Time-Series Analysis Approach for Improving Energy Efficiency of a Fixed-Route Vessel in Short-Sea Shipping

Mohamed Abuella, Hadi Fanaee, Slawomir Nowaczyk, Simon Johansson, and Ethan Faghani.

***Abstract*—Several approaches have been developed for improving the ship energy efficiency, thereby reducing operating costs and ensuring compliance with climate change mitigation regulations. Many of these approaches will heavily depend on measured data from onboard IoT devices, including operational and environmental information, as well as external data sources for additional navigational data. In this paper, we develop a framework that implements time-series analysis techniques to optimize the vessel's speed profile for improving the vessel's energy efficiency. We present a case study involving a real-world data from a passenger vessel that was collected over a span of 15 months in the south of Sweden. The results indicate that the implemented models exhibit a range of outcomes and adaptability across different scenarios. The findings highlight the effectiveness of time-series analysis approach for optimizing vessel voyages within the context of constrained landscapes, as often seen in short-sea shipping.**

***Index Terms*—Energy efficiency, voyage speed optimization, time-series analysis, short-sea shipping.**

I. Introduction

THE transportation of commercial freight by sea is the most environment-friendly transportation mode, since it emits less greenhouse gases (GHGs) per tonne of capacity and kilometer of distance, thus making resulting a smaller carbon footprint and lower impact on global climate. Short-Sea Shipping (SSS) is a commercial transportation mode that does not involve intercontinental cross-ocean. The SSS provides a cost-efficient and environment-friendly alternative for transportation by utilizing inland and coastal waterways to transport the commercial freight [1].

On the other hand, the SSS produces some negative effects on the natural habitats and polluting the air along the coasts of populated cites [2]. As a response to this, the International Maritime Organization (IMO) have conducted many studies and recommended standards and imposed policies for the maritime sector to reduce the carbon dioxide (CO_2) to 40% by 2030 and cut 50% of all GHGs by 2050, based on the emissions in 2008 [3].

Furthermore, COVID-19 pandemic has accelerated the digitalization of the entire shipping industry globally, and hence attracted a profound consideration to data collection and preparation stages [4]. The operational and some environmental conditions can be accessed through an Automatic Identification System (AIS) messages, which is a service developed by the International Maritime Organization (IMO) in 2002 to record the sensor measurements and send the vessel position information for the traffic between other ships and neighboring shores [5].

In a broader perspective, for improving the vessel's energy efficiency and harnessing more fuel savings and less GHG emissions, there are mainly two procedures. The first strategy is in the ship design stage, where the ship is built to obtain a body and equipped machinery that work efficiently. The second strategy is during the ship operation over the water or at the ports. This latter procedure can be achieved by adopting energy management plans that optimally enhance the energy efficiency and fuel consumption [6].

This paper proposes a data-driven framework for optimizing vessel speed profiles to improve energy efficiency in short-sea shipping.

The main contributions of this paper can be summarized as follow:

- Modeling of energy efficiency: Develop a data-driven model for voyage energy efficiency, including:
  A spatiotemporal aggregation of operation and navigation data from onboard and external sources to capture the impact of both spatial and temporal factors on voyage energy efficiency.
  Introduce an efficiency score that considers both total fuel consumption and voyage duration to measure the voyage energy efficiency.
- Data clustering: Clustering the data of voyages and sorting them based on their efficiency scores. This clustering enables the voyage optimization algorithm to learn more insights for better actions, either by selecting the best voyages or by eliminating the worst voyages.
- Time-series analysis models and comparative analysis: Four time-series based models are implemented as algorithms of voyage speed optimization. Then, a rigorous evaluation of their performance is conducted across different data clusters and using metrics that account for voyage efficiency.
- Practical implication: Demonstrate the significant effectiveness and practicality of the proposed approach for fixed-route vessels in short sea shipping, where the options for obtaining efficient voyages are limited. The

Manuscript received November 1, 2023; revised November 16, 2023.

This work was supported by the Sweden's innovation agency (Vinnova).(*Corresponding author: Mohamed Abuella*.)

Mohamed Abuella, Hadi Fanaee, and Slawomir Nowaczy are with Center for Applied Intelligent Systems Research (CAISR), Halmstad University, Kristian IV:s väg 80523, Halmstad, 30118 Sweden (e-mail: mohamed.abuella@hh.se; hadi.fanaee@hh.se; Slawomir.Nowaczyk@hh.se). Simon Johansson and Ethan Faghani are with CetaSol AB, Gothenburg, 41251, Sweden (e-mail: simon.johansson and ethan.faghani}@cetasol.com)

approach also aligns with the guidance of domain experts, adhering to safety and traffic considerations.

The remainder of the paper is organized as follows: related work is presented in Section II. The methodology of the proposed framework is covered in Section III. Section IV introduces the case study, including a description of the data used and the framework's implementation. The results and outcomes of the models are discussed in Section V. Finally, conclusions are addressed in Section VI.

II. Related Work

Various techniques can be applied for the voyage optimization problem, as demonstrated in prior studies such as [7]–[10]. Some of these approaches have been implemented specifically for short-sea shipping, as indicated in the domain of [11] and [12].

Different modeling and optimization algorithms for ship weather routing have been investigated thoroughly in [13]. It was found that the effectiveness of these algorithms strongly depends on specific requirements concerning the objectives, control variables and constraints as well as the implementation. The majority of these voyage optimization approaches are applied to ocean-crossing ships by controlling mainly the vessel speed and its route. Whereas, in the SSS, especially for fixed-route vessels like ours, there are fewer options available for voyage optimization. Therefore, the scope of this paper is mainly focused on short-sea shipping and the pertinent research literature.

