvers Literature dataset is able to capture the 10–35 zigzag maneuver for both the vessels. Since the dataset contains the structured excitation of zigzag maneuver, the identified model is able to give accurate trajectory for a different variant of the signal structured in the same way. For the random step signal, the trajectory accuracy is marginally lower compared to the zigzag case for both vessels. This signal introduces unstructured variations in rudder input that are not represented in the standard maneuvers, and the degradation in performance reflects the model's limited exposure to such patterns during training. For the random uniform test case, the deviation becomes substantially more pronounced. Unlike the random step signal which still contains piecewise-constant segments, the random uniform signal lacks any temporal structure and induces rapid, irregular variations in the commanded rudder angle. This excites the vessel dynamics in a far broader and less predictable manner. Since the identification dataset consists exclusively of structured maneuvering inputs, the resulting model has not encountered excitation of this nature. Consequently, the prediction errors observed for the random uniform signal highlight the restricted generalization capability of models trained solely on standard maneuver datasets.

Figures 16 and 17 present the identified hydrodynamic coefficients using the *Standard Maneuver Literature* dataset for the KCS and Mariner vessels, and compare them against the original reference values. The deviations are particularly pronounced for higher-order nonlinear terms and cross-coupled hydrodynamic coefficients. While the linear terms are generally recovered with moderate accuracy, the dataset shows consistently poor identification of nonlinear damping and cross coupled interaction terms. These mismatches directly contribute to the trajectory errors observed in Figure 15.

An important observation from the bar plots is the lack of sparsity in the identified models. Even though the optimal thresholding parameters selected using BIC are small up to the order of 10^{-3}, the resulting models do not exhibit the expected sparse structure. This is especially evident in the

surge equation for both the vessels where physically, only a few coefficients should be active, yet the identified model retains a large number of coefficients with non-zero magnitude. This behavior reflects a fundamental identifiability limitation of the standard maneuvers. Since the inputs excite only a restricted portion of the dynamics, many terms appear weakly correlated with the outputs and the estimator is unable to suppress the unnecessary coefficients. This is also seen with the high value of condition number for Standard Maneuver Literature dataset in Figure 14b and Figure 14a. Several nonlinear and coupled coefficients display either inflated magnitudes or incorrect signs. This indicates that the excitation provided by the standard maneuvers is insufficient not only for sparsification but also for determining the correct functional dependence of the dynamics.

Overall, this analysis highlights a critical limitation of using standard maneuver datasets for model identification. Although models trained on these datasets can reproduce similar structured standard maneuvers with reasonable accuracy, they fail to generalize beyond them. Most literature describing system identification of maneuvering hydrodynamic coefficients usually also test accuracy on trajectories similar to those used in training. The apparent correctness obtained when validating only on standard maneuvers may therefore create a false sense of accuracy, masking the poor recovery of underlying dynamics and the lack of robustness under more varied or unstructured inputs.

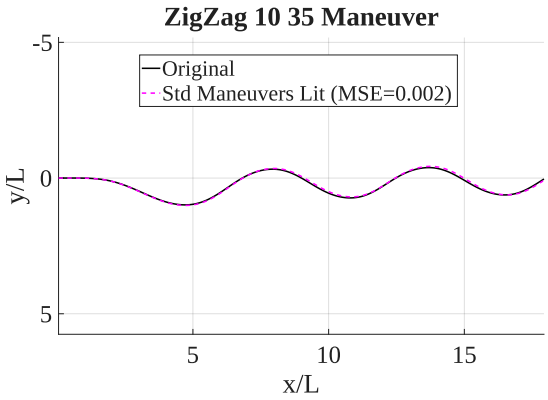
6.3 Improvement with Extended Standard Maneuver Set

To investigate whether the excitation limitations of the conventional maneuver set can be alleviated, an extended dataset comprising additional zigzag and turning circle maneuvers was considered. This *Standard Maneuvers Extended* dataset preserves the general structure of the standard tests while broadening the excitation spectrum. The identification procedure follows the framework outlined in section 3.

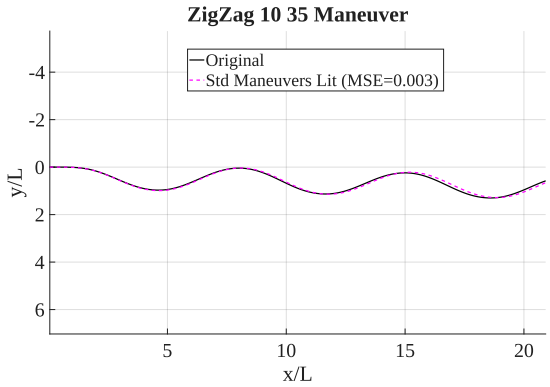
Figure 18 compares the reconstructed trajectories obtained using the models identified from the *Standard Maneuvers Literature* and *Standard Maneuvers Extended* datasets for both vessels. The use of the extended dataset results in visibly improved trajectory fidelity for the unseen test cases. The improvement is more prominent in the KCS vessel, where trajectory drift observed for random uniform test case is significantly reduced compared to the literature dataset. For the Mariner vessel, improvements are present but less pronounced, with some residual offsets in lateral position still observed for the random uniform test case. This behavior is consistent with its stronger nonlinear coupling. For both the vessels, marginal improvement is observed for the random step test case because of the additional input space coverage by the extended dataset.

To quantify the improvement in trajectory reconstruction, the average MSE over 50 random rudder inputs sampled from a uniform distribution is compared for the *Standard Maneuvers Literature* and *Extended* datasets. The results are summarized in Table 4. For both the KCS and Mariner vessels, the mean squared trajectory error improves by approximately 90% and 92% respectively with the extended dataset, demonstrating improved generalization to unseen inputs. Though the percentage improvements in the error are significant, the coefficient comparison between the identified and reference coefficients are required for complete analysis.

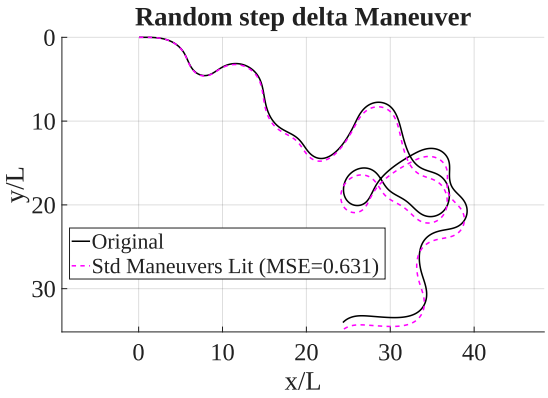
The coefficient comparisons in Figures 19 and 20 highlight different trends observed from trajectory-based validation. For the KCS vessel, the identified coefficients align more closely with the reference values, especially the surge hydrodynamic coefficients. In contrast, the Mariner vessel coefficients exhibit noticeable deviations in several higher-order nonlinear terms. The BIC-selected thresholding parameters summarized in Table 5 provide additional insight into the structure of the identified models. For the KCS vessel, the optimal thresholding parameter (λ) values increase when moving from the *Standard Maneuvers Literature* to the *Extended* dataset, particularly in the surge and sway modes. The larger λ values indicate that the additional excitation present in the extended dataset enables the BIC criterion to suppress weak or spurious coefficients more effectively. This behavior results in a noticeably sparser and more physically consistent coefficient set, with the surge dynamics closely matching the true sparsity pattern and marginal improvements also visible in the yaw dynamics. For sway dynamics, significant parameter drift is still present as seen with third order coeffi-



