Acknowledgement. This is the software infrastructure of ***the IAMU (International Association of Maritime Universities) research project titled "Data fusion and machine learning for ship fuel efficiency analysis: a small but essential step towards green shipping through data analytics" (Research Project No. 20210205_AMC). The materials and data in this project have been obtained through the support of IAMU and The Nippon Foundation in Japan***. This project has the following investigators: Yuquan (Bill) Du (Coordinator, Lead Applicant), Peggy Shu-Ling Chen, Nataliya Nikolova, Prashant Bhaskar, and Jiangang Fei from University of Tasmania (UTAS); Alessandro Schönborn from World Maritime University (WMU) as WMU-side Chief Investigator; and Zhuo Sun from Dalian Maritime University (DMU) as DMU-side Chief Investigator. Dr Xiaohe Li laid a good Python code fundation during his University Associate role at UTAS and conducted some preliminary experiments. Ms Yanyu Chen is the major Research Assistant of this project who has major contributions to Python code and conducting experiments under the supervision of Dr Yuquan (Bill) Du. Mr Jean-Louis Bertholier developed the Python code of collecting meteorological data for ships during his Assistant Engineer internship at World Maritime University. Warm help are received from ECMWF (Centre for Medium-range Weather Forecasts) and Copernicus Marine Service (CMEMS) when we wanted to automate the download process of meteorological data. This study has been conducted using E.U. Copernicus Marine Service Information; https://doi.org/10.48670/moi-00050. Hersbach et al. (2018) was downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store. The results of this study and trained machine learning models published contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

Reference. Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2018): ERA5 hourly data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 10-Sep-2021), 10.24381/cds.adbb2d47.

If you use our trained machine learning models to calculate or predict the fuel consumption rate (ton/day, ton/hour), kindly give us credits by citing our following papers because the models are from these papers.

*Xiaohe Li, Yuquan Du, Yanyu Chen, Son Nguyen, Wei Zhang, Alessandro Schönborn, Zhuo Sun, 2022. "Data fusion and machine learning for ship fuel efficiency modeling: Part I – voyage report data and meteorological data". JOURNAL NAME TO BE CONFIRMED, Vol XX, No. XX, pp XXX-XXX***.**

Yuquan Du, Yanyu Chen, Xiaohe Li, Alessandro Schönborn, Zhuo Sun, 2022a. "Data fusion and machine learning for ship fuel efficiency modeling: Part II – voyage report data, AIS data and meteorological data". JOURNAL NAME TO BE CONFIRMED, Vol XX, No. XX, pp XXX-XXX.

Yuquan Du, Xiaohe Li, Yanyu Chen, Alessandro Schönborn, Zhuo Sun, 2022b. "Data fusion and machine learning for ship fuel efficiency modeling: Part III – sensor data and meteorological data". JOURNAL NAME TO BE CONFIRMED, Vol XX, No. XX, pp XXX-XXX.

There are 130 trained models in total provided here (in the folder of "Trained Models/") , for four best datasets found in our three papers including ***Set1***, ***Set3Precise***, ***AIS5Precise***, and ***Sensor2***, using machine learning models including ***Extremely randomized trees (ET)***, ***Gradient Tree Boosting (GB)***, ***XGBoost (XG)***, ***Support Vector Machine (SVM)***, and ***Artificial Neural Networks (ANN)***, for ***Ships S1, S2, S3, S4, S5, S6, S7 and S8*** described in our papers.

Each trained model is stored in a separate file. For instance, the model stored in the file named "Ship_S1_GB_AIS5Precise" is the trained GB model for ship S1 over the dataset *AIS5Precise*. Similary, **the model stored in the file named "Ship_S5_ET_Sensor2" is the trained ET model for ship S5 over the dataset *Sensor2*.**

The following Python code demonstrates how we can load a machine learning model into Python workspace and forecast the fuel consumption rate (ton/day) given sailing speed, displacement/draft, trim, weather conditions, and sea conditions.

