Forecasting based on machine learning methods in the fuel monitoring system

Aleksandr Gozhyj[1,*,†], Irina Kalinina[1,†], Sergii Shiyan[1,†], Victor Gozhyi[1,†] and Ainura Ormanbekova[2,†]

[1] *Petro Mohyla Black Sea National University, Desantnykiv Street 68, Building 10, 54000 Mykolaiv, Ukraine*
[2] *Almaty Technological University, Tole bi Street 100, 050012 Almaty, Republic of Kazakhstan*

Abstract
The article considers a forecasting information system based on machine learning methods for fuel monitoring. The system solves the following monitoring tasks: data analysis and evaluation, model building and forecasting values or decision-making. The system is developed based on information technologies for fuel monitoring. The information system consists of the following subsystems: information collection and storage subsystem, data preparation subsystem, data analysis and pre-processing subsystem, modeling subsystem and forecasting subsystem. An important place in the modeling and forecasting subsystems is occupied by modules for assessing the quality of models and forecast values based on quality metrics. In the forecasting subsystem, in particular, the forecasting module based on basic alternative models has a forecast value combination module, which implements seven different methods for combining forecast values. In most cases, combination helps to improve the quality of forecasts. The experimental part of the study considers the problem of predicting the volumes of possible filling of storage systems with fuel based on a report on regular data collection on the level and amount of fuel in the tanks on the ship. The following machine learning methods were used for forecasting: exponential smoothing, regression neural network models and Bayesian structural time series models. The quality assessment of the obtained forecast values was carried out using the following quality metrics: MAE, MSE, RMSE. The information system makes it possible to obtain high-quality forecasts of the amount of fuel for tanks of various types, as well as generalized indicators.

Keywords
Forecasting information system, machine learning methods, fuel monitoring.

1. Introduction

The development of modern technologies for managing complex technical objects and systems has given rise to specialized observation methods and analysis methods that constantly collect, process and evaluate information and data about the state of the object and the system as a whole. This process is called monitoring.

Monitoring is a system of constant assessment and forecasting of changes in the state of any technical, natural, social and other objects in other industries based on constant observations. Within the framework of the observation system, control over the object, assessment of the state of the object and management of the object depending on the influence of certain factors take place. Monitoring is formally defined as a systematic process of collecting and analyzing information about a certain object, phenomenon or process in order to track changes, control and make informed decisions based on the information collected.

Monitoring can occur in real time or periodically, allowing to assess the dynamics of changes and certain trends. The main function of monitoring is to control the execution of the process, measure performance indicators and the reaction of the process to the changes made. Monitoring of technical objects and systems is carried out by monitoring systems.

AIT&AIS'2025: International Scientific Workshop on Applied Information Technologies and Artificial Intelligence Systems, December 18–19 2025, Chernivtsi, Ukraine
[1*] Corresponding author.
[†] These authors contributed equally.
alex.gozhyj@gmail.com (A. Gozhyj); irina.kalinina1612@gmail.com (I. Kalinina); shyiansi@gmail.com (S. Shiyan); gozhyi.v@gmail.com (V. Gozhyi); ainura.alibek@gmail.com (A. Ormanbekova)
0000-0002-3517-580X (A. Gozhyj); 0000-0001-8359-2045 (I. Kalinina); 0000-0001-9255-9511 (S. Shiyan); 0000–0002–5341–0973 (V. Gozhyi); 0000-0001-8663-006X (A. Ormanbekova)

Monitoring systems are complex observation and analysis systems that constantly collect, process and evaluate information and data on the state and parameters of an object or system (for example, the state of equipment, the state of the environment, etc.) to control and predict parameters, detect deviations from the norm and make timely decisions.

The main functions of monitoring systems:

1. Observation – the continuous collection of data about the performance of a system or facility.
2. Information collection and processing – automated collection and initial processing of information.
3. Data analysis – evaluation of collected data to identify trends, anomalies, or deviations from specified parameters.
4. Storage – accumulation of information for further analysis and reporting.
5. Forecasting – prediction of possible future states based on data analysis.
6. Response: – providing information for timely decision-making or automatic response to problems.

