

## **Research Proposal**

# **Identifying Sources of Operational Efficiency in Maritime Transport**

**Student:** August Bjerg-Heise (114300315)

**Date of Submission:** January 1st, 2025

**Number of Pages:** 10

# 1 Introduction

The operational efficiency of a vessel can largely be defined as its ability to achieve the highest transport work for the lowest fuel consumption. (source)

With maritime transport constituting 3% of global CO<sub>2</sub> emissions, this is a pressing issue that goes beyond the shipping industry

## 2 Literature Review

Ever since the first ships were built, operators have sought to optimize energy use, first in manpower and wind, and later in modern fuels. The earliest academic contributions are commonly traced to 19th-century work on the physics of ship resistance and propulsion (Russell, 1839), alongside the pioneering of systematic towing-test approaches that enabled empirical inference about performance (Froude et al., 1955). This early literature cemented a central operational hypothesis: efficiency is governed primarily by the relationship between speed and required propulsive power, a principle embedded today in the “speed–power curve,” which maps the power needed to maintain each steady speed under specified conditions. Through almost two centuries of research, it has been found that the relationship between speed and propulsion power is approximately cubic, i.e., power need rises steeply with speed (see example in Figure 1).

As measurement and computation matured, subsequent research increasingly combined physics-grounded models with real operational data, enabling more mathematically complex and empirically testable approaches to operational efficiency analysis (Kim et al., 2025). More contemporary work can be grouped into three broad methodological categories: i) Computational Fluid Dynamics, ii) Empirical Naval Architecture and iii) Data Science.

Computational fluid dynamics (CFD) is a physics-based approach that numerically solves the flow equations around a ship hull to predict resistance and related propulsion requirements as a function of speed and operating conditions (Russell, 1839). Conceptually, this computational programme extends early resistance-focused science, including Russell’s experimental

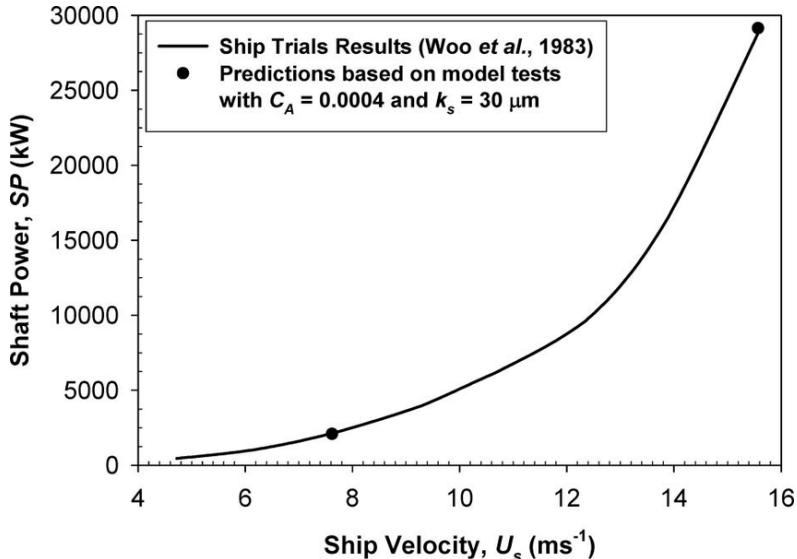


Figure 1: Example of a speed-power curve derived from empirical tests (Schultz, 2007)

work that treated resistance as a mechanistic phenomenon. Contemporary studies broaden the optimization target from calm-water performance to operational conditions, for example panel-method modelling of wave added resistance and related wave-load computations that refine how power penalties are represented under realistic sea states (Dong et al., 2023; Wang et al., 2022). Relative to purely empirical curves and data-driven models, CFD’s key strengths are physical interpretability and theoretical rigor, while its main limitations remain sensitivity to modelling assumptions and the need for careful verification and validation, alongside high computational cost.