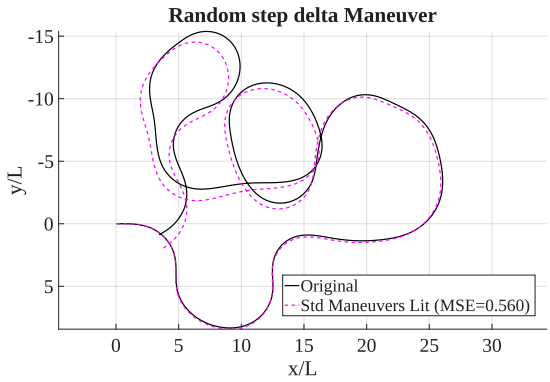
(a) Zig Zag 10-35 : KCS



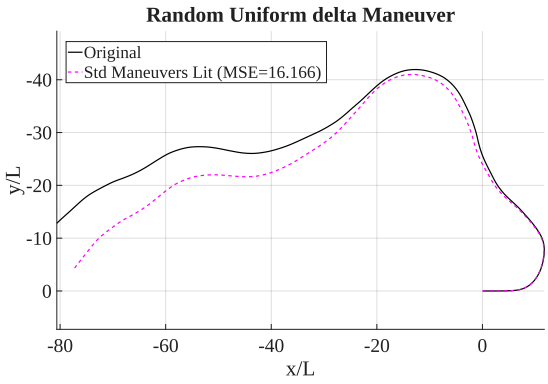
(b) Zig Zag 10-35 : Mariner



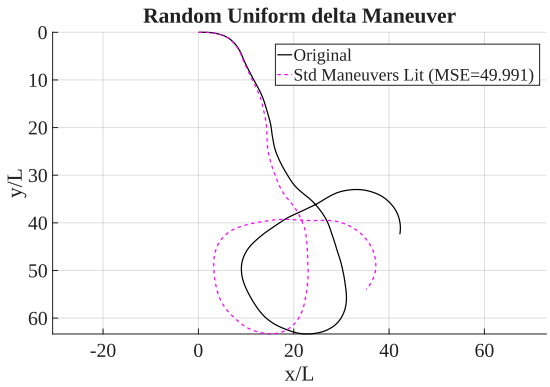
(c) Random step rudder signal : KCS



(d) Random step rudder signal : Mariner

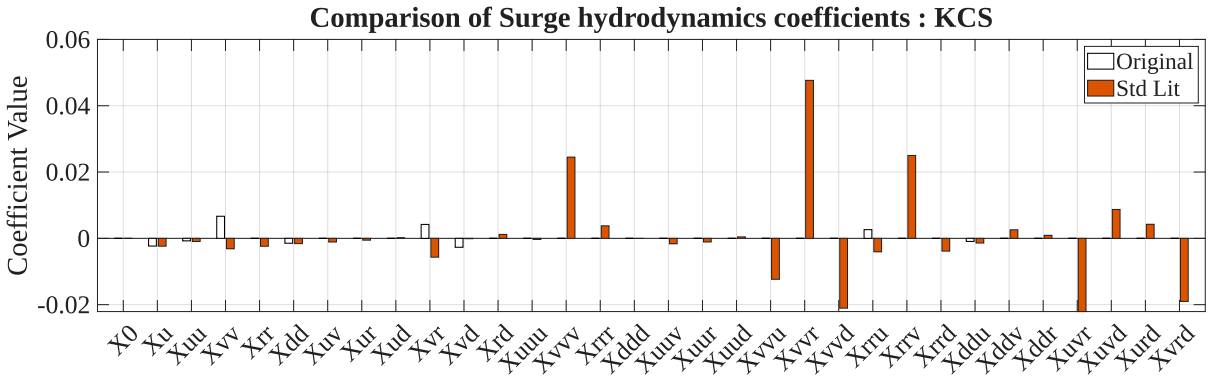


(e) Random uniform rudder signal : KCS

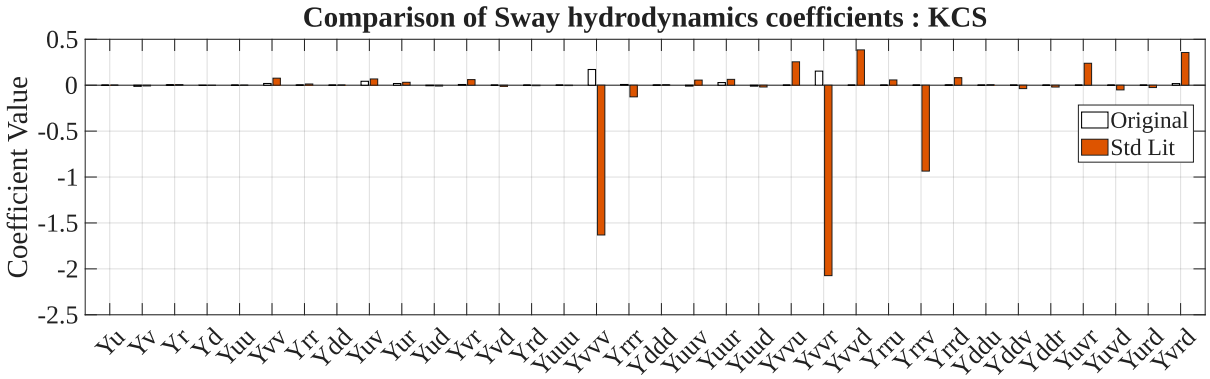


(f) Random uniform rudder signal : Mariner

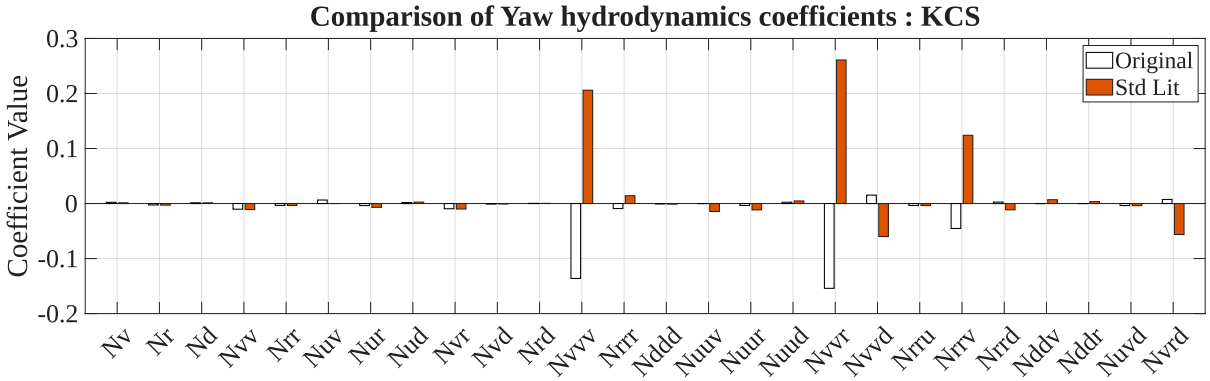
Figure 15: Trajectory reconstruction performance for the KCS (left) and Mariner (right) vessels using the *Standard Maneuvers Literature* dataset. Predicted (dashed) and reference (solid) trajectories are shown for (a,b) Zig Zag 10-35, (c,d) random step rudder signal, and (e,f) random uniform rudder signal.



(a) Optimal thresholding parameter for surge $\lambda = 0.00166$

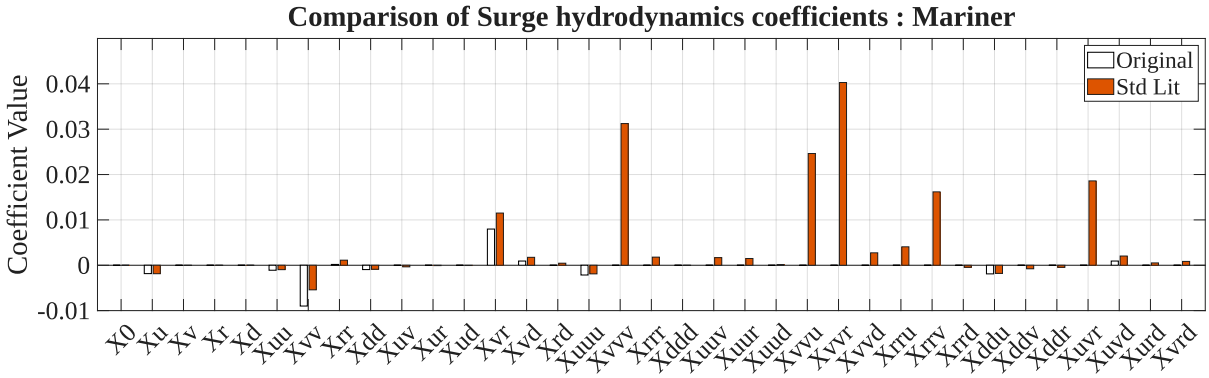


(b) Optimal thresholding parameter for sway $\lambda = 0.001$

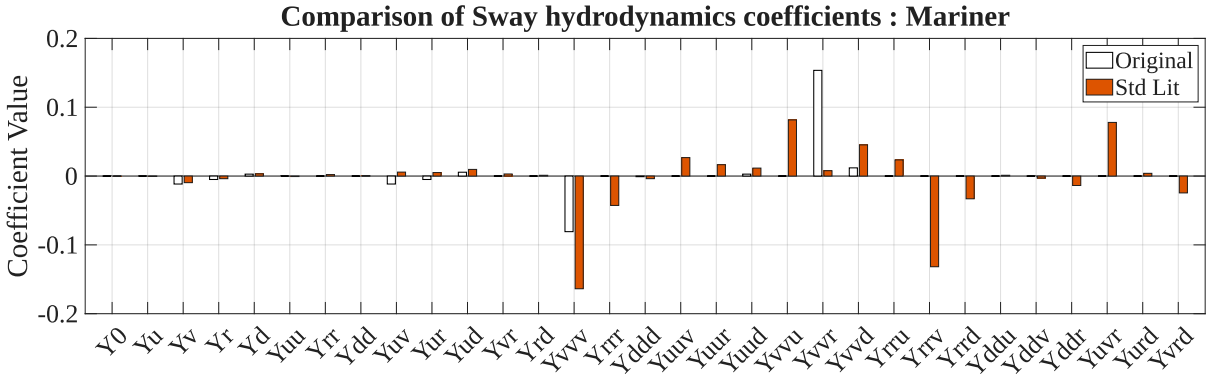


(c) Optimal thresholding parameter for yaw $\lambda = 0.0359$

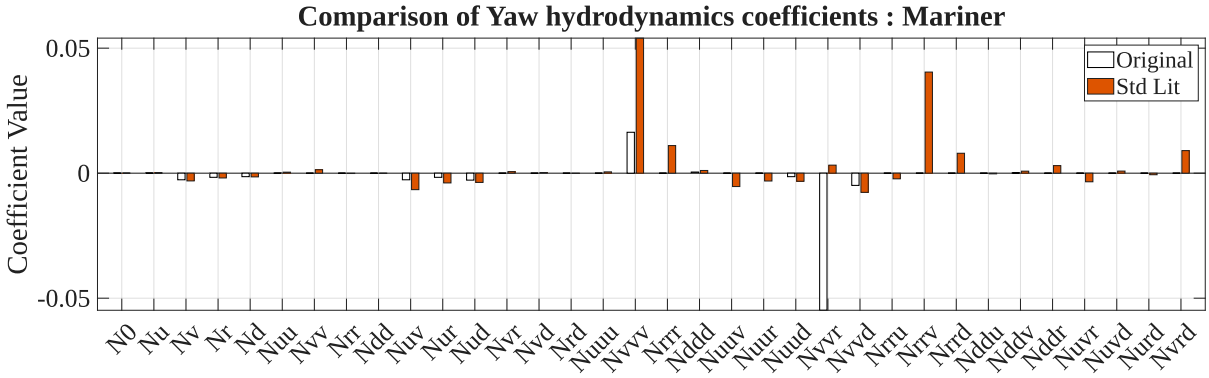
Figure 16: Comparison of identified vs. reference hydrodynamic coefficients for the KCS vessel using the *Standard Maneuvers Literature* dataset.



(a) Optimal thresholding parameter for surge $\lambda = 0.001$



(b) Optimal thresholding parameter for sway $\lambda = 0.0215$



(c) Optimal thresholding parameter for yaw $\lambda = 0.0215$

Figure 17: Comparison of identified vs. reference hydrodynamic coefficients for the Mariner vessel using the *Standard Maneuvers Literature* dataset.

Table 4: Mean squared trajectory error for models identified using Standard Maneuvers (Literature) and Standard Maneuvers (Extended) averaged over 50 random rudder signals. The values represented are in terms of non dimensional lengths.

Vessel	**Standard (Lit)**	**Standard (Ext)**
KCS	13.135	1.280
Mariner	94.4459	7.1757

Table 5: BIC–optimal thresholding parameters λ for the Standard Maneuvers Literature and Extended datasets.

KCS Vessel	Surge		Sway		Yaw	
Std Lit λ	1.67 10^{-3}	$\times$	1.00 10^{-3}	$\times$	3.59 10^{-2}	$\times$
Std Ext λ	2.15 10^{-2}	$\times$	1.67 10^{-3}	$\times$	1.29 10^{-2}	$\times$
Mariner Vessel	Surge		Sway		Yaw	
Std Lit λ	1.00 10^{-3}	$\times$	2.15 10^{-2}	$\times$	2.15 10^{-2}	$\times$
Std Ext λ	7.74 10^{-3}	$\times$	1.29 10^{-2}	$\times$	1.00 10^{-3}	$\times$

cient values of $Y_{vvv}, Y_{vvr}, Y_{vvd}, Y_{rrv}$

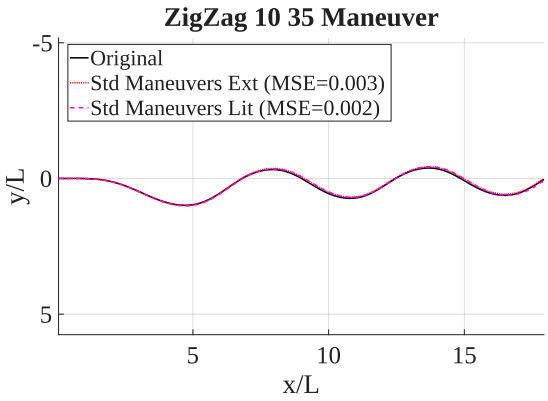
For the Mariner vessel, the changes in optimal λ values are not consistent, and the resulting coefficient sets remain less sparse in all three modes. Several higher-order and cross-coupled hydrodynamic coefficients continue to deviate from the reference values despite the additional maneuvers. While the extended dataset leads to sparsity and accuracy improvements for the KCS vessel, the gains for the Mariner vessel remain limited.