In [3]:
```
# Import machine learning models packages in Python
import os
import time
import joblib
import openpyxl
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

from sklearn import svm
from sklearn.neural_network import MLPRegressor
from sklearn.model_selection import learning_curve
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor, ExtraTreesRegressor, GradientBoostingRegressor
from sklearn.metrics import r2_score, mean_squared_error, mean_absolute_error, explained_variance_score

import xgboost as xgb
from xgboost.sklearn import XGBRegressor
```

In **Du et al. (2022a)**, the best dataset is ***AIS5Precise*. Using the trained model "Ship_S1_GB_AIS5Precise" as an example**, the following Python code demonstrates how we can load a machine learning model into Python workspace and forecast the fuel consumption rate (ton/day) given sailing speed, displacement/draft, trim, weather conditions, and sea conditions.

In [4]:
```
# Location of the trained machine learning model. Here, we use the model "Ship_S1_GB_AIS5Precise" for example.
# All the trained models are in the folder of "Trained Models/".
path_model = 'Trained Models/Ship_S1_GB_AIS5Precise'
```

```python
# If a model trained by AIS5Precise is used, arrange the values of X (input variables)
# and y (target/output variable, "Fuel consumption rate") in the following form:
# x_expr = ['Sailing speed', 'Trim','Displacement', 'Sea water temperature', 'Wind speed', 'Wind direction (Rel.)',
#           'Sea current speed', 'Sea current direction (Rel.)', 'Combined wave height', 'Combined wave direction (Rel.)']
# y_expr = 'Fuel consumption rate'.

# Sailing speed (knots); Min value = 12; Maximum value = 26;
sailingSpeed = 23;
# Trim (m); Min value = -5.0; Maximum value = 5.0;
trim = 2.3;
# Displacement (ton); Min value = 65270; Maximum value = 165070;
displacement = 82500
# Sea water temperature (degree); Min value = 1.2; Maximum value = 31;
seaWaterTemp = 23
# Wind speed (knots); Min value = 0 ; Maximum value = 30;
windSpeed = 3
# Wind direction (Rel.) (degree); Min value = 0; Maximum value = 180;
windDirection = 120
# Sea current speed (knots); Min value = 0; Maximum value = 8;
currentSpeed = 2
# Sea current direction (Rel.) (degree); Min value = 0; Maximum value = 180;
currentDirection = 4.5
# Combined wave height (m); Min value = 0; Maximum value = 8;
combinedWaveHight = 3.8
# Combined wave direction (Rel.) (degree); Min value = 0; Maximum value = 180;
combinedWaveDirection = 60


x_expr = pd.DataFrame([[sailingSpeed], [trim], [displacement], [seaWaterTemp], [windSpeed], [windDirection], [currentSpeed],
                       [currentDirection], [combinedWaveHight], [combinedWaveDirection]])

x_expr = x_expr.values.reshape(1,-1)
print(x_expr.shape)


# Load the trained machine learning model
model = joblib.load(path_model)

# Predict the y (fuel consumption rate, MT/day) values of the experimental data "x_expr"
y_pre = model.predict(x_expr)

print("The prediction of ship fuel consumption rate of machine learning model is (metric ton/day):", y_pre)
```

```
(1, 10)
```

```
The prediction of ship fuel consumption rate of machine learning model is (metric ton/day): [137.97731548]
```

In **Li et al. (2022)**, one of the best datasets is ***Set1*** (original voyage report data). **Using the trained model "Ship_S1_ET_Set1" as an example**, the following Python code demonstrates how we can load a machine learning model into Python workspace and forecast the fuel consumption rate (ton/day) given sailing speed, displacement/draft, trim, weather conditions, and sea conditions.