Empirical naval architecture (ENA) models estimate the speed–power relationship from physical measurements, most commonly towing-tank tests where a scale model is towed at controlled speeds and resistance is measured for inference. In contrast to CFD, which derives the same relationship by numerically solving flow equations, towing tests provide benchmark data that is often used to validate and calibrate computational predictions. This experimental tradition is commonly linked to William Froude and Robert Froude’s (1955) development of systematic model testing and scaling concepts for predicting full-scale ship performance. More recent applications include towing-tank studies of operational interventions such as hull coatings (Schultz, 2007), while the core trade-off remains that experiments can be highly accurate under controlled conditions but are expensive, time-consuming, and not always fully representative of real sea states and in-service hull condition.

Data-science approaches model operational efficiency directly from observed data, typically without imposing any physical constraints. However, physics-informed feature choice and engineering remain necessary to avoid spurious inference. Recent work leverages increasingly dense IoT (sensor) data streams to estimate performance baselines and degradation signals, for example by building data-driven digital twins that predict “expected” speed and then interpret systematic deviations as fouling-induced speed loss (Coraddu et al., 2019). Related studies fuse operational logs with metocean data and apply machine-learning regressors to quantify hull condition, enabling counterfactual evaluation of cleaning timing under controlled operating scenarios (Duan et al., 2024). The main strength of these methods is flexibility in capturing ship-specific, nonlinear relationships at scale as data availability and ML toolchains have improved, but their central limitation is dependence of very high volume, frequency and quality of data (Kim et al., 2025). From a research perspective, exploring new avenues within data-science based approaches is a particularly interesting due to the recent rapid developments in the quantity and quality of data.

### **3 Theoretical Framework (1.5 pages)**

- 1: Some shipping- specific theory about operational inefficiency? - Maybe some theory that leads us to how we can i) quantify it and ii) identify sources of it
- 2: Data-science based approach to identify sources of inefficiency

#### **3.1 Hypotheses**

### **4 Dataset (1 page)**

#### **4.1 Origin**

- background on the workshop - Quick high-level rundown of a couple relevant ship particulars

## 4.2 Contents

# 5 Methodology

## 5.1 Modelling framework

It is intended to use statistical modelling to separate an estimated “optimal” operating speed from observed performance. First, an “optimal-speed” benchmark will be constructed by fitting a predictive model of vessel speed using data from a period in which operation can reasonably be assumed close to optimal, such as the first month following a hull and propeller cleaning event. This benchmark model will then be used to compute speed loss at each operating point as the difference between predicted optimal speed and observed speed. Second, a separate explanatory model will be estimated on data from a later, non-optimal period, in which fouling, engine depreciation, and other degradation mechanisms are expected to be present, with the dependent variable defined as the estimated speed loss and the regressors given by operational and environmental covariates in the dataset. The latter will be built for an inferential purpose, i.e., the primary purpose will be to deconstruct to infer what variables led to a given prediction. This will allow for the attribution of developments in inefficiency to certain variables and, in extension, parts of the ship. The approach to modelling is shown in 2.