6.4 DEM Identification Performance

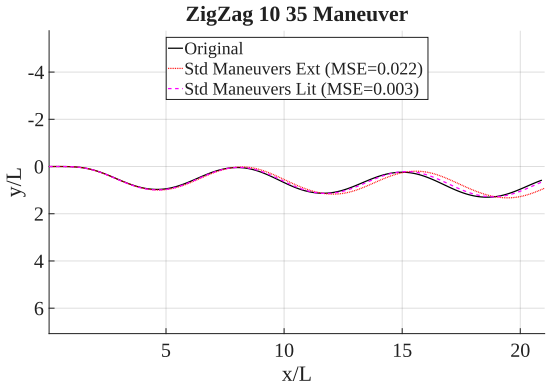
The Dynamic Excitation Maneuver (DEM) dataset has been designed to overcome the limited excitation characteristics of standard maneuvers by combining multiple characteristics of standard maneuvers and additional insights from input space occupancy of different maneuvers. This dataset provides a richer distribution of velocities and accelerations as seen with entropy measure in Figure 11b and Figure 11a. To evaluate the impact of the richer dataset on both trajectory reconstruction and coefficient recovery, the model is identified following the methodology described in section 3. Figure 22 illustrates the trajectory reconstruction results for both the KCS and Mariner vessels using models identified using all three datasets including the DEM dataset. The reconstructed trajectories with coefficients identified using DEM dataset closely follow the reference motion across all test cases, including zig-zag, random step and random rudder inputs from uniform distribution, with negligible drift. For both vessels, the model identified using DEM dataset consistently achieves the lowest errors on unseen trajectories. The improved performance for the test case with random rudder signal can be attributed to better coverage of the admissible input space, leading to more robust estimation of nonlinear cross-coupled hydrodynamic coefficients.

The statistical comparison of average MSE value between trajectory position for the DEM dataset over the same 50 randomly sampled rudder signals as in Table 4 is summarized in Table 6. The analysis shows that for the KCS vessel, the mean squared trajectory error improves by approximately 45% with the DEM dataset over the *Standard Maneuvers Extended* dataset, demonstrating improved generalization to unseen inputs. For the Mariner, the improvement is much more significant around 99% over the *Standard Maneuvers Extended* dataset. The marked improvement across both vessels demonstrates that the quality of excitation in terms of its variability and transition richness is more influential than the quantity of data.

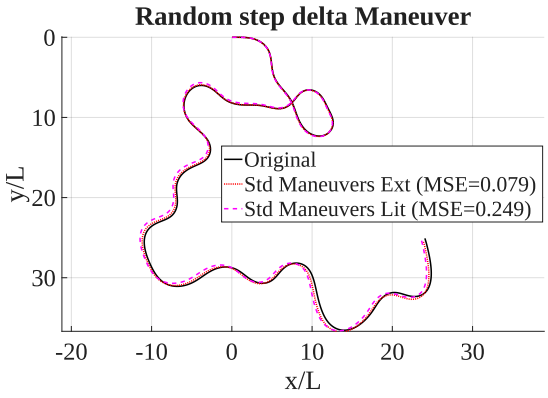
Figures 24 and 25 present the identified hydrodynamic coefficients using the DEM dataset for KCS and Mariner vessel, respectively. Compared to the standard-maneuver extended dataset, the DEM dataset yields coefficient estimates for the KCS vessel that match the reference values more closely, especially for the surge and yaw hydrodynamic coefficients. The third-order sway coefficients do not exhibit significant improvement, even with the increased variability provided by the DEM dataset. This suggests that additional analysis is needed to understand the underlying causes and to refine the identification approach. For the mariner vessel,



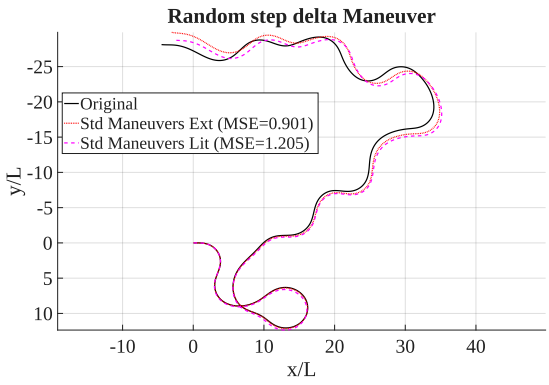
(a) Zig Zag 10-35 : KCS



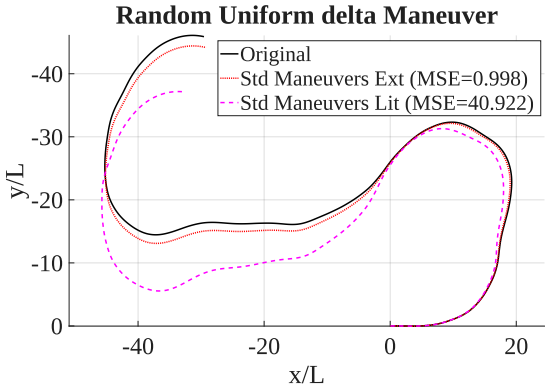
(b) Zig Zag 10-35 : Mariner



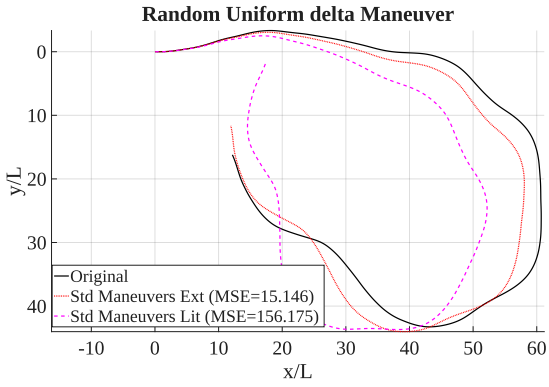
(c) Random step rudder signal : KCS



(d) Random step rudder signal : Mariner

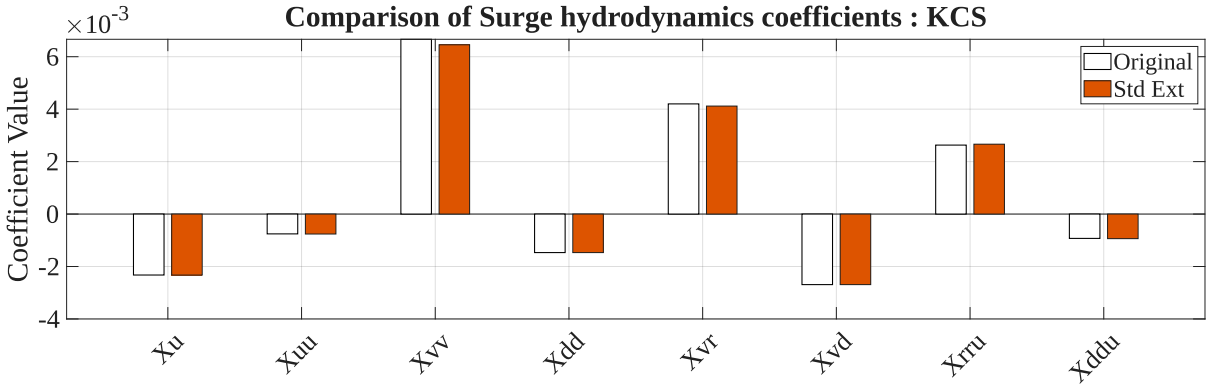


(e) Random uniform rudder signal : KCS

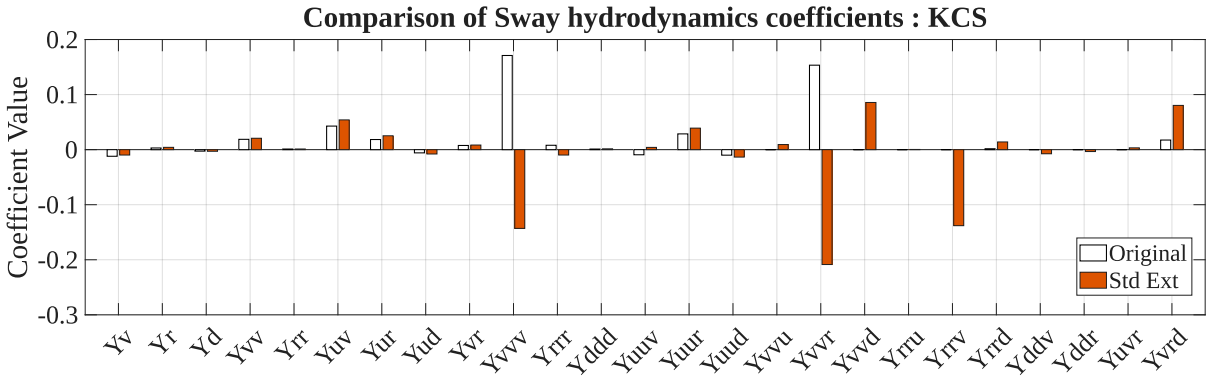


(f) Random uniform rudder signal : Mariner

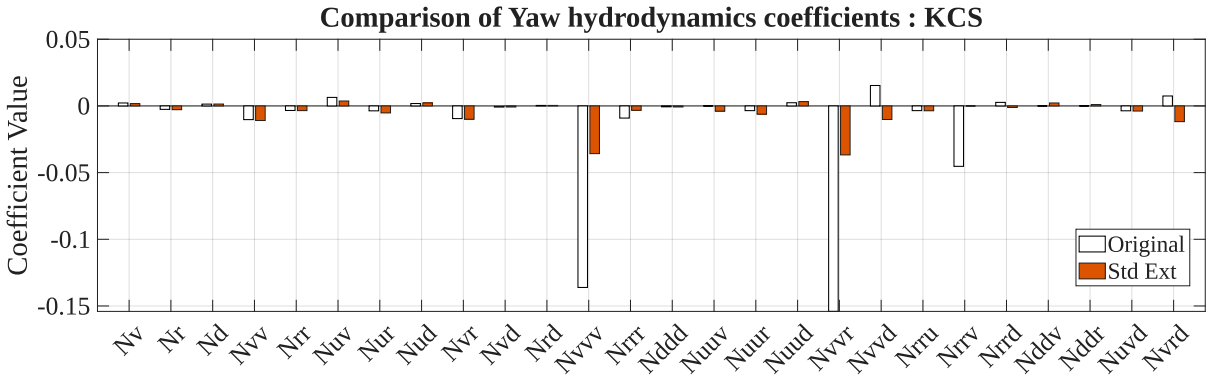
Figure 18: Trajectory reconstruction performance for the KCS (left) and Mariner (right) vessels using the *Standard Maneuvers Extended* dataset. Predicted (dashed) and reference (solid) trajectories are shown for (a,b) Zig Zag 10-35, (c,d) random step rudder signal, and (e,f) random uniform rudder signal.