In the third IMO GHG study 2014 [14], it was assumed that weather effects alone would be responsible for 15% of additional power margin on top of the theoretical propulsion requirements of ocean-going ships, and a 10% additional power requirement for coastal ships.

In a recent adaption of the Ship Traffic Emission Assessment Model (STEAM), propelling power is determined by wave height and directions, accounting for the environmental conditions in a highly detailed manner [5].

Recent research studies [15], [16] have explored energy-efficient routing for an electric ferry in Western Norway. They rely on operational data from onboard measurements and environmental conditions from the Norwegian Meteorological Institute, and proposed a hybrid physics-guided machine learning model for optimizing the ship route. Based on their findings, the hybrid model was showing an energy reduction of 3.7% compared to the actual consumption, simply by applying minor route and speed profile alterations as guided by the provided weather forecasts.

Researchers from Napa Ltd. in Finland conducted several studies on voyage optimization. In two of their studies [17] and [18]. These experts stated that their products of voyage optimization can achieve a fuel cost reduction of more than 10% with 2% to 4% savings from trim optimization and 6% to 8% from speed and route optimization. [19].

In the study conducted by Huotari et al. [20], where they used a combined model with both dynamic programming and convex optimisation to obtain optimal speed profiles. The fuel savings were around 1.1% and for voyages with substantial variance in environmental conditions, the fuel savings reached as high as 3.5%.

In the review paper by Wang et al. [21] numerous studies are explored, including coastal and inland shipping, also revealing a diverse range of fuel savings outcomes.

Meanwhile, regarding weather routing, specifically through speed and route optimization, it has been demonstrated in [22] that the potential savings of carbon dioxide emissions and fuel costs are in range of 5% to 10%.

III. Methodology

The main objective of this study is to improve the vessel voyage by optimizing its speed to enhance the vessel's energy efficiency. In other words, reducing the vessel's fuel consumption within constrained arrival time.

A. Framework of Voyage Optimization

The framework of the developed approach for improving the vessel voyages is depicted in Figure 1.

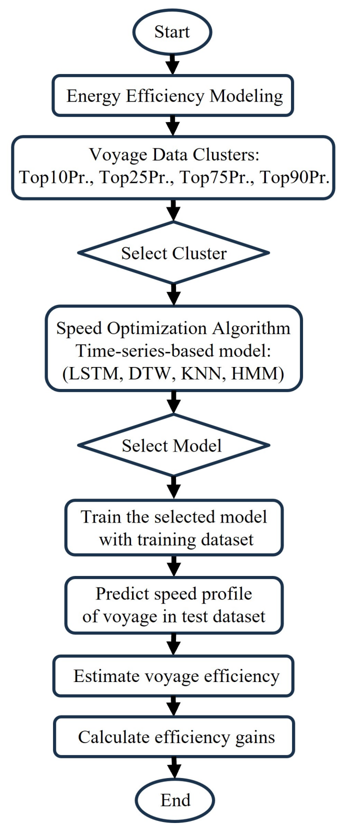


Fig. 1: Framework of vessel voyage optimization.

Data Processing and Clustering is a vital component of our framework. The primary objective is to identify and sort the voyages based on their efficiency scores. Then, train the models with the sorted data clusters iteratively, to distill insights from the voyages with different behaviours. Thus, the trained models will gain valuable insights into the performance and operational patterns of the vessel.

B. Energy Efficiency Modeling

This part presents a mathematical and visual overview of the fundamental theoretical background that forms the basis for modeling vessel energy efficiency—an indispensable element within our comprehensive framework.

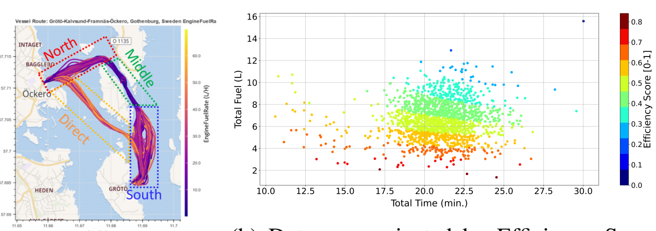
To estimate the vessel's energy efficiency in this framework of voyage optimization, we employed a previously developed model equipped with artificial intelligence (XAI) and machine learning techniques. More details about this energy efficiency modeling approach can be found in [23].

The efficiency score ($\mathrm{Eff}_{\mathrm{score}}$) is calculated from the normalized total fuel and time for every voyage, as following:

$$\mathrm{Eff}_{\mathrm{Score}} = 1 - \frac{2 \times [Fuel_{Tl_{Nm}} \times Time_{Tl_{Nm}}]}{[Fuel_{Tl_{Nm}} + Time_{Tl_{Nm}}]} \quad (1)$$

The efficiency score considers the proportional reduction in both fuel consumption and time, assessing the vessel's efficient use of resources during the voyage.

Figure 2 facilitates to visualize the process of aggregation for the vessel's voyages. First by illustrating the voyages in terms of space (i.e., latitude and longitude) as shown in Figure 2a, and second by representing the aggregated voyages as points in new dimensions of Efficiency Score versus total fuel and total time. The aggregated data and its new dimensions are projected as in Figure 2b.



(a) Vessel's route. (b) Data are projected by Efficiency Score vs. fuel and time.