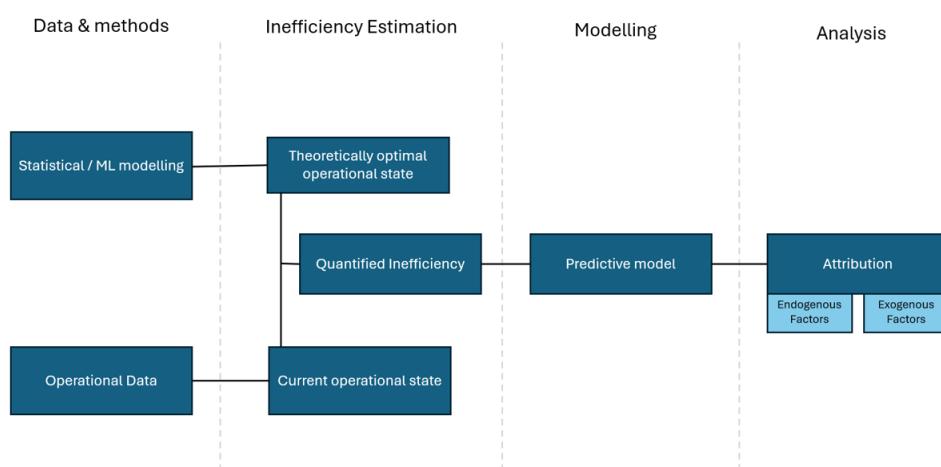


Figure 2: Visualisation of approach

## **5.2 models**

To predict the speed loss, It is suggested to employ and compare multiple models of varying complexity for two reasons. Firstly, it increases the likelihood of achieving a satisfactory predictive power. Secondly, it allows for the evaluation of the accuracy/complexity trade-off which is particularly relevant in edge computing scenarios such as onboard deployment. Applying three different models are suggested as a starting point. As a simple baseline model, it is suggested to use a linear regression due to its simplicity and intuitive interpretation of coefficients. As a natural progression of complexity, it is suggested to use a decision tree since it can model more complex relationships while still staying relatively interpretable. Lastly, it is suggested to include neural network-based model to have a high complexity benchmark to inform the complexity/trade-off analysis. The best predictive model should be chosen for a inferential analysis as described in 5. It is suggested to take an experimental approach and also look into other models to find the best model.

For interpretation and attribution, the inferential analysis will be tied to the best-performing predictive model, while maintaining a consistent attribution framework across model classes. For the linear regression, attribution follows directly from estimated coefficients, interpreted as marginal associations conditional on the included covariates and modelling assumptions. For the decision tree model, local and global drivers of predicted speed loss will be quantified using TreeSHAP, which decomposes each prediction into additive feature contributions while preserving consistency properties for tree-based models. For the neural-network model, SHAP values will be used analogously to obtain feature-level contributions to each predicted speed-loss outcome, enabling a comparable attribution of operational and environmental drivers even under highly nonlinear specifications.

## **6 Expected Results and Applications (1.5 page)**

## References

- Coraddu, A., Oneto, L., Baldi, F., Cipollini, F., Atlar, M., & Savio, S. (2019). Data-driven ship digital twin for estimating the speed loss caused by the marine fouling. *Ocean Engineering*, 186, 106063. <https://doi.org/10.1016/j.oceaneng.2019.05.045>
- Dong, G., Huang, J., Yao, C., Feng, D., & Meng, F. (2023). Numerical study of second-order wave forces using a rankine panel method. *IOP Conference Series: Materials Science and Engineering*, 1288(1), 012028. <https://doi.org/10.1088/1757-899X/1288/1/012028>
- Duan, K., Li, Q., Liu, Y., Jiang, H., Wang, S., & Zhang, M. (2024). Data-driven approach to evaluate the impact of hull roughness on main engine load of river-sea ships. *Ocean Engineering*, 312, 119239. <https://doi.org/10.1016/j.oceaneng.2024.119239>
- Froude, W., Abell, S. W., Gawn, R., & Duckworth, A. (1955). *The papers of william froude, 1810-1879: With a memoir by sir westcott abell and an evaluation of william froude's work by rwl gawn: Collected into one volume*. Institution of Naval Architects.
- Kim, Y., Gupta, P., & Steen, S. (2025). A comprehensive review of data processing for ship performance analysis. *Applied Ocean Research*, 162, 104737. <https://doi.org/10.1016/j.apor.2025.104737>
- Russell, J. S. (1839). Iii.—experimental researches into the laws of certain hydrodynamical phenomena that accompany the motion of floating bodies, and have not previously been reduced into conformity with the known laws of the resistance of fluids. *Transactions of the Royal Society of Edinburgh*, 14(1), 47–109. <https://doi.org/10.1017/S0080456800021451>
- Schultz, M. (2007). Effects of coating roughness and biofouling on ship resistance and powering. *Biofouling*, 23, 331–41. <https://doi.org/10.1080/08927010701461974>
- Wang, X., Zhang, R., Zhao, J., & Cao, P. (2022). Study on wave added resistance of ships in oblique waves based on panel method. *Journal of Ocean University of China*, 21, 773–781. <https://doi.org/10.1007/s11802-022-5074-3>

# **Appendix**

## **1 [Appendix Section Title]**

[Your appendix content goes here.]