In [23]:
```
# Location of the trained machine learning model. Here, we use the model "Ship_S1_ET_Set1" for example.
# All the trained models are in the folder of "Trained Models/".
path_model = 'Trained Models/Ship_S1_ET_Set1'

# If a model trained by Set1 is used, arrange the values of X (input variables)
# and y (target/output variable, "Fuel consumption rate") in the following form:
# x_expr = ['Sailing speed', 'Trim','Displacement', 'Sea water temperature', 'Wind speed', 'Wind direction (Rel.)',
#           'Sea current speed', 'Sea current direction (Rel.)', 'Swell height', 'Swell direction (Rel.)']
# y_expr = 'Fuel consumption rate'.

# Sailing speed (knots); Min value = 12; Maximum value = 26;
sailingSpeed = 23;
# Trim (m); Min value = -5.0; Maximum value = 5.0;
trim = 2.3;
# Displacement (ton); Min value = 65270; Maximum value = 165070;
displacement = 82500
# Sea water temperature (degree); Min value = 1.2; Maximum value = 31;
seaWaterTemp = 23
# Wind speed (Level); Min value = 2 ; Maximum value = 9;
windSpeed = 3
# Wind direction (Rel.); Min value = 1; Maximum value = 5;
# For wind direction definition, see Figure 1 of Li et al. (2022): "A" - 1; "B"/"H" - 2; "C"/"G" - 3; "D"/"F" - 4; "E" - 5.
windDirection = 3
# Sea current speed (knots); Min value = 0.1; Maximum value = 3;
currentSpeed = 2
# Sea current direction (Rel.); Min value = 1; Maximum value = 5;
# For sea current direction definition, see Figure 1 of Li et al. (2022). "A" - 5; "B"/"H" - 4; "C"/"G" - 3; "D"/"F" - 2; "E" - 1.
currentDirection = 4
# Swell height (m); Min value = 1; Maximum value = 6.5;
swellHeight = 3.8
# Swell direction (Rel.) (Level); Min value = 1; Maximum value = 5;
swellDirection = 2


x_expr = pd.DataFrame([[sailingSpeed], [trim], [displacement], [seaWaterTemp], [windSpeed], [windDirection], [currentSpeed],
                      [currentDirection], [swellHeight], [swellDirection]])
x_expr = x_expr.values.reshape(1,-1)
```

```
print(x_expr.shape)

# Load the trained machine learning model
model = joblib.load(path_model)

# Predict the y (fuel consumption rate, MT/day) values of the experimental data "x_expr"
y_pre = model.predict(x_expr)

print("The prediction of ship fuel consumption rate of machine learning model is (metric ton/day):", y_pre)
```

```
(1, 10)
The prediction of ship fuel consumption rate of machine learning model is (metric ton/day): [129.27202961]
```

In **Li et al. (2022)**, another best dataset is ***Set3Precise*** (a fusion of voyage report and meteorological data). **Using the trained model "Ship_S5_SVM_Set3Precise" as an example**, the following Python code demonstrates how we can load a machine learning model into Python workspace and forecast the fuel consumption rate (ton/day) given sailing speed, displacement/draft, trim, weather conditions, and sea conditions.

In [26]:

```
# Location of the trained machine learning model. Here, we use the model "Ship_S5_SVM_Set3Precise" for example.
# All the trained models are in the folder of "Trained Models/".
path_model = 'Trained Models/Ship_S5_SVM_Set3Precise'

# If a model trained by Set3Precise is used, arrange the values of X (input variables)
# and y (target/output variable, "Fuel consumption rate") in the following form:
# x_expr = ['Sailing speed', 'Trim','Displacement', 'Sea water temperature', 'Wind speed', 'Wind direction (Rel.)',
#           'Sea current speed', 'Sea current direction (Rel.)', 'Swell height', 'Swell direction (Rel.)',
#           'Wind wave height', 'Wind wave direction (Rel.)', 'Combined wave height', 'Combined wave direction (Rel.)']
# y_expr = 'Fuel consumption rate'.

# Sailing speed (knots); Min value = 12; Maximum value = 26;
sailingSpeed = 23;
# Trim (m); Min value = -3.0; Maximum value = 3.5;
trim = 2.1;
# Displacement (ton); Min value = 68000; Maximum value = 135000;
displacement = 128391
# Sea water temperature (degree); Min value = 0; Maximum value = 35;
seaWaterTemp = 23
# Wind speed (knots); Min value = 0 ; Maximum value = 18;
windSpeed = 3.5
# Wind direction (Rel.) (degree); Min value = 0; Maximum value = 180;
windDirection = 65
# Sea current speed (knots); Min value = 0; Maximum value = 2;
currentSpeed = 0.55
# Sea current direction (Rel.) (degree); Min value = 0; Maximum value = 180;
```

```
currentDirection = 158
# Swell height (m); Min value = 0; Maximum value = 6;
swellHeight = 2.6
# Swell direction (Rel.) (degree); Min value = 0; Maximum value = 180;
swellDirection = 122
# Wind wave height (m); Min value = 1; Maximum value = 6.5;
windWaveHeight = 2.8
# Wind wave direction (Rel.) (degree); Min value = 0; Maximum value = 180;
windWaveDirection = 25
# Combined wave height (m); Min value = 0; Maximum value = 8;
combinedWaveHeight = 3.8
# Combined wave direction (Rel.) (degree); Min value = 0; Maximum value = 180;
combinedWaveDirection = 60


x_expr = pd.DataFrame([[sailingSpeed], [trim], [displacement], [seaWaterTemp], [windSpeed], [windDirection],
                       [currentSpeed], [currentDirection], [swellHeight], [swellDirection],
                       [windWaveHeight], [windWaveDirection], [combinedWaveHeight], [combinedWaveDirection]])
x_expr = x_expr.values.reshape(1,-1)
print(x_expr.shape)

# Load the trained machine learning model
model = joblib.load(path_model)

# Predict the y (fuel consumption rate, MT/day) values of the experimental data "x_expr"
y_pre = model.predict(x_expr)

print("The prediction of ship fuel consumption rate of machine learning model is (metric ton/day):", y_pre)
```

```
(1, 14)
The prediction of ship fuel consumption rate of machine learning model is (metric ton/day): [121.7081992]
```