Fig. 2: The vessel voyages and the aggregated data projected by the Efficiency Score.

The representation of voyages in terms of fuel and time is done by adopting the concept of the Efficiency Score (Eff-Score). The efficiency scores for all vessel's voyages are presented in Figure 2b. It is evident that the voyage with lower fuel and shorter time have higher efficiency scores, and vice versa.

There are four voyage data clusters, namely $Top75Pr$, $Top50Pr$, $Top25Pr$, and $Top10Pr$, as shown in Figure 3, These clusters are categorized on their respective Eff-Score percentiles, enabling a structured analysis of voyage efficiency across various percentile groups.

C. Time-Series Analysis Models for Voyage Optimization

For the purpose of vessel voyage optimization, we need a model to optimize the vessel's speed profile for improving the vessel's energy efficiency. This ideal model should mainly be able to:

- Model the temporal dependencies in vessel speed profiles.

Algorithm 1: Modeling of Energy Efficiency and Clustering of Voyages Data Based on Their Energy Efficiency

Data: Voyages data of the vessel
Result: Clusters of Voyages
Load the operational and navigational data, including speed, course, fuel, position, distance, and weather;
Tag the datapoints to its corresponding voyage, V_{id};
for *each voyage V_i in voyages data* **do**
  Calculate total fuel consumption and time for V_i;
  Normalize total fuel and time for V_i based on their maximum values of all voyages;
  Calculate the *Eff-Score* as described in Eq. (1), and assign it to all datapoints of this voyage V_i;
Initialize four empty lists for each cluster: $Top75Pr$, $Top50Pr$, $Top25Pr$, $Top10Pr$ (Percentiles of Eff-Scores);
for *each data point in all data* **do**
  Extract the Eff-Score of the data point;
  if *Eff-Score $\geq$ 0.4070* **then**
    Append the data point to $Top75Pr$;
  else if *Eff-Score $\geq$ 0.4623* **then**
    Append the data point to $Top50Pr$;
  else if *Eff-Score $\geq$ 0.5220* **then**
    Append the data point to $Top25Pr$;
  else if *Eff-Score $\geq$ 0.5730* **then**
    Append the data point to $Top10Pr$;

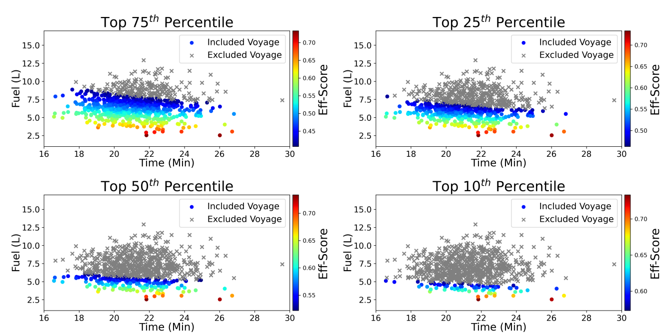


Fig. 3: Four clusters of voyages based on their efficiency.

- Incorporate external variables, such as weather conditions, into the modeling process.
- Adapt to changing conditions to provide real-time optimized speed profiles.

In order to meet these requirements, we adopt several time-series analysis models, which are described with some details as follow.

1) Long Short-Term Memory: Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture designed for sequential data analysis. It is particularly useful for modeling and predicting time-series data, making it a valuable tool for improving energy efficiency in vessel operations [24].

LSTM networks are well-suited for tasks involving time-series sequences of varying lengths, capturing long-term dependencies, and handling irregular temporal patterns. The LSTM architecture deals specifically to address the problem of vanishing gradients that often occurs in other RNN structures [25].

The LSTM's capability to learn the hidden insights from selected data, particularly in terms of vessel's efficiency scores, comes with a trade-off. The data clustering may result in a smaller dataset, which potentially limiting the LSTM's deep learning capabilities. Additionally, data clustering increases the susceptibility of LSTM to overfitting.

LSTM can be employed to model and predict the vessel's speed to improve its fuel efficiency, taking into account various factors such as weather conditions and operational parameters

2) Dynamic Time Warping: Dynamic Time Warping (DTW) is a distance-wise alignment algorithm that allows for the comparison of two time-series sequences that may have different lengths, time shifts, and speed variations [26]–[30]. The DTW algorithm can be utilized to measure the similarity between the given speed profile and other measured vessel speed profiles with higher efficiency scores. DTW is capable to capture the temporal dependency of speed profiles, even in short-sea shipping where the options of control the ship to improve its energy efficiency are limited.

Thus, the DTW identifies the efficient observed speed profile that most similar to the given speed profile to be chosen as suggested an optimized speed profile for the current journey of the vessel that can lead to improved energy efficiency.

3) k-Nearest Neighbors: The k-Nearest Neighbors (kNN) algorithm is a non-parametric model that makes no assumptions and operates on the principle of determining an unknown observation's class by measuring distances to nearby observations, attributing the observation to the majority class of its nearest neighbors. The parameter k represents the number of neighboring observations considered when classifying a given observation, and its value is determined through a search for the optimal choice that maximizes accuracy in the training set.

The paper by Cover and Hart [31] is a foundational work in pattern recognition and machine learning, introduced the KNN algorithm for pattern classification by utilizing the nearest neighbors to classify data points based on majority vote.

KNN can be employed to analyze and predict energy consumption patterns based on similar historical data. By considering the nearest neighbors of a given operational scenario, KNN provides a straightforward approach to making energy-efficient decisions. The choice of the parameter 'k,' representing the number of neighbors, plays a crucial role in the accuracy of KNN-based models.