(a) Optimal thresholding parameter for surge $\lambda = 0.0215$

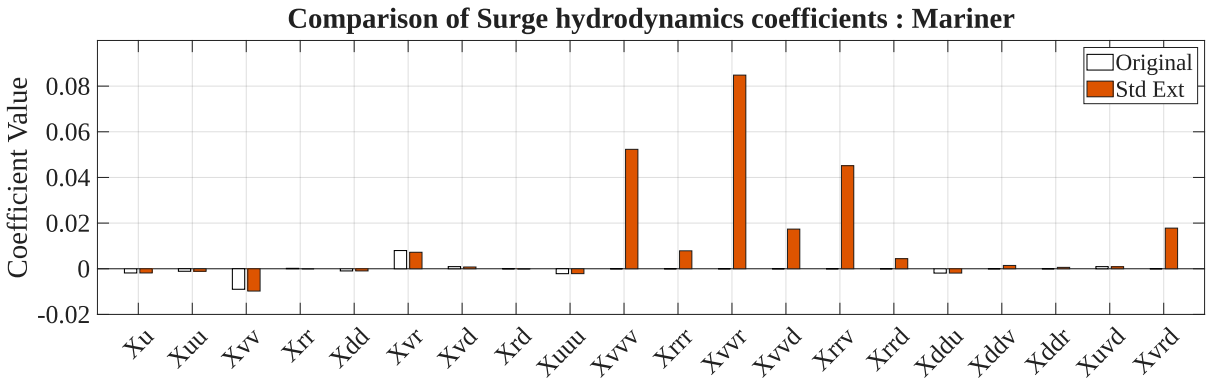


(b) Optimal thresholding parameter for sway $\lambda = 0.00166$

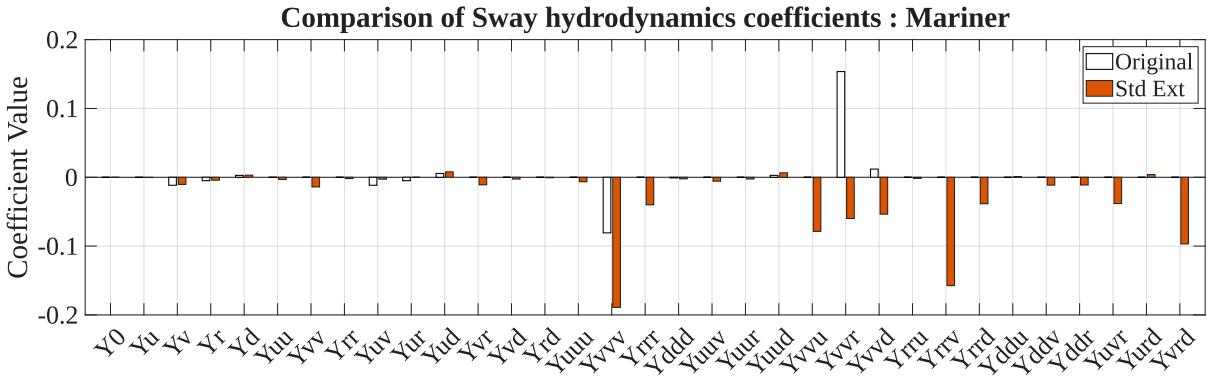


(c) Optimal thresholding parameter for yaw $\lambda = 0.0129$

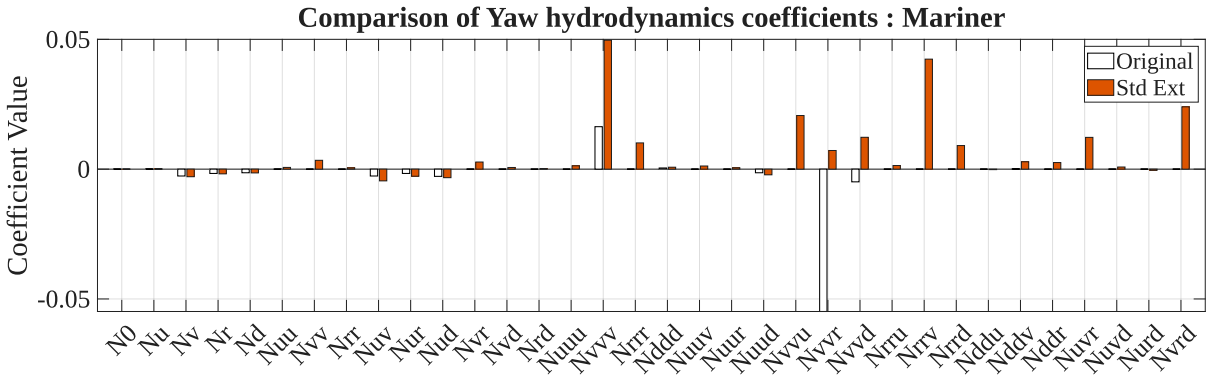
Figure 19: Comparison of identified vs. reference hydrodynamic coefficients for the KCS vessel using the *Standard Maneuvers Extended* dataset.



(a) Optimal thresholding parameter for surge $\lambda = 0.00774$



(b) Optimal thresholding parameter for sway $\lambda = 0.0129$



(c) Optimal thresholding parameter for yaw $\lambda = 0.001$

Figure 20: Comparison of identified vs. reference hydrodynamic coefficients for the Mariner vessel using the *Standard Maneuvers Extended* dataset.

Table 6: Mean squared trajectory error for models identified using Standard Maneuvers (Literature), Standard Maneuvers (Extended) and DEM averaged over 50 random rudder signals. The values represented are in terms of non dimensional lengths.

Vessel	**Standard (Lit)**	**Standard (Ext)**	**DEM**
KCS	13.135	1.280	0.700
Mariner	94.4459	7.1757	0.0629

the coefficient estimates from DEM show improvement across all three modes compared to standard-maneuver extended dataset. The predicted model is also more parsimonious as compared to the standard maneuver datasets. The richness of the excitation in this case provides enough variability to correctly estimate the necessary hydrodynamic coefficients. It reduces the correlations that make extra parameters seem useful under limited excitation.

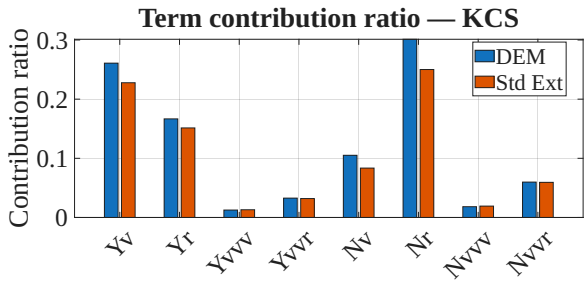
In addition to the trajectory reconstruction results shown, another important maneuver used in ship maneuvering analysis is the turning circle maneuver. Unlike the zig–zag and random rudder inputs considered earlier, the turning circle maneuver captures several standard maneuvering characteristics that are commonly used to evaluate ship maneuverability. These include the *advance*, which corresponds to the maximum distance traveled by the vessel in the direction of its initial motion during the initial phase of turning when the vessel begins to deviate from its straight path, the *tactical diameter*, which represents the maximum lateral displacement of the vessel as the sway velocity and yaw rate evolve toward their steady turning values, and the *steady turning radius*, which corresponds to the radius of the circular trajectory once the vessel settles into a steady turning motion.

Figure 23 presents the turning circle trajectories obtained from the identified model and the reference model for both the KCS and Mariner vessels. Results are shown for both starboard and port turning maneuvers to ensure that the identified model captures the turning dynamics consistently in both directions.

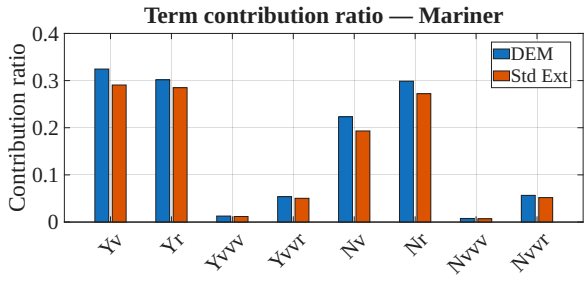
To better understand the difficulty in identifying certain nonlinear and cross–coupled hydrodynamic derivatives such as $Y_{vvr}, N_{vvr}, Y_{vvv}, N_{vvv}$, the relative contribution of individual regression terms along the trajectory is analyzed. For each term, the contribution is computed as the time-averaged ratio of the magnitude of its force contribution to the sum of magnitudes of force contribution of all terms. This metric provides an estimate of how strongly each of these terms influences the predicted forces and moments along the observed motion.

The resulting contributions for the KCS and Mariner vessels are shown in Figure 21. The analysis indicates that linear terms such as Y_v, Y_r, N_v, and N_r dominate the overall force generation along the trajectories. In contrast, higher–order nonlinear and cross–coupled terms such as Y_{vvv}, Y_{vvr}, N_{vvv}, and N_{vvr} contribute significantly less to the total forces and moments. Because these terms have relatively small contributions throughout the trajectory, their influence on the system dynamics is comparatively weak, making their coefficients more difficult to estimate reliably from the available data.

A comparison between the DEM and Standard Maneuvers (Extended) datasets shows that the dominant linear terms remain the primary contributors in both cases. While the DEM dataset exhibits slightly larger contributions for some of the linear derivatives, the nonlinear and cross–coupled terms remain relatively small contributors overall. This observation suggests that the weaker excitation of these higher–order interaction terms contributes to the difficulty in accurately identifying their corresponding coefficients.

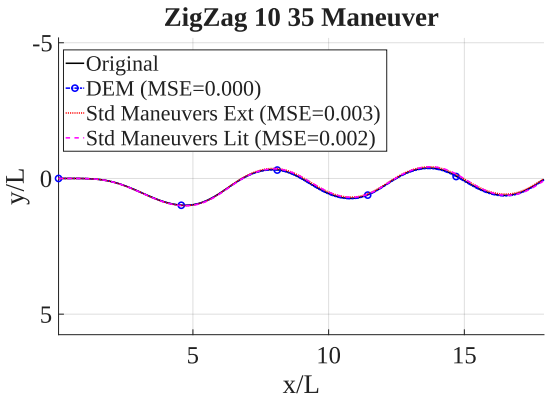


(a) KCS vessel

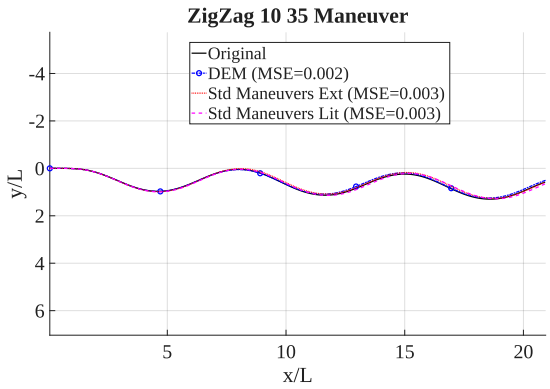


(b) Mariner vessel

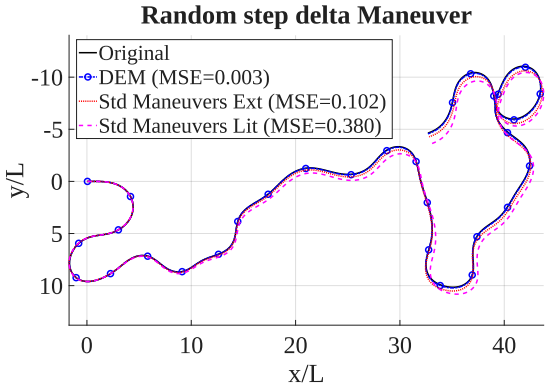
Figure 21: Relative contribution of selected hydrodynamic derivative terms to the reconstructed forces and moments along the trajectory. The contribution of each term is computed as $\text{mean}_t\left(|\theta_j\phi_j(t)| / \sum_k |\theta_k\phi_k(t)|\right)$, where θ_j denotes the original coefficient value and $\phi_j(t)$ is the corresponding regression term.