In **Du et al. (2022b)**, the best dataset is ***Sensor2*** (a fusion of voyage report data, AIS data, and meteorological data). **Using the trained model "Ship_S5_XG_Sensor2" as an example**, the following Python code demonstrates how we can load a machine learning model into Python workspace and forecast the fuel consumption rate (ton/day) given sailing speed, displacement/draft, trim, weather conditions, and sea conditions.

In [6]:
```
# Location of the trained machine learning model. Here, we use the model "Ship_S5_XG_Sensor2" for example.
# All the trained models are in the folder of "Trained Models/".
path_model = 'Trained Models/Ship_S5_XG_Sensor2'

# If a model trained by Sensor2 is used, arrange the values of X (input variables)
# and y (target/output variable, "Fuel consumption rate") in the following form:
# x_expr = ['Sailing speed', 'Displacement', 'Trim', 'Wind speed', 'Wind direction (Rel.)',
```

```
#           'Sea current speed', 'Sea current direction (Rel.)',
#           'Combined wave height', 'Combined wave direction (Rel.)', 'Combined wave period, 'Sea water temperature',]
# y_expr = 'Fuel consumption rate'.

# Sailing speed (knots); Min value = 12; Maximum value = 26;
sailingSpeed = 23;
# Draft (m); Min value = 7; Maximum value = 14;
draft = 2
# Trim (m); Min value = -2.0; Maximum value = 5;
trim = 2.1;
# Wind speed (knots); Min value = 0 ; Maximum value = 40;
windSpeed = 20
# Wind direction (Rel.) (degree); Min value = 0; Maximum value = 180;
windDirection = 65
# Sea current speed (knots); Min value = 0; Maximum value = 3.5;
currentSpeed = 2.2
# Sea current direction (Rel.) (degree); Min value = 0; Maximum value = 180;
currentDirection = 56
# Combined wave height (m); Min value = 0; Maximum value = 8;
combinedWaveHeight = 3.8
# Combined wave direction (Rel.) (degree); Min value = 0; Maximum value = 180;
combinedWaveDirection = 60
# Combined wave period; Min value = 0; Maximum value = 15
combinedWavePeriod = 12
# Sea water temperature (degree); Min value = 0; Maximum value = 35;
seaWaterTemp = 23

x_expr = pd.DataFrame([[sailingSpeed], [draft], [trim], [windSpeed], [windDirection], [currentSpeed], [currentDirection],
                       [combinedWaveHeight], [combinedWavePeriod], [combinedWaveDirection], [seaWaterTemp]])
x_expr = x_expr.values.reshape(1,-1)
print(x_expr.shape)

# load the trained machine learning model
model = joblib.load(path_model)

# Predict the y (fuel consumption rate, MT/day) values of the experimental data "x_expr"
y_pre = model.predict(x_expr)

print("The prediction of ship fuel consumption rate of machine learning model is (metric ton/day):", y_pre* 24 / 1000)
```

```
(1, 11)
The prediction of ship fuel consumption rate of machine learning model is (metric ton/day): [115.13089]
```