4) Hidden Markov Model: Hidden Markov Models (HMMs) are based on probabilistic modeling and use a variety of techniques from the statistical modeling, and they are widely used in time-series analysis and pattern recognition.

The work by Rabiner in 1989 [32] is an essential reference in the application of HMM. This tutorial delved into the formulation of a statistical method for representing speech, showcasing a successful HMM system implementation with a focus on discrete or continuous density parameter distributions.

In the context of improving energy efficiency in constrained environments, HMMs can be applied to model the underlying patterns and transitions in ship operational data, allowing for more informed decisions on optimizing energy consumption.

The HMM estimates three main weather states and dynamically adjusts the SpeedOverGround (SOG) predictions. During calm weather, it selects the maximum observed speed in the historical data at that condition. In moderate conditions, it uses the average speed, while in rough weather, it relies on the minimum speed profiles. For instance, Eq. 2 indicates that the predicted speed ($\mathrm{SOG}_{\mathrm{Pred}_i}$) by HMM is the minimum value of measured speed profiles in rough weather state, and likewise for other weather states.

$$SOG_{\text{Pred.}_i} = \min(SOG_{\text{Meas.}} \,|\, \text{Rough weather}) \tag{2}$$

This adaptability to changing weather conditions enhances voyage optimization, making it a crucial component of the overall framework.

IV. Case Study

A. Data Collection

The ship's onboard data have been received from our industry partner CetaSol AB in Gothenburg [33]. The data has been gathered over a period of 15 months, between January 2020 and March 2021. It has a 3Hz frequency and records about the ship's position, course direction, and speed. It is also including some of operational and meteorological data, such as fuel rate, engine speed, torque, acceleration, wind speed and direction.

Some information about the ship and its voyage can be found on Marine Traffic website [34].

Other weather variables such as wave height and sea current speed and direction have been collected from external sources, Copernicus Marine Service [35] and Stormglass [36] APIs.

TABLE I: The navigational variables and their data sources.

Variable	Source	Variable	Source
Latitude	Onboard	WindSpeed_cps	Copernicus
Longitude	Onboard	WindDirection_cps	Copernicus
SpeedOverGround	Onboard	WaveHeight	Copernicus
HeadingMagnetic	Onboard	WaveDirection	Copernicus
Pitch	Onboard	WindSpeed_sg	Stormglass
Roll	Onboard	WindDirection_sg	Stormglass
WindSpeed_onb	Onboard	CurrentSpeed	Stormglass
WindDirection_onb	Onboard	CurrentDirection	Stormglass

B. Data Preparation and Validation

The external weather data are past forecasts (hindcasts), which have reanalysed to become hourly in temporal resolution and with 0.25 to 0.5 degree as a spatial resolution. Trilinear interpolation in time and space dimensions has been applied on external weather data to be more suitable for time and position frames of the given vessel routing. Therefore, the weather and onboard data are used in this analysis with a temporal resolution of 1-minute in average.

The data validation is conducted through the cruising-speeds mode is to reduce the other vessel effects on the fuel consumption, and thus, producing graphs that can be then compared with the general ship's standard performance. The operational and weather data validation is carried out visually, as shown in Figure 4.

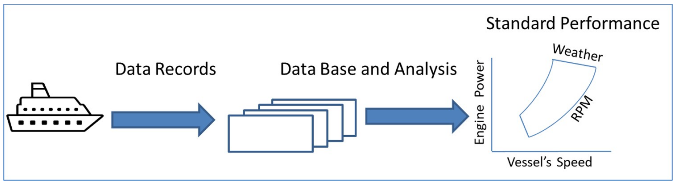


Fig. 4: Vessel's data analytics and standard performance graph for the case study [37]

In a prior publication [23], we have addressed the challenges of data representation and energy modeling in short-sea shipping (SSS), by proposing a data-driven modeling approach of voyage efficiency, which combines and aggregates data from multiple sources and seamlessly integrates explainable artificial intelligence (XAI) to attain clear insights about the energy efficiency for a vessel in SSS.

C. Implementation of the Framework

The implementation of our approach of a time-series analysis-based voyage optimization framework for a fixed-route vessel of our case study is depicted in Figure 1, and the step-by-step process is described by algorithms 1 and 2. For more detailed information about setting up the models and their specifications including various parameters, you may refer to the source codes, which are developed in Python 3.9.7 to produce the results of this study. These source codes are available at: https://github.com/MohamedAbuella/TSA4EESSS.

V. Results and Discussion

The vessel has a fixed-route which starts from the southern port to the northern port or vice versa. This route can be divided into four segments, specifically North, Middle, South, and Direct, as depicted in Figure 2a.
Cruising speeds are more common in North, South, and Direct segments of the vessel's route. Meanwhile, in the Middle segment, the vessel typically operates at maneuvering speeds, due to the presence of two ports located on west and east sides of the canal.

We have first conducted a statistical analysis on the dataset. Figure 5 illustrates some important statistic for all aggregated voyage, with regard to the accumulated fuel, time, and distance at different route segments. The route segments are depicted in Figure 2a.