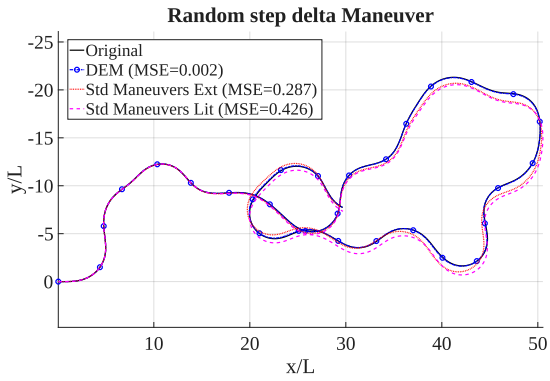
(a) Zig Zag 10-35 : KCS



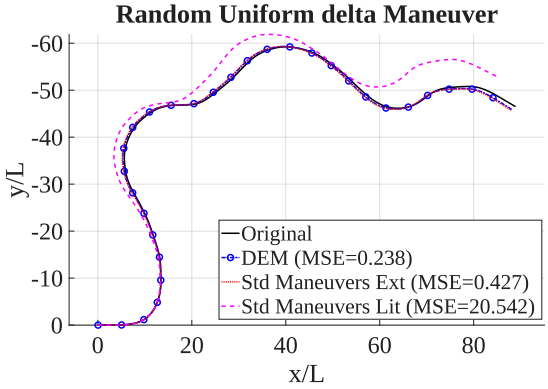
(b) Zig Zag 10-35 : Mariner



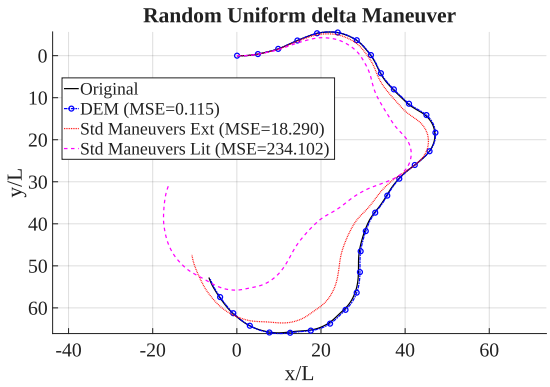
(c) Random step rudder signal : KCS



(d) Random step rudder signal : Mariner

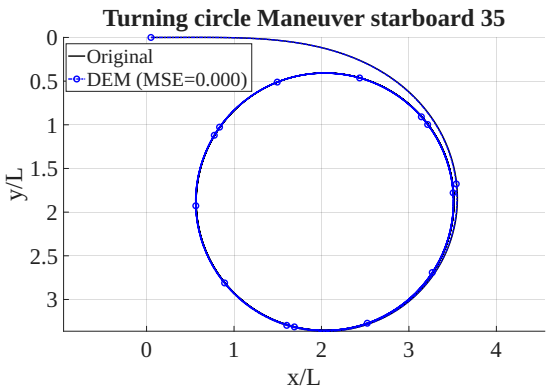


(e) Random uniform rudder signal : KCS

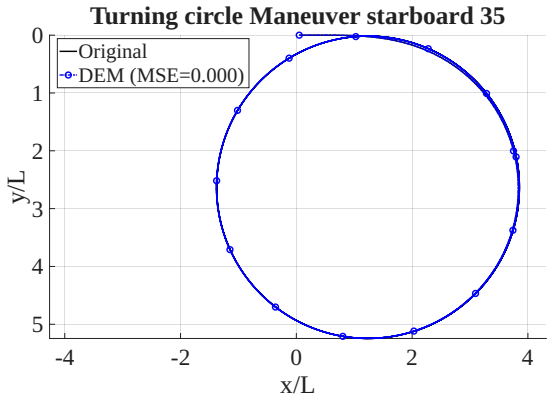


(f) Random uniform rudder signal : Mariner

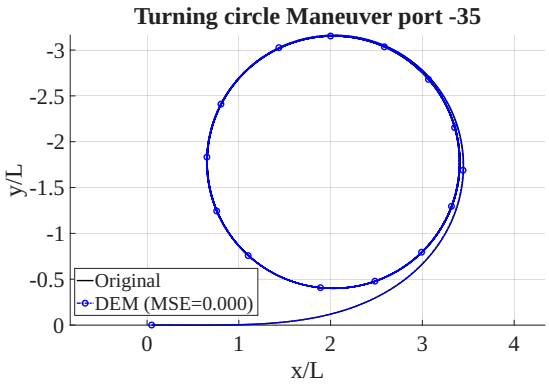
Figure 22: Trajectory reconstruction performance for the KCS (left) and Mariner (right) vessels using the *DEM* dataset. Predicted (dashed) and reference (solid) trajectories are shown for (a,b) Zig Zag 10-35, (c,d) random step rudder signal, and (e,f) random uniform rudder signal.



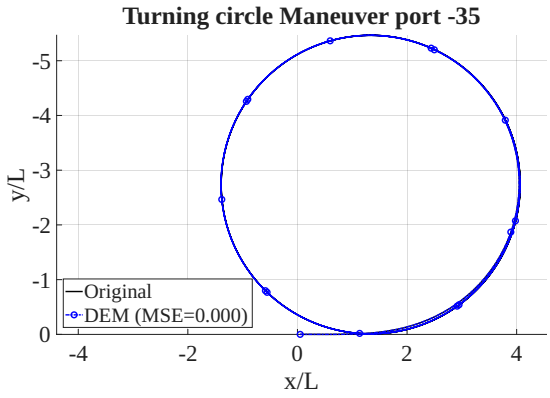
(a) Starboard turning circle $\delta = 35°$: KCS



(b) Starboard turning circle $\delta = 35°$: Mariner

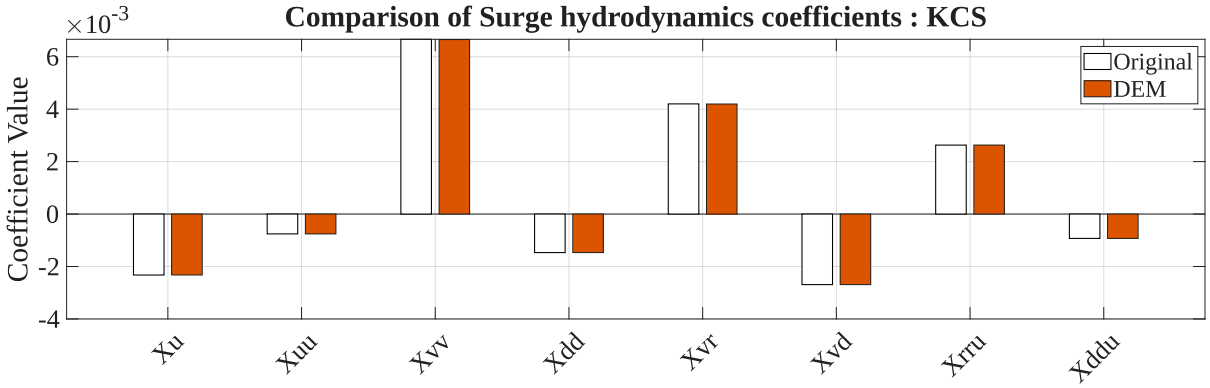


(c) Port turning circle $\delta = -35°$: KCS

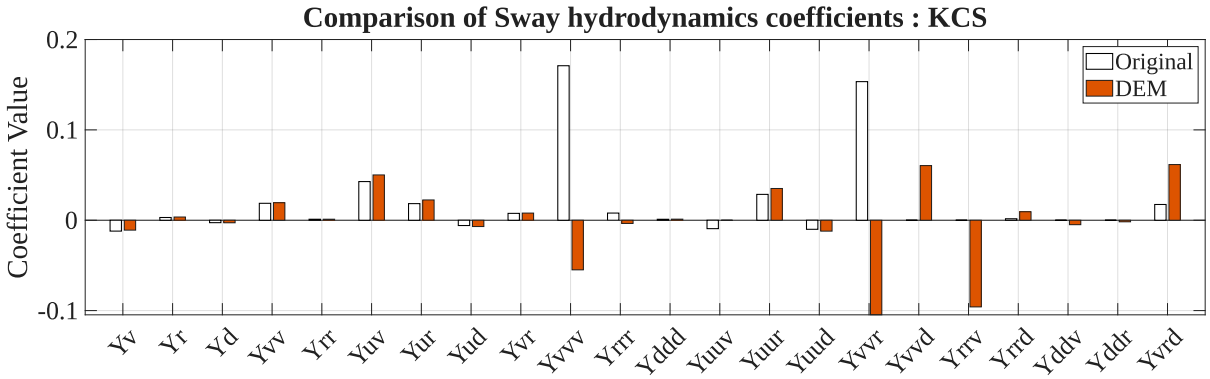


(d) Port turning circle $\delta = -35°$: Mariner

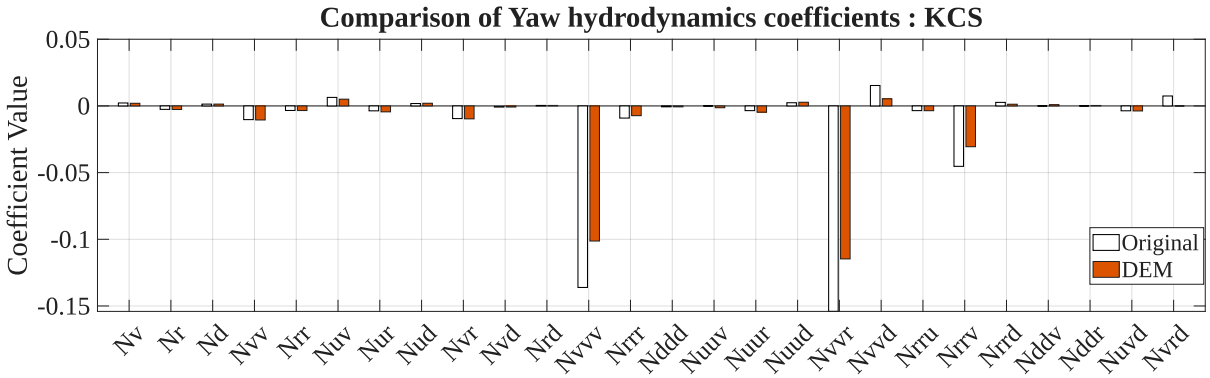
Figure 23: Turning circle trajectory comparison between the reference model (solid) and the model identified using the DEM dataset (dashed). Results are shown for both starboard and port turning maneuvers for the KCS and Mariner vessels. The turning circle maneuver captures key maneuvering characteristics including advance, tactical diameter and steady turning radius, which are widely used to evaluate ship maneuverability. The close agreement between the predicted and reference trajectories indicates that the DEM dataset provides sufficient excitation to accurately capture the nonlinear hydrodynamic effects governing turning dynamics.



