

Supplementary Materials

Supplementary Materials for paper "Data-Driven Explainable Artificial Intelligence for Energy Efficiency in Short-Sea Shipping" submitted to ECML/PKDD 2023.

Source Codes

The source codes that are implemented on Python 3.9.7 to produce the results are available at: <https://anonymous.4open.science/r/ST4EESSS-37A7>

The dataset is private and cannot be shared due to the crucial commercial interests of the startup company operating the iHelm system.

The List of Available Signals

The list of the main collected variables of onboard data is presented in Table 7 below.

Table 7: List of the main collected variables in onboard data.

Name	Unit	Name	Unit
Timestamp	time	WindSpeed	m/s
Latitude	degree	WindDirection	rad
Longitude	degree	SpeedApparent	m/s
SpeedOverGround	m/s	Temperature	K
CourseOverGroundTrue	rad	Pressure	Pa
HeadingMagnetic	rad	AcceleratorPedalPosition	%
Pitch	rad	ActualEnginePercentTorque	%
Roll	rad	EngineSpeed	RPM
RateOfTurn	rad/s	Enginefuelrate	l/h

Trilinear Interpolation

Trilinear interpolation is a method of multivariate interpolation on a 3-dimensional regular grid. It approximates the value of a function at an intermediate point (x, y, z) within the local axial rectangular prism linearly, using function data on the lattice points.

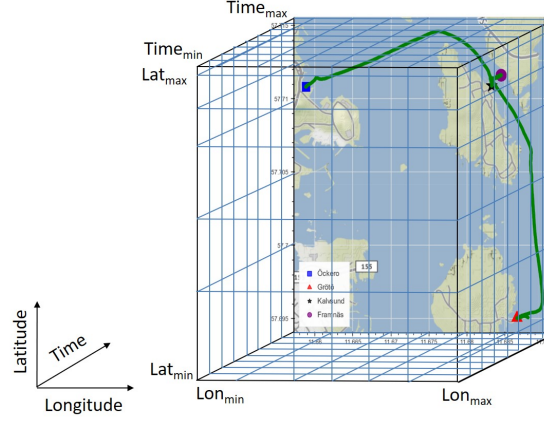


Fig. 10: Schematic diagram of data interpolation (trilinear interpolation).

Results of XGBoost Regression

The following tables contain the results for XGBoost regression models with four input cases and three outputs: EngineFuelRate, Total Fuel, and Efficiency Score.

Table 8: Results of XGBoost regression, where EngineFuelRate is the XGBoost output, and inputs are different cases.

Input Cases	Number of XGBoost Inputs	RMSE	R2	MAE
I	6	0.0948	0.6491	0.0696
II	12	0.0782	0.7615	0.0558
III	10	0.0792	0.7550	0.0572
IV	14	0.0783	0.7604	0.0559

Table 9: Results of XGBoost regression, where Total Fuel-Total is the XGBoost output, and XGBoost inputs with different cases.

Input Cases	Number of XGBoost Inputs	RMSE	R2	MAE
I	6	0.1001	0.1676	0.0789
II	12	0.0439	0.8400	0.0326
III	10	0.0542	0.7563	0.0411
IV	14	0.0447	0.8340	0.0330

Table 10: Results of XGBoost regression, where Efficiency Score is the XGBoost output, and with different cases of XGBoost inputs.

Input Cases	Number of XGBoost Inputs	RMSE	R2	MAE
I	6	0.0824	0.1625	0.0646
II	12	0.0387	0.8149	0.0283
III	10	0.0472	0.7256	0.0348
IV	14	0.0368	0.8324	0.0272

Ethical Statement

Hereby, **Mohamed Abuella** consciously assures that for the submitted research work data from electronic devices on the vessel had approval through the data access agreement with our industry partner.

Consent from the vessel's personnel was obtained through opt-out and their personal information kept private. In addition, there was an option was given for personnel to request a removal of their data from the study or analysis.

We have tried to rule out unrelated human characteristics from the data as much as possible.

Therefore, the data of this research is only relative to vessel operation and irrelevant to the personnel and passengers.

Date: 24/03/2023

Mohamed Abuella