As it can be seen from Table II, the difference of fuel consumption of the and cruising speeds is 5.47%, so that and also based on the recommendations from domain experts in compliance with maritime regulations including safety and traffic considerations, it might be more practical to primarily focus on optimizing cruising speed. After that we have implemented the framework for improving the vessel's energy

Algorithm 2: Speed Optimization Models for Improving Voyage Efficiency

Data: Refer to **Algorithm 1** for data processing and clustering.
Add $\mathrm{SOG}_{\mathrm{Meas.}}$ and Weathers to Inputs.
for *each C_k in CT (sorted by Eff-Score)* **do**
  Set training dataset to voyages $\in C_k$;
  Set test dataset to voyages $\notin C_k$;
  for *each **model** in [LSTM, DTW, KNN, HMM]* **do**
    Train the **model** with training dataset C_k;
    `// Skip DTW`
    for *each Voyage v_i in test dataset* **do**
      if *model is LSTM* **then**
        Recall trained LSTM to predict $\mathrm{SOG}_{\mathrm{Pred.}_i}$.
      else if *model is DTW* **then**
        for *each voyage v_j in training dataset cluster C_k* **do**
          Measure the similarity of voyage v_i compared to v_j;
        Set $\mathrm{SOG}_{\mathrm{Pred.}_i}$ to $\mathrm{SOG}_{\mathrm{Meas.}_j}$ of the most similar v_j;
      else if *model is KNN* **then**
        Recall the trained kNN to predict $\mathrm{SOG}_{\mathrm{Pred.}_i}$;
      else if *model is HMM* **then**
        Recall the trained HMM to estimate three weather states.
        if *Weather is Calm* **then**
          Set $\mathrm{SOG}_{\mathrm{Pred.}_i}$ to $\mathbf{max}(\mathrm{SOG}_{\mathrm{Calm}})$;
        else if *Weather is Moderate* **then**
          Set $\mathrm{SOG}_{\mathrm{Pred.}_i}$ to $\mathbf{mean}(\mathrm{SOG}_{\mathrm{Moderate}})$;
        else
          Set $\mathrm{SOG}_{\mathrm{Pred.}_i}$ to $\mathbf{min}(\mathrm{SOG}_{\mathrm{Rough}})$
      Energy Efficiency Estimation:;
      Estimate voyage v_i efficiency of both $\mathrm{SOG}_{\mathrm{Meas.}_i}$ and $\mathrm{SOG}_{\mathrm{Pred.}_i}$.;
    Evaluation Stage:;
    Calculate efficiency gains of test voyages by **model** trained with C_k;

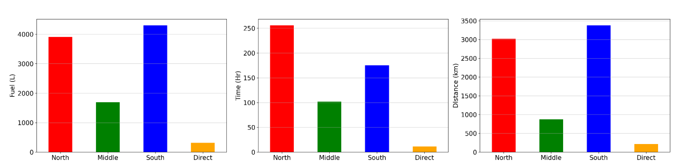


Fig. 5: Barplots for statistics of Fuel, Time, and Distance in different route segments

efficiency, as presented in IV-C of Section IV. Then, we evaluate the results, and for a sake of a fair comparison, we are injected both the actual measured and optimized speed profiles into the same estimation model of energy efficiency to predict fuel and time before and after the improving framework of energy efficiency is being implemented. Once the fuel and

TABLE II: Statistics of the dataset for different speed modes

Variable	Speeds		Difference (%)
	All	Cruising	
Fuel, total (Liter)	1329.2	1256.62	5.47%
Time, total (Hour)	48.04	22.32	53.57%
Distance, total (km)	608.72	349.2	42.68%
Speed, average (m/s)	2.67	1.67	37.5%

TABLE III: Average efficiency gains (Eff. Gains %, see Eq. 3) and counts of improved voyages (Eff. Improved #) out of 162 voyages in the test dataset.

Cluster	Efficiency Metric	LSTM	DTW	KNN	HMM
Top10Pr	Eff. Gains (%)	2.61	3.20	2.13	6.05
	Eff. Improves (#)	134	127	114	139
Top25Pr	Eff. Gains (%)	2.38	3.23	1.58	1.30
	Eff. Improves (#)	129	128	107	107
Top50Pr	Eff. Gains (%)	0.97	2.58	0.98	7.34
	Eff. Improves (#)	100	117	106	140
Top75Pr	Eff. Gains (%)	-0.84	2.28	0.50	9.31
	Eff. Improves (#)	60	119	93	141
Average	Eff. Gains (%)	1.28	2.82	1.30	6.00
	Eff. Improves (#)	105.75	122.75	105.00	131.75

time estimated, we compute the efficiency to determine how much energy has been saved.

One of our main metrics to evaluate the model performance of voyage efficiency optimization is the gain of efficiency scores, as represented in (3).

$$Eff._{Gain} = \frac{Eff.Score_{Pred.} - Eff.Score_{Meas.}}{Eff.Score_{Meas.}} \times 100 \tag{3}$$

Where $Eff.Score_{Meas.}$ and $Eff.Score_{Pred.}$ represent the voyage efficiency obtained with measured and predicted speed profiles, respectively.