(a) Optimal thresholding parameter for surge $\lambda = 0.00774$

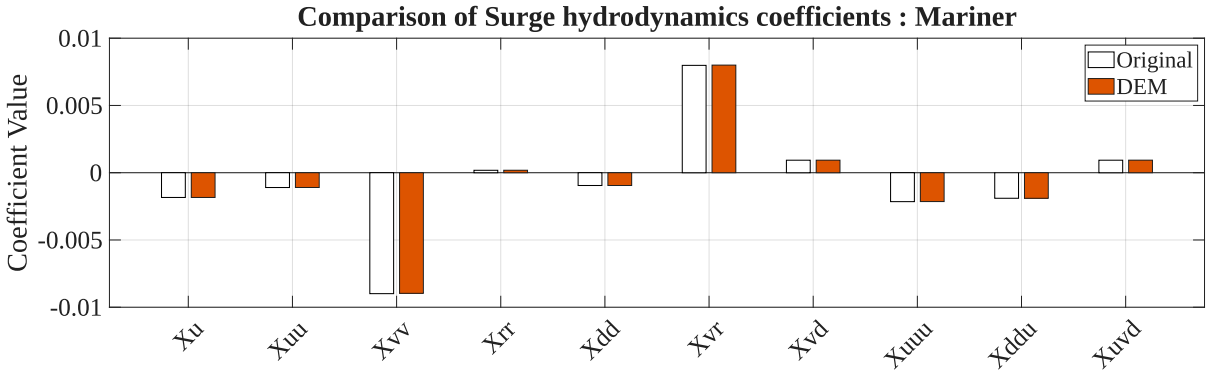


(b) Optimal thresholding parameter for sway $\lambda = 0.00774$

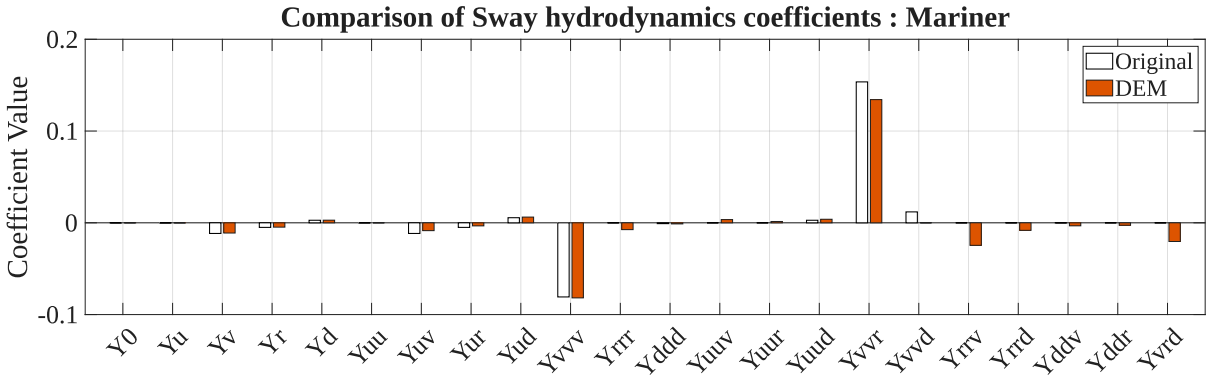


(c) Optimal thresholding parameter for yaw $\lambda = 0.0129$

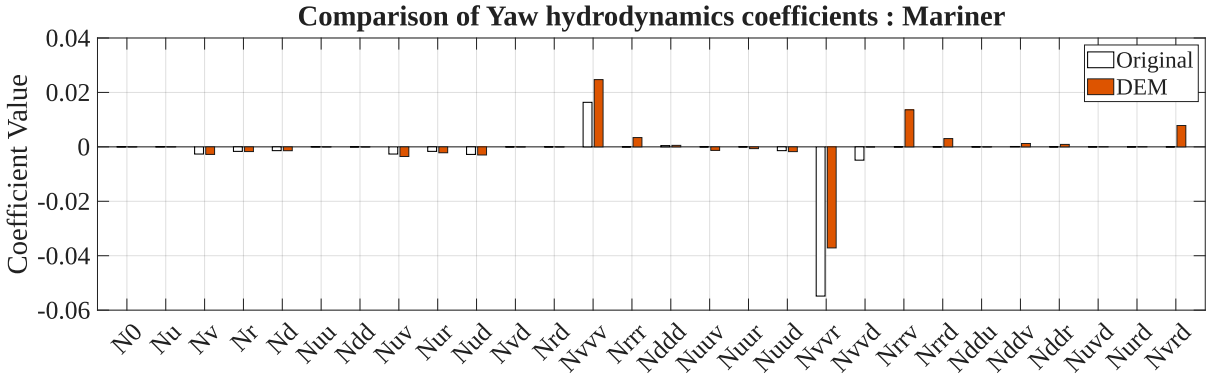
Figure 24: Comparison of identified vs. reference hydrodynamic coefficients for the KCS vessel using the *DEM* dataset.



(a) Optimal thresholding parameter for surge $\lambda = 0.00166$



(b) Optimal thresholding parameter for sway $\lambda = 0.00278$



(c) Optimal thresholding parameter for yaw $\lambda = 0.00278$

Figure 25: Comparison of identified vs. reference hydrodynamic coefficients for the Mariner vessel using the *DEM* dataset.

7 Coefficient Prediction Metrics and Observations

To quantitatively compare the quality of parameter identification across datasets, several complementary performance metrics are evaluated: Symmetric Mean Absolute Percentage Error (sMAPE), Normalized Root Mean Square Error (NRMSE), Cosine Similarity, and the classification-based Precision, Recall, and F1-score. Each metric captures a different aspect of how well the estimated coefficients match the true physical model parameters.

The description of each metric and the motivation for using it is as follows:

1. **Symmetric Mean Absolute Percentage Error (sMAPE)**

   sMAPE measures the relative deviation between the predicted and true coefficients, normalized by their local average magnitude. Considering $\hat{c}_i$ is the predicted value of the coefficient and c_i is the true value and the averaging is performed over n coefficients, the sMAPE is defined as follows

$$\text{sMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{c}_i - c_i|}{(|\hat{c}_i| + |c_i| + \varepsilon)/2}. \tag{36}$$

   This metric provides a scale-independent proportional error, which is particularly useful when coefficient magnitudes differ significantly across terms.

2. **Normalized Root Mean Square Error (NRMSE)**

   The Normalized Root Mean Square Error quantifies the overall magnitude of the prediction error relative to the energy of the true coefficient vector. The NRMSE used in this work is defined as

$$\text{NRMSE} = \frac{\|\hat{\mathbf{c}} - \mathbf{c}\|_2}{\|\mathbf{c}\|_2 + \varepsilon}, \tag{37}$$

   where $\hat{\mathbf{c}}$ denotes the vector of identified coefficients, $\mathbf{c}$ denotes the true coefficients, and ε is a small numerical stability constant included to avoid division by zero.

   The numerator and denominator are given explicitly by

$$\|\hat{\mathbf{c}} - \mathbf{c}\|_2 = \sqrt{\sum_{i=1}^{n}(\hat{c}_i - c_i)^2}, \tag{38}$$

$$\|\mathbf{c}\|_2 = \sqrt{\sum_{i=1}^{n} c_i^2}. \tag{39}$$

   This formulation normalizes the root-mean-square deviation between the predicted and true parameters by the L_2 norm of the true coefficient vector, resulting in a scale-independent measure of identification accuracy. While NRMSE captures absolute numerical deviation, it may be misleading when the coefficient magnitudes vary across several orders of magnitude.

3. **Cosine Similarity**

   Cosine Similarity evaluates the directional alignment between the predicted and true coefficient vectors:

$$\text{CosSim} = \frac{\hat{\mathbf{c}} \cdot \mathbf{c}}{\|\hat{\mathbf{c}}\|_2 \, \|\mathbf{c}\|_2 + \varepsilon}. \tag{40}$$

   A value close to one indicates that the relative distribution and sign structure of coefficients is preserved, even when magnitudes differ. This metric therefore quantifies how well the structural pattern of the physical model is retained.

4. **Precision, Recall, and F1-score (Sparsity Consistency Metrics)**

   These metrics evaluate how accurately the sparsity pattern of active vs. inactive terms is recovered by the identified model. They are not used here for classification, but rather for comparing the identified model structure. A coefficient is considered active if $c_i \neq 0$, and inactive if $c_i = 0$. Using the following definitions,

$$\text{TP} = |\{\, i : c_i \neq 0 \text{ and } \hat{c}_i \neq 0 \,\}|,$$

$$\text{FP} = |\{\, i : c_i = 0 \text{ but } \hat{c}_i \neq 0 \,\}|,$$

$$\text{FN} = |\{\, i : c_i \neq 0 \text{ but } \hat{c}_i = 0 \,\}|.$$

the metrics are defined as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{41}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{42}$$

$$\text{F1-score} = 2\,\frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{43}$$

These structural metrics are especially relevant because Sequentially Thresholded Least Squares (STLS) enforces sparsity. As a result, a model may exhibit a small NRMSE yet misidentify which terms should remain active or inactive, or conversely recover the structure correctly while differing in coefficient magnitudes.

The value of ε taken in denominators during metric computation is 10^{-12}. Relying on a single error-based metric can give an incomplete or misleading picture of identification quality. A holistic comparison using both magnitude-based and structure-based metrics ensures that models are evaluated both for quantitative accuracy and physical interpretability.

Observations from Tables 7 and 8 suggest that for both the KCS and Mariner vessels, the model identified using the DEM dataset consistently provides lower sMAPE and NRMSE values compared to those trained with Standard Maneuvers (Extended and Literature) across all modes—surge, sway, and yaw. The DEM-based identification also yields Cosine Similarity values near unity in most cases except in the sway dynamics of KCS vessel, indicating good alignment in coefficient direction and magnitude. In contrast, both Standard Maneuver datasets exhibit either large errors and relatively small or negative Cosine Similarity values in several cases signifying poor structural alignment with the true model.

The precision–recall metrics further reinforce this observation. DEM achieves both higher Precision and F1 score consistently, reflecting that the sparsity pattern of the identified coefficients has good consistency with the true system dynamics. It can be noted though that the Standard Maneuver datasets sometimes maintain higher Recall than DEM. This implies that most true coefficients are retained in the identification but this results mainly due to less sparsity of the models identified from Standard Maneuver datasets. The higher Recall with lower Precision for these Standard Maneuever datasets shows that they tend to include spurious coefficients. This can be attributed to insufficient excitation diversity in the training data which leads to higher correlation among the columns of the regressor matrix and poor conditioning.

Overall, these results indicate that richer and dynamically varied datasets like DEM enable the identification of models that are numerically accurate, parsimonious and physically consistent. The combination of all six metrics provides a complete understanding of both the structural fidelity and numerical reliability of the identified models, which would not be evident from any single metric alone.