Fuel system monitoring is the process of tracking and managing fuel levels and consumption using technology to improve operational efficiency, reduce emissions and waste, and prevent fraud and abuse. Accurate and continuous fuel level monitoring is a major challenge for many businesses where effective fuel management is essential. While traditional methods such as manual checks and float-based systems may seem simple, they often fall short of the accuracy, reliability, and efficiency requirements. Decision-making and effective fuel management are complicated by the inability of these methods to provide real-time data, the requirement for constant manual monitoring, and the susceptibility to error.

Modern fuel system monitoring systems are built on the basis of embedded systems that use sensors and a common bus (e.g.: CAN) for data transmission, or by external fuel management systems that combine GPS tracking with fuel level sensors, flow meters, and software [1–3]. In a broader sense, this also includes on-board monitoring systems that check air-fuel ratios and detect malfunctions in fuel storage systems and engine fuel and lubrication systems.

But existing fuel monitoring methods have certain limitations. Many traditional approaches are based on static standards calculated based on averaged conditions, which makes them ineffective for accurately assessing efficiency in real, dynamically changing operating conditions. Even modern telemetry systems that collect large amounts of data mostly perform the function of registration and basic reporting, but do not provide tools for in-depth analysis of the causes of costs. This increases the risk of financial losses due to technical malfunctions, inefficient operation or unauthorized actions.

The purpose of this work is to study machine learning methods in solving the problem of monitoring ship fuel systems. Increasing the efficiency of fuel use by developing a forecasting information system for solving monitoring problems based on machine learning methods.

2. Related works

Analysis of literature sources that highlight the problems of monitoring fuel systems shows that the main trend in creating modern monitoring systems is the use of innovative technological solutions in hardware and machine learning technologies in the data processing process in order to increase the accuracy of monitoring and increase the efficiency of decision-making when managing the fuel system. Fuel monitoring management is a critical technology, where fuel level monitoring systems are key, so the industry is increasingly looking for automated and remote solutions for monitoring and controlling fuel storage systems.

Current monitoring technologies are wireless fuel monitoring systems. In [4], a project of a monitoring system is considered, in which, in order to maximize the efficient use of diesel fuel and simplify timely refueling in remote locations, the Web Socket protocol is used to connect fuel sensors and a mobile application in real time. Thanks to the use of Wi-Fi networks, customers can receive information about consumption, monitor the current level of diesel fuel in different tanks and log in to the mobile application, which has a simple interface. Diesel fuel levels are displayed as a percentage converted to voltage readings, and the system sends notifications when levels fall below set limits. A reliable and seamless diesel fuel level monitoring solution is made possible by integrating an ESP32 microcontroller, Web Socket connectivity, and a capacities fuel level sensor. This cutting-edge technology is based on the basic idea of capacitance. By carefully measuring and electrically converting the change in capacitance to a corresponding level number, very accurate statistics about the amount and condition of fuel in the tank are provided in real time. This project advances the field of remote monitoring systems by offering a robust and scalable method for monitoring diesel fuel levels in a variety of applications.

In [5], the design and implementation of a high-precision monitoring system based on capacitive sensors is investigated. The system is reliable and economical. The project aims to eliminate the shortcomings of traditional methods and set a new standard for efficient and reliable fuel management by carefully selecting sensors, developing sophisticated electronic circuits for accurate measurement, implementing sophisticated software for signal processing and level conversion, and creating an intuitive user interface for real-time data visualisation.

In [6], a project was proposed that used intelligent devices to automate fuel level measurement, for more precise control and to guarantee accurate gasoline supply at gas stations. The system uses GSM modules and intelligent fuel sensors. The system has an alarm system that is triggered when the gasoline level changes abnormally, possibly indicating theft or anomalies in the fuel consumption network. This concept can be used for various types of vehicles.

The work proposes a system [7] to create a fuel level monitoring system that combines the configuration of the Aplicom 12 GSM module with the developed sensor. This allows you to transmit control signals from a mobile device for remote fuel monitoring. The work corresponds to previously published studies of control systems based on mobile phones, integration of GSM modules and fuel level monitoring. The work [8] describes a project that proposes to use IoT technology to solve the problem of fuel monitoring. The hardware features of the project implementation are presented in detail. The project is implemented on the Arduino platform.