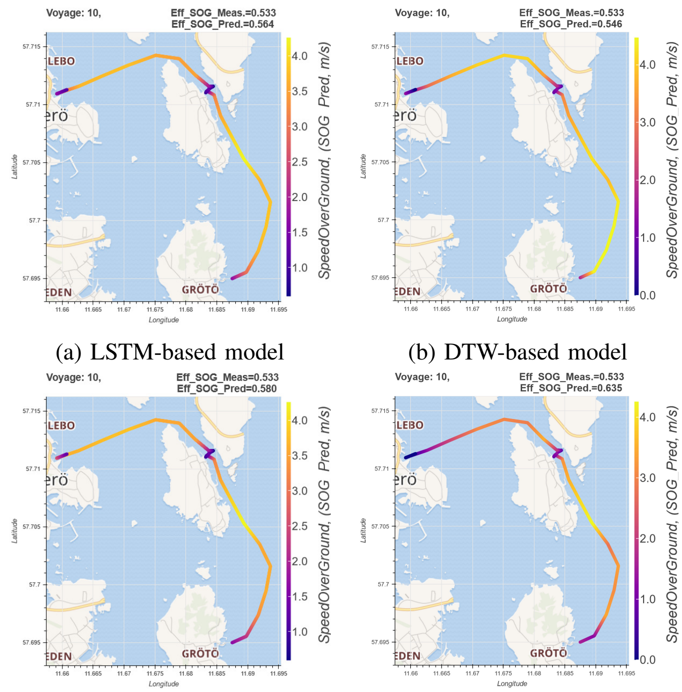
TABLE IV: Average and deviation of gains (%) of Eff-Scores in three weather states, when models are trained with the four data clusters

Model	LSTM		DTW		KNN		HMM	
Weather	Avg	Std	Avg	Std	Avg	Std	Avg	Std
Calm	0.17	4.72	2.21	4.3	1.17	3.29	3.96	6.19
Moderate	1.53	4.12	3.4	4.47	1.26	4.19	5.33	6.56
Rough	1.94	3.48	2.84	4.9	1.42	3.65	8.17	9.08
Average	1.21	4.11	2.82	4.56	1.28	3.71	5.82	7.28

As shown in Table III, the HMM model achieves the highest average efficiency gain of 6.00%, followed by the DTW model (2.82%), the KNN model (1.30%), and the LSTM model (1.28%). In terms of the number improved voyages out of 162 voyages in test dataset, the HMM model also improves the energy efficiency of the most average number of improved voyages (131.75 out of 162 voyages).

The HMM model achieves its best performance when trained on the Top75Pr cluster, which includes voyages with lower Eff-Scores and frequently encountered adverse weather conditions. Such performance underscores the HMM model's capability to learn the hidden patterns between the vessel speed and weather states, ultimately facilitates for developing more efficient speed profiles.

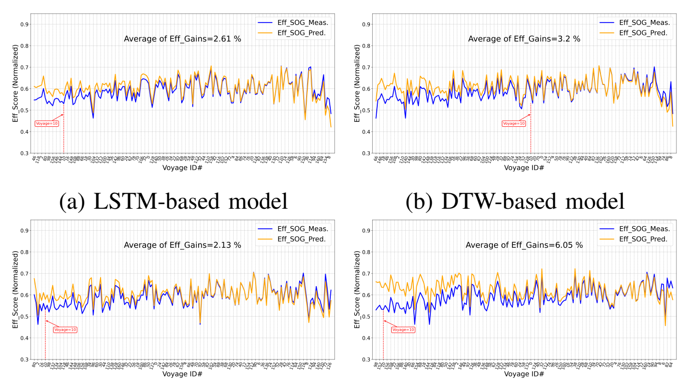
Table IV indicates that the HMM model yields the highest average efficiency gain in all three weather states, namely calm, moderate, and rough. The HMM model also has highest deviation of gains in all three weather states, indicating its adaptability to produce optimized speed profiles for more efficient voyages.



(a) LSTM-based model (b) DTW-based model

(c) KNN-based model (d) HMM-based model

Fig. 6: Predicted speed profile for a test voyage. From four time-series based models incorporate weather data as inputs and are trained by Top10Pr cluster.



(a) LSTM-based model (b) DTW-based model

(c) KNN-based model (d) HMM-based model

Fig. 7: Sorted Eff-Scores based on their gain, for 162 test voyages. From four time-series based models incorporate weather data as inputs and are trained by Top10Pr cluster, $Eff._{\text{Gain}}$ as in Eq. 3.

We also present the results of our analysis through two key visualizations. The first plot illustrates the predicted speed profiles for a test voyage, which are generated by four time-series based models that incorporate weather data as inputs. These models were trained using data from the Top10Pr cluster. The second plot showcases the sorted Eff-Scores for 162 test voyages, based on their efficiency gains. These models, similar to the first plot, also integrate weather data as inputs and were trained by the Top10Pr cluster. The efficiency gain is quantified using the formula as defined in Eq. 3.

In summary, the HMM-based model is the most effective model for improving energy efficiency for a vessel voyage in short sea. The HMM model is able to learn the complex relationships between the input features (e.g., speed and weather) and the output feature (Eff-Score), even in different weather conditions.

VI. Conclusion

The study employs four distinct models: LSTM, DTW, KNN, and HMM, to optimize vessel speed profiles with the objective of enhancing energy efficiency in short sea voyages. The key observation is that model performance varies significantly across these algorithms.
However, the performance of the models varies depending on the data cluster used to train the model and the weather conditions.

We developed a data-driven framework for optimizing vessel speed profiles to improve energy efficiency in SSS. The framework integrates a data-driven modeling approach to energy efficiency with the DTW algorithm. We evaluated the added value of the framework using a real-world dataset and found that it can effectively improve vessel energy efficiency, especially with limited options, which are common in short-sea shipping.

DTW exhibits the ability to capture temporal dependencies within speed profiles, especially within the constraints of short-sea shipping where opportunities for actively controlling the vessel to enhance its energy efficiency are restricted

Although the KNN can handle multivariate data and incorporating additional features like weather conditions, in this case study, the DTW performs better due to its specialized handling of time-dependent data and inherent patterns.