7.1 BIC-based Model Selection and λ Analysis

To identify an optimal level of model sparsity, sequentially thresholded least-squares regression was performed across a logarithmic range of the thresholding parameter λ whose values ranged between 10^{-3} and 10^{-1}. This range was selected based on preliminary tests that showed:

1. For $\lambda < 10^{-3}$, the BIC values began to decrease only marginally, suggesting overfitting tendencies where redundant nonlinear terms are retained.

2. For $\lambda > 10^{-1}$, excessive penalization caused rapid growth in the BIC score due to underfitting, as important nonlinear contributions were pruned out.

Thus, the interval $[10^{-3}, 10^{-1}]$ provides a balance between bias and variance in the identified hydrodynamic models. The Bayesian Information Criterion (BIC) is used to quantify the trade-off between model fit and complexity as described in subsection 3.4. For each λ, the model yielding the minimum BIC is considered the most parsimonious representation of the underlying dynamics.

The results of BIC based model selection are presented in Table 9 and Table 10, showing the optimal λ, number of active terms K, residual variance σ^2, and corresponding minimum BIC for each

Table 7: Comparison of coefficient prediction metrics for the KCS vessel across three datasets: DEM, Standard Maneuvers Extended, and Standard Maneuvers Literature. Each mode shows metrics for all datasets in separate rows for readability.

Mode	**Dataset**	**Mean sMAPE**	**NRMSE**	**Cosine Sim.**	**Precision**	**Recall**	**F1-score**
Surge	DEM	0.00005	1.31×10^{-3}	1.00000	1.00000	1.00000	1.00000
	Std Ext	0.00087	2.48×10^{-2}	0.99987	1.00000	1.00000	1.00000
	Std Lit	0.58938	7.91×10^{0}	-0.06842	0.25000	1.00000	0.40000
Sway	DEM	0.20823	1.53×10^{0}	-0.44499	0.80952	0.94444	0.87179
	Std Ext	0.28755	2.14×10^{0}	-0.65837	0.72000	1.00000	0.83721
	Std Lit	0.55407	1.29×10^{1}	-0.85111	0.52941	1.00000	0.69231
Yaw	DEM	0.12691	2.64×10^{-1}	0.99781	0.88000	0.95652	0.91667
	Std Ext	0.20991	7.74×10^{-1}	0.87713	0.88000	0.95652	0.91667
	Std Lit	0.27320	2.70×10^{0}	-0.96665	0.88000	0.95652	0.91667

Table 8: Comparison of coefficient prediction metrics for the Mariner vessel across three datasets: DEM, Standard Maneuvers Extended, and Standard Maneuvers Literature. Each mode shows metrics for all datasets in separate rows for readability.

Mode	**Dataset**	**Mean sMAPE**	**NRMSE**	**Cosine Sim.**	**Precision**	**Recall**	**F1-score**
Surge	DEM	0.00041	2.57×10^{-3}	1.00000	1.00000	1.00000	1.00000
	Std Ext	0.23365	8.90×10^{0}	0.11220	0.47368	0.90000	0.62069
	Std Lit	0.56307	4.92×10^{0}	0.19567	0.28571	1.00000	0.44444
Sway	DEM	0.22612	2.36×10^{-1}	0.97301	0.65000	0.86667	0.74286
	Std Ext	0.54113	1.87×10^{0}	0.10946	0.43750	0.93333	0.59574
	Std Lit	0.50944	1.46×10^{0}	0.33925	0.48387	1.00000	0.65217
Yaw	DEM	0.33252	4.52×10^{-1}	0.89435	0.51852	0.93333	0.66667
	Std Ext	0.51861	1.59×10^{0}	0.09041	0.44118	1.00000	0.61224
	Std Lit	0.51710	1.43×10^{0}	0.19429	0.44118	1.00000	0.61224

motion mode (surge, sway, and yaw) and maneuver type. The results show that for both vessels, the DEM data consistently yields lower residual variances and smaller model orders (K) for each motion mode. This implies that the DEM signals provide richer excitation, allowing the identification of compact yet accurate nonlinear dynamic models. In contrast, standard maneuvers extended and literature variants show larger number of active terms (K) and greater residual variances, reflecting comparatively greater redundancy and poor accuracy in the identification. The overall trend across both vessels indicates that data richness directly affects the model selection sensitivity and sparsity characteristics governed by λ. The BIC trends further reinforce the observation that, richer datasets drive the model toward an optimal region of parsimony where unnecessary terms are naturally pruned without sacrificing fidelity.

8 Conclusion and Future work

This study systematically demonstrates that the composition of the excitation dataset plays a critical role in the accuracy of ship maneuvering model identification. Across both vessels analyzed, richer

Table 9: Optimal λ, residual variance σ^2, active terms (K), and minimum BIC for the KCS vessel across different maneuver types.

Maneuver Type	**Mode**	$\boldsymbol{\lambda_{\text{opt}}}$	$\boldsymbol{\sigma^2}$ **(**$\boldsymbol{\times 10^{-6}}$**)**	$\boldsymbol{K}$	**BIC (Best)**
DEM	Surge	7.74×10^{-3}	0.950	8	-2.6019×10^5
DEM	Sway	7.74×10^{-3}	1.591	21	-2.4791×10^5
DEM	Yaw	1.29×10^{-2}	1.408	25	-2.5075×10^5
Standard (Ext.)	Surge	2.15×10^{-2}	11.897	8	-2.2935×10^5
Standard (Ext.)	Sway	1.67×10^{-3}	3.945	25	-2.5908×10^5
Standard (Ext.)	Yaw	1.29×10^{-2}	2.364	25	-2.7291×10^5
Standard (Lit.)	Surge	1.67×10^{-3}	17.645	32	-7.2673×10^4
Standard (Lit.)	Sway	1.00×10^{-3}	5.784	34	-8.2693×10^4
Standard (Lit.)	Yaw	1.00×10^{-3}	4.950	34	-8.5570×10^4

Table 10: Optimal λ, residual variance σ^2, active terms (K), and minimum BIC for the Mariner vessel across different maneuver types.

Maneuver Type	**Mode**	$\boldsymbol{\lambda_{\text{opt}}}$	$\boldsymbol{\sigma^2}$ **(**$\boldsymbol{\times 10^{-6}}$**)**	$\boldsymbol{K}$	**BIC (Best)**
DEM	Surge	1.6681×10^{-3}	0.444	10	-2.781184×10^5
DEM	Sway	2.7826×10^{-3}	8.291	20	-2.089524×10^5
DEM	Yaw	2.7826×10^{-3}	17.349	27	-1.914568×10^5
Standard (Ext.)	Surge	7.7426×10^{-3}	2.990	19	-2.666236×10^5
Standard (Ext.)	Sway	1.2915×10^{-2}	23.138	32	-2.112501×10^5
Standard (Ext.)	Yaw	1.0000×10^{-3}	43.902	34	-1.939362×10^5
Standard (Lit.)	Surge	1.0000×10^{-3}	5.396	35	-8.330855×10^4
Standard (Lit.)	Sway	1.0000×10^{-3}	34.972	35	-6.648921×10^4
Standard (Lit.)	Yaw	1.0000×10^{-3}	43.902	34	-1.939362×10^5

and more varied rudder excitation signals yield models that are simultaneously more accurate and parsimonious. The proposed Dynamic Excitation Maneuver (DEM), by integrating diverse dynamic transitions within a single trajectory, enables reliable identification of nonlinear and cross-coupled hydrodynamic coefficients. Quantitatively, this results in up to two orders of magnitude improvement in trajectory reconstruction error compared to models identified from standard maneuver datasets.

These results establish that, in system identification for ships with constant propeller speed, data richness outweighs data volume. Datasets that contain a diverse set of transitions across the vessel's admissible state space enable better recovery of the underlying sparse dynamical structure while improving predictive accuracy. Quantitative metrics using entropy and condition number are also proposed to evaluate dataset richness, and a systematic analysis is presented to characterize their influence on identification quality and model parsimony. Overall, this study reframes the identification problem from a purely algorithmic challenge to one of data design.

From the perspective of autonomous surface vessels, these findings highlight the importance of dataset design in ensuring reliable model-based decision making (Leo et al., 2026). Maneuvering models identified from insufficiently rich datasets may fail to capture critical nonlinear behaviors (Tian et al., 2025), potentially degrading the performance of guidance, navigation, and control algorithms in safety-critical scenarios such as collision avoidance or close-quarter operations. The proposed DEM framework offers a systematic approach to gen-

erating informative datasets that improve model fidelity, thereby contributing toward more robust and reliable autonomy in marine systems.

Despite the promising results, it is important to clearly delineate the limitations of the proposed framework to properly interpret its applicability. The identification results show significant improvement in model parsimony with increasing data richness; however, parameter drift in sway dynamics for the KCS vessel indicates that certain aspects of the model remain sensitive to data and system characteristics. It is observed that the dominant first and second-order hydrodynamic coefficients are recovered with correct signs and physically consistent magnitudes under DEM excitation but minor discrepancies persist in certain higher-order nonlinear terms. Unlike lower-order damping terms, these coefficients do not generally possess strict sign constraints and can vary depending on vessel characteristics. The present study focuses on isolating the effect of data richness on identification performance. The incorporation of additional physical constraints within the regression framework, while potentially beneficial for robustness, is not considered here to avoid confounding the influence of excitation design. The entropy analysis for quantifying the richness of each dataset is performed over a simulation-based approximation of the explored operating region and does not attempt to rigorously characterize the complete physically admissible state space of the nonlinear vessel dynamics.

The identification framework assumes a predefined polynomial library consistent with the Abkowitz maneuvering model structure. As a result, the methodology estimates coefficients within this assumed functional form rather than discovering new governing dynamics. Furthermore, the performance evaluation is conducted under controlled simulation conditions, focusing primarily on trajectory reconstruction and coefficient recovery. Environmental disturbances such as wind, waves, and currents, as well as sensor uncertainties, are not explicitly modeled. Although the rudder actuator is represented using a simplified first-order steering gear model with rudder-rate saturation, higher-fidelity actuator and servo dynamics such as backlash and nonlinearity are beyond the scope of the present study.

The present study was conducted under noise-free conditions in order to isolate and systematically examine the influence of data richness on the identification of ship maneuvering models. In this context, the analysis focused specifically on how the variability and coverage of maneuvering trajectories affect the identifiability, sparsity, and generalization capability of the identified hydrodynamic models. Introducing measurement noise would couple the effects of excitation richness with the robustness of the estimation framework to noisy observations. Consequently, any degradation in coefficient recovery or trajectory prediction could arise either from insufficient state-space exploration or from noise amplification during derivative estimation and sparse regression, making it difficult to isolate the individual contribution of data richness.

While the present work establishes the role of excitation design in a controlled setting, the interaction between data richness and measurement noise remains an important direction for future research. In practical scenarios, identification performance is influenced not only by the richness of the collected trajectories but also by factors such as sensor noise, filtering strategies, derivative estimation methods, regularization approaches, and uncertainty propagation. Future work will therefore investigate the robustness of the proposed framework under realistic noise conditions and explore the integration of noise-robust identification techniques with richness-aware maneuver design. Although physics-constrained sparse identification is not required to establish the present contribution, it represents a promising extension for robust identification using noisy experimental or full-scale measurements.