The issue of fuel economy for maritime transport is also relevant. The maritime sector also depends on oil prices, like all other industrial sectors. Fuel costs for the maritime industry are the most important expense item. Since the dominant type of fuel in maritime transport is hydrocarbon fuels.

Effective monitoring of fuel use allows you to reduce total emissions of harmful gases into the atmosphere by at least 20%. Therefore, research aimed at reducing emissions from ship operations through the implementation of innovative steps, as well as high fuel prices, is particularly important. Ship fuel consumption is monitored by daily midday reports during the voyage, as well as by companies that perform this service on behalf of shipping agencies. for these reasons [9], [10]. Therefore, the maritime industry focuses on fuel efficiency through methods such as waste heat recovery, loading optimization, maintenance and efficient hull design [11–13]. In addition to all these methods, fuel consumption prediction is used as part of monitoring activities, which is also important for optimizing ship operating conditions.

Estimates of fuel consumption on ships are difficult to process due to the varying operational and environmental conditions, as well as the operation of power and propulsion systems [14–17]. Over the past decade, various fuel economy methods have been proposed to predict the energy efficiency of ships [18–21]. One of these methods examined actual data from reports related to fuel consumption and attempted to predict consumption [22, 23]. Fuel consumption was also estimated based on weather forecasts for the ships' sailing route [24–26] using Automatic Identification System (AIS) [27–29]. Although there are many studies on this topic, usually internal and external factors such as environmental conditions, wind, waves, currents, main engine speed, ship speed, etc. are neglected.

The use of machine learning methods allows for significant improvements in efficiency[30–33]. For example, the multiple linear regression method can be used to find the relationship between several variables mentioned above. This method has proven its success through its use in various forecasting applications. For example, multiple linear regression can be used to find the relationship between variables and, especially, to estimate energy consumption [34–36].

In these works, actual voyage data obtained from a ship were investigated and internal and external factors affecting the ship's fuel consumption were analyzed. To understand the influence of these elements on the ship's fuel consumption during the voyage, a midday report was used. Then, the data was divided into two parts as training and test data. In the next step, the data was calculated using the multiple linear regression method. Based on these calculation data, an estimation and forecasting method for efficient fuel use was developed.

3. Models and methods

3.1. Main aspects of building information technology to solve monitoring tasks

When building information technology to solve fuel monitoring tasks in storage systems, the main aspects of information processing were highlighted. The main procedures for processing and managing information are based on systems analysis methods [37–39]. A systems approach to applied monitoring tasks is used in the process of determining goals at each stage of information processing. The following aspects of building information technology to solve monitoring tasks are highlighted:

1. Observation, measurement and storage of fuel material (FM) parameters for monitoring tasks.
2. Analysis and assessment of data quality for monitoring tasks.
3. Formation of data sets.
4. Preparation and pre-processing of data for monitoring information systems.
5. Definition and construction of model structures for solving forecasting tasks in monitoring systems.
6. Analysis and construction of mathematical models for machine learning procedures when solving problems of predicting FM parameters.
7. Assessment of the quality of model and forecast solutions.
8. Improving the quality and efficiency of forecast solutions through modern approaches.
9. Multi-criteria evaluation when building forecast models and final forecasts at the stage of their creation and in the process of implementation.
10. Creation of an information and analytical system for solving the problem of FM monitoring.

Fig. 1 presents the developed structure of information processing when solving monitoring tasks. As can be seen from the figure, information for further analysis and processing is tracked, accumulated and stored by analysts and operators.

To process information, probabilistic-statistical methods of data analysis and pre-processing are used. This stage is carried out using the appropriate group of methods. At the next stage, predictive models are built using modeling methods. Next, the structure of predictive solutions is determined and the corresponding forecasts are built. This is carried out using forecasting methods. Thus, to solve monitoring tasks, it is necessary to use the following groups of methods [39–41]:

1. Methods and technologies for monitoring and measuring FM parameters.
2. Methods for preparing, analyzing and pre-processing data for FM monitoring systems.
3. Modeling methods for FM monitoring systems.
4. Methods for forecasting and decision-making for FM monitoring systems.
5. Methods for assessing the quality of monitoring.