The result findings emphasize that in terms of searching the best behavior of vessel from the observed data, the DTW exhibits superior performance compared to LSTM and KNN, since the DTW selects the best measured speed profiles. On the other hand, the HMM is the most effective approach in our study. where the HMM optimizes these measured speed profiles further by offering strategies to them informed by their weather states. The study also reveals that the HMM model exhibits notable adaptability to different weather states (Calm, Moderate, and Rough). In each weather state, the HMM consistently delivers efficiency gains, indicating its ability to adapt speed profiles according to varying environmental conditions. This adaptability is crucial for real-world maritime applications, where weather can change rapidly.

Acknowledgment

This research project is funded by Sweden's innovation agency (Vinnova).
The authors wish to thank the diverse group at the Center for Applied Intelligent Systems Research (CAISR), Halmstad University, for helpful discussions.

Appendix
Supplementary Materials

The source codes that are implemented on Python 3.9.7 to produce the results are available at: https://github.com/MohamedAbuella/TSA4EESSS The dataset is private and cannot be shared due to the crucial commercial interests of the startup company operating the iHelm system.

References

[1] "Development of short sea shipping," [Online]. Available: https://www7.transportation.gov/testimony/development-short-sea-shipping, Feb. 2007.

[2] P. Donner and T. Johansson, "Sulphur directive, short sea shipping and corporate social responsibility in a eu context," *corporate social responsibility in the maritime industry*, pp. 149–166, 2018.

[3] J. D. Ampah, A. A. Yusuf, S. Afrane, C. Jin, and H. Liu, "Reviewing two decades of cleaner alternative marine fuels: Towards imo's decarbonization of the maritime transport sector," *Journal of Cleaner Production*, vol. 320, p. 128871, 2021.

[4] W. Bank, *Accelerating Digitalization: Critical Actions to Strengthen the Resilience of the Maritime Supply Chain*. Washington: World Bank, 2020.

[5] C. IMO, "Fourth imo ghg study 2020," 2020.

[6] T. P. Zis, H. N. Psaraftis, and L. Ding, "Ship weather routing: A taxonomy and survey," *Ocean Engineering*, vol. 213, p. 107697, 2020.

[7] H. Chen and P. Ballou, "Art and science of ship voyage optimization: A critical review," 2021.

[8] L. Walther, A. Rizvanolli, M. Wendebourg, and C. Jahn, "Modeling and optimization algorithms in ship weather routing," *International Journal of e-Navigation and Maritime Economy*, vol. 4, pp. 31–45, 2016.

[9] A. Fan, J. Yang, L. Yang, D. Wu, and N. Vladimir, "A review of ship fuel consumption models," *Ocean Engineering*, vol. 264, p. 112405, 2022.

[10] M. H. Moradi, M. Brutsche, M. Wenig, U. Wagner, and T. Koch, "Marine route optimization using reinforcement learning approach to reduce fuel consumption and consequently minimize co2 emissions," *Ocean Engineering*, vol. 259, p. 111882, 2022.

[11] A. Zakaria, A. Md Arof, and A. Khabir, "Instruments utilized in short sea shipping research: A review," *Design in Maritime Engineering*, pp. 83–108, 2022.

[12] M. Grifoll, F. X. Martínez de Osés, and M. Castells, "Potential economic benefits of using a weather ship routing system at short sea shipping," *WMU Journal of Maritime Affairs*, vol. 17, no. 2, pp. 195–211, 2018.

[13] L. Walther, A. Rizvanolli, M. Wendebourg, and C. Jahn, "Modeling and optimization algorithms in ship weather routing," *International Journal of e-Navigation and Maritime Economy*, vol. 4, pp. 31–45, 2016.

[14] T. W. Smith, J. Jalkanen, B. Anderson, J. Corbett, J. Faber, S. Hanayama, E. O'keeffe, S. Parker, L. Johansson, L. Aldous *et al.*, "Third imo greenhouse gas study 2014," 2015.

[15] P. R. Bellingmo, A. Pobitzer, U. Jørgensen, and S. P. Berge, "Energy efficient and safe ship routing using machine learning techniques on operational and weather data," in *20th International Conference on Computer Applications and Information Technology in the Maritime Industries*, 2021.

[16] U. Jørgensen, P. R. Belingmo, B. Murray, S. P. Berge, and A. Pobitzer, "Ship route optimization using hybrid physics-guided machine learning," in *Journal of Physics: Conference Series*, vol. 2311, no. 1. IOP Publishing, 2022, p. 012037.

[17] K. Sugimoto, "Digital twin for monitoring remaining fatigue life of critical hull structures."

[18] M. Haranen, S. Myöhänen, and D. S. Cristea, "The role of accurate now-cast data in shi p efficiency analysis," in *2nd Hull Performance & Insight Conference*, 2017, pp. 25–38.

[19] M. Wingrove, "Owners cut fuel costs by 10 per cent," vol. 10, no. 4, p. 42–43, 2016.

[20] J. Huotari, T. Manderbacka, A. Ritari, and K. Tammi, "Convex optimisation model for ship speed profile: Optimisation under fixed schedule," *Journal of Marine Science and Engineering*, vol. 9, no. 7, p. 730, 2021.