Building on the insights from this study, future work will focus on integrating data-rich excitation design with more robust identification strategies. In particular, incorporating physics-informed constraints and noise-resilient regression techniques may improve model reliability under realistic operating conditions. Further investigation is also needed to address the interaction between data richness, model structure, and noise sensitivity in practical identification scenarios. Extending the framework to accommodate alternative model representations beyond predefined polynomial libraries, as well as validating the approach using experimental data, represents an important next step.

The next stage of this research will therefore ex-

plore adaptive and hybrid identification strategies that combine excitation design with improved estimation robustness, enabling reliable maneuvering model identification for real-world autonomous marine systems.

Acknowledgment

The authors would like to acknowledge the support from the Center of Excellence for Maritime Experiments to Maritime Experience (ME2ME), Department of Ocean Engineering, IIT Madras. This work was supported by the ANRF Advanced Research Grant (ARG) scheme of Anusandhan National Research Foundation (ANRF), India, Grant No. ANRF/ARG/2025/000738/ENS. The author, Vallabh Deogaonkar (PMRF Research Scholar), gratefully acknowledges the support of the Prime Minister's Research Fellowship, India, for supporting the completion of this study.

Data Availability Statement

The trajectory datasets used in this study are available in the Zenodo repository at https://doi.org/10.5281/zenodo.19030617 The simulation model and hydrodynamic coefficients used to generate the datasets are provided in the manuscript.

References

Abkowitz, M. A. (1964). Lectures on ship hydrodynamics–steering and manoeuvrability. Technical report.

Abkowitz, M. A. (1980). Measurement of hydrodynamic characteristics from ship maneuvering trials by system identification. Technical report.

Araki, M., Sadat-Hosseini, H., Sanada, Y., Tanimoto, K., Umeda, N., and Stern, F. (2012). Estimating maneuvering coefficients using system identification methods with experimental, system-based, and cfd free-running trial data. *Ocean Engineering*, 51:63–84.

Bendat, J., Palo, P., and Coppolino, R. (1992). A general identification technique for nonlinear differential equations of motion. *Probabilistic Engineering Mechanics*, 7(1):43–61.

Birk, L. (2019). *Fundamentals of ship hydrodynamics: Fluid mechanics, ship resistance and propulsion*. John Wiley & Sons.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Brunton, S. L., Proctor, J. L., and Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, 113(15):3932–3937.

Cheng, J. and Falzarano, J. M. (2003). System identification of nonlinear coupled ship/offshore platform dynamics in beam seas. In *International Conference on Offshore Mechanics and Arctic Engineering*, volume 36819, pages 531–540.

Costa, A. C., Xu, H., and Guedes Soares, C. (2021). Robust parameter estimation of an empirical manoeuvring model using free-running model tests. *Journal of Marine Science and Engineering*, 9(11):1302.

Deogaonkar, V. V., Jadhav, A. K., Ramachandran, K., and Somayajula, A. S. (2023). Data driven identification of ship maneuvering coefficients. In *International Conference on Offshore Mechanics and Arctic Engineering*, volume 86878, page V005T06A050. American Society of Mechanical Engineers.

Deraj, R., Kumar, R. S., Alam, M. S., and Somayajula, A. (2023). Deep reinforcement learning based controller for ship navigation. *Ocean Engineering*, 273:113937.

Deshpande, S. and Rivera, D. E. (2013). Optimal input signal design for data-centric estimation methods. In *2013 American control conference*, pages 3924–3929. IEEE.

Diez, M., Serani, A., Campana, E. F., and Stern, F. (2021). Data-driven modelling of ship maneuvers in waves via dynamic mode decomposition. *arXiv preprint arXiv:2105.13062*.

Fossen, T. I. (2011). *Handbook of marine craft hydrodynamics and motion control*. John Wiley & Sons.

Gil, M. (2021). A concept of critical safety area applicable for an obstacle-avoidance process for manned and autonomous ships. *Reliability Engineering & System Safety*, 214:107806.

Gil, M., Krata, P., Kozioł, P., and Hinz, T. (2024). A multiparameter simulation-driven analysis of ship turning trajectory concerning a required number of irregular wave realizations. *Ocean Engineering*, 299:117293.

Heinz, T. O. and Nelles, O. (2018). Excitation signal design for nonlinear dynamic systems with multiple inputs–a data distribution approach. *at-Automatisierungstechnik*, 66(9):714–724.

Huang, L., Pena, B., and Thomas, G. (2023). Towards a full-scale cfd guideline for simulating a ship advancing in open water. *Ship Technology Research*, pages 1–17.

Jadhav, A. K., Pandi, A. R., and Somayajula, A. (2023). Collision avoidance for autonomous surface vessels using novel artificial potential fields. *Ocean Engineering*, 288:116011.

Jegorova, M., Smith, J., Mistry, M., and Hospedales, T. (2020). Adversarial generation of informative trajectories for dynamics system identification. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7109–7115. IEEE.

Källström, C. G. and Åström, K. J. (1981). Experiences of system identification applied to ship steering. *Automatica*, 17(1):187–198.

Kanazawa, M., Wang, T., Ichinose, Y., Skulstad, R., Li, G., and Zhang, H. (2023). Bridging similar ships' dynamics for safeguarding the system identification of maneuvering models. *Ocean Engineering*, 280:114874.

Kiss, M., Tóth, R., and Schoukens, M. (2024). Space-filling input design for nonlinear state-space identification. *IFAC-PapersOnLine*, 58(15):562–567.

Krata, P., Gil, M., Hinz, T., and Kozioł, P. (2024). A multiparameter simulation-driven analysis of ship response when turning concerning a required number of irregular wave realizations. *Ocean Engineering*, 302:117701.

Leo, D., Somayajula, A., and Murali, K. (2026). Ship maneuvering in shallow water with varying depth and draft: a nonlinear parameter-varying mmg-nmpc framework. *Ocean Engineering*, 357:125411.

Li, S., Wang, T., Li, G., and Zhang, H. (2023). Ship maneuvering model optimization for improved identification with less excitation. *Ocean Engineering*, 280:114540.

Mitsuyuki, T., Kuribayashi, K., Fernandez, R. F. S., Shimozawa, H., Kakuta, R., Niki, R., and Matsushita, R. (2024). Mmg 3dof model identification with uncertainty of observation and hydrodynamic maneuvering coefficients using mcmc method. *Journal of Marine Science and Technology*, 29(3):668–682.

Mohan, A. and Somayajula, A. S. (2023). Analyzing robustness and accuracy of different controllers for underactuated ships. In *2023 4th International Conference for Emerging Technology (INCET)*, pages 1–6. IEEE.

Paninski, L. (2003). Estimation of entropy and mutual information. *Neural computation*, 15(6):1191–1253.

Perez, T., Smogeli, O., Fossen, T., and Sorensen, A. J. (2006). An overview of the marine systems simulator (mss): A simulink toolbox for marine control systems. *Modeling, identification and Control*, 27(4):259–275.

Rajesh, G. and Bhattacharyya, S. K. (2008). System identification for nonlinear maneuvering of large tankers using artificial neural network. *Applied Ocean Research*, 30(4):256–263.

Selvam, R. P., Bhattacharyya, S., and Haddara, M. (2005). A frequency domain system identification method for linear ship maneuvering. *International shipbuilding progress*, 52(1):5–27.

Simpkins, A. (2012). System identification: Theory for the user, (ljung, l.; 1999)[on the shelf]. *IEEE Robotics & Automation Magazine*, 19(2):95–96.

Somayajula, A. and Falzarano, J. (2016). Estimation of roll motion parameters using r-miso system identification technique. In *ISOPE International Ocean and Polar Engineering Conference*, pages ISOPE–I. ISOPE.

Somayajula, A. and Falzarano, J. (2017). Application of advanced system identification technique to extract roll damping from model tests in order to accurately predict roll motions. *Applied Ocean Research*, 67:125–135.

Sutulo, S. and Soares, C. G. (2014). An algorithm for offline identification of ship manoeuvring mathematical models from free-running tests. *Ocean engineering*, 79:10–25.

Tian, K., Wen, Y.-Q., Zhu, M., Lei, T., Wang, S.-Y., and Huang, Y.-M. (2025). Hybrid modeling and online prediction of 4-dof ship maneuvering motion in disturbed waters. *Ocean Engineering*, 341:122566.

Vijay, A. and Somayajula, A. (2022). Identification of hydrodynamic coefficients using support vector regression. In *OCEANS 2022-Chennai*, pages 1–7. IEEE.

Vijayakumar, A., Atman, M., and Somayajula, A. (2024). Learning autonomous docking operation of fully actuated autonomous surface vessel from expert data. In *OCEANS 2024-Halifax*, pages 1–5. IEEE.

Wakita, K., Maki, A., Umeda, N., Miyauchi, Y., Shimoji, T., Rachman, D. M., and Akimoto, Y. (2022). On neural network identification for low-speed ship maneuvering model. *Journal of Marine Science and Technology*, 27(1):772–785.

Wakita, K., Miyauchi, Y., Akimoto, Y., and Maki, A. (2025). Data augmentation methods of dynamic model identification for harbor maneuvers using feedforward neural network. *Journal of Marine Science and Technology*, 30(1):18–33.

Wang, H., Somayajula, A., and Falzarano, J. (2020). Application of system identification technique in efficient model test correlations for a floating power system. *Applied Ocean Research*, 98:102126.

Wang, X.-G., Zou, Z.-J., Hou, X.-R., and Feng, X. (2015). System identification modelling of ship manoeuvring motion based on support vector regression. *Journal of Hydrodynamics, Ser. B*, 27(4):502–512.

Wang, Z., Zou, Z., and Soares, C. G. (2019). Identification of ship manoeuvring motion based on nu-support vector machine. *Ocean Engineering*, 183:270–281.

Willems, J. C., Rapisarda, P., Markovsky, I., and De Moor, B. L. (2005). A note on persistency of excitation. *Systems & Control Letters*, 54(4):325–329.

Woo, J., Park, J., Yu, C., and Kim, N. (2018). Dynamic model identification of unmanned surface vehicles using deep learning network. *Applied Ocean Research*, 78:123–133.

Xu, H. and Guedes Soares, C. (2025). Review of system identification for manoeuvring modelling of marine surface ships. *Journal of Marine Science and Application*, 24(3):459–478.

Xue, Y., Liu, Y., Xue, G., and Chen, G. (2021). Identification and prediction of ship maneuvering motion based on a gaussian process with uncertainty propagation. *Journal of Marine Science and Engineering*, 9(8):804.

Yang, R., Xu, J., Wang, X., and Zhou, Q. (2019). Parallel trajectory planning for shipborne autonomous collision avoidance system. *Applied Ocean Research*, 91:101875.

Yu, D. and Roh, M.-I. (2024). Method for anti-collision path planning using velocity obstacle and a* algorithms for maritime autonomous surface ship. *International Journal of Naval Architecture and Ocean Engineering*, 16:100586.

Zheng, J., Yan, D., Yan, M., Li, Y., and Zhao, Y. (2022). An unscented kalman filter online identification approach for a nonlinear ship motion model using a self-navigation test. *machines*, 10(5):312.