[21] K. Wang, J. Wang, L. Huang, Y. Yuan, G. Wu, H. Xing, Z. Wang, Z. Wang, and X. Jiang, "A comprehensive review on the prediction of ship energy consumption and pollution gas emissions," *Ocean Engineering*, vol. 266, p. 112826, 2022.

[22] L. Walther, *Development of a Weather Routing System for Analysis and Optimization of Ship Voyages*. Fraunhofer Verlag, 2021.

[23] M. Abuella, M. A. Atoui, S. Nowaczyk, S. Johansson, and E. Faghani, "Data-driven explainable artificial intelligence for energy efficiency in short-sea shipping," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2023, pp. 226–241.
[24] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[25] R. C. Staudemeyer and E. R. Morris, "Understanding lstm–a tutorial into long short-term memory recurrent neural networks," *arXiv preprint arXiv:1909.09586*, 2019.
[26] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proceedings of the 3rd international conference on knowledge discovery and data mining*, 1994, pp. 359–370.
[27] E. Keogh and C. A. Ratanamahatana, "Exact indexing of dynamic time warping," *Knowledge and information systems*, vol. 7, pp. 358–386, 2005.
[28] H. Ding, G. Trajcevski, P. Scheuermann, X. Wang, and E. Keogh, "Querying and mining of time series data: experimental comparison of representations and distance measures," *Proceedings of the VLDB Endowment*, vol. 1, no. 2, pp. 1542–1552, 2008.
[29] T. Rakthanmanon, B. Campana, A. Mueen, G. Batista, B. Westover, Q. Zhu, J. Zakaria, and E. Keogh, "Searching and mining trillions of time series subsequences under dynamic time warping," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2012, pp. 262–270.
[30] A. Bagnall, J. Lines, A. Bostrom, J. Large, and E. Keogh, "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances," *Data mining and knowledge discovery*, vol. 31, pp. 606–660, 2017.
[31] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.
[32] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
[33] "CetaSol AB ," [Online]. Available: https://cetasol.com.
[34] "Marine Traffic," [Online]. Available: https://www.marinetraffic.com/en/ais/details/ships/shipid:1088282/mmsi:265513810/imo:8602713/vessel:BURO, Jul. 2022.
[35] "Copernicus Marine Service ," [Online]. Available: https://marine.copernicus.eu.
[36] "StormGlass API ," [Online]. Available: https://stormglass.io.
[37] J. Carlton, *Marine propellers and propulsion*, 2nd ed. Oxford: Butterworth-Heinemann, January 2007.

Mohamed Abuella received his M.S. and PhD degrees in Electrical and Computer Engineering from Southern Illinois University at Carbondale and University of North Carolina at Charlotte, in 2012 and 2018 respectively. He is a postdoctoral researcher at Halmstad University since 2022. His research interests include energy analytics and AI for sustainability.

Hadi Fanaee-T is an Associate Professor at Halmstad University, Sweden. He completed his PhD (with distinction) in Computer Science, under supervision of Professor Joao Gama at the Faculty of Science of University of Porto, Portugal in November 2015. He was the finalist in ERCIM Cor Baayen Young Researcher Award 2017. Prior to this position he was a postdoctoral fellow at the department of biostatistics, University of Oslo, Norway, and also worked as a postdoctoral researcher in European FP7 Project "MAESTRA" at INESC TEC research institute, Portugal. His main research interests are interdisciplinary applications of tensor decompositions, data fusion, anomaly/event detection and spatiotemporal data mining. He is the first-author of several journal and conference papers. He has served as a PC member to over 20 prestigious conferences (e.g. IJCAI, AAAI, ECML-PKDD, ISMIS, ACM SAC, IEEE DSAA, etc.) and also reviewer to several high-impact journals (e.g. TDKE, TKDD, DAMI, ML, KAIS, KBS, CSUR.

Slawomir Nowaczyk is a Professor in Machine Learning at the Center for Applied Intelligent Systems Research, Halmstad University, Sweden. He received his MSc degree from Poznan University of Technology in 2002 and his PhD from the Lund University of Technology in 2008. During the last decades, his research has focused on machine learning, knowledge representation, and self-organising systems. The majority of his work concerns industrial data streams, often with predictive maintenance as the goal. Given that accurate and relevant labels are usually impossible to obtain in such settings, Slawomir's contributions primarily take advantage of proxy labels, such as transfer learning and multi-task learning, or concern semi-supervised and unsupervised modelling. He is a board member of the Swedish AI Society and a research leader for the School of Information Technology at Halmstad University. Slawomir has led multiple research projects on applying Artificial Intelligence and Machine Learning in different domains, such as transport and automotive, energy, smart cities, and healthcare. In most cases, this research was done in collaboration with industry and public administration organisations – inspired by practical challenges and leading to tangible results and deployed solutions.

Simon Johansson is a MSc graduate of Chalmers University of Technology's program in Engineering Mathematics and Computational Science in 2020, currently works in Cetasol, a marine company specialising in CO2 reduction and energy optimisation for vessels. His practical application of computational methods and dedication to environmental sustainability align with his role, contributing to global efforts to mitigate climate change. Simon's commitment to advancing eco-friendly solutions in the marine industry reflects a seamless integration of academic excellence and real-world impact.

Ethan Faghani is the CEO and founder of Cetasol. Before Cetasol, Ethan was Chief Engineer of Automation and AI at Volvo Penta. Ethan has experience working with cutting-edge technologies in other transportation segments in both big enterprises and his own founded startup. Ethan obtained his Ph.D. in mechatronics from UBC and Innovation and Entrepreneurship from Stanford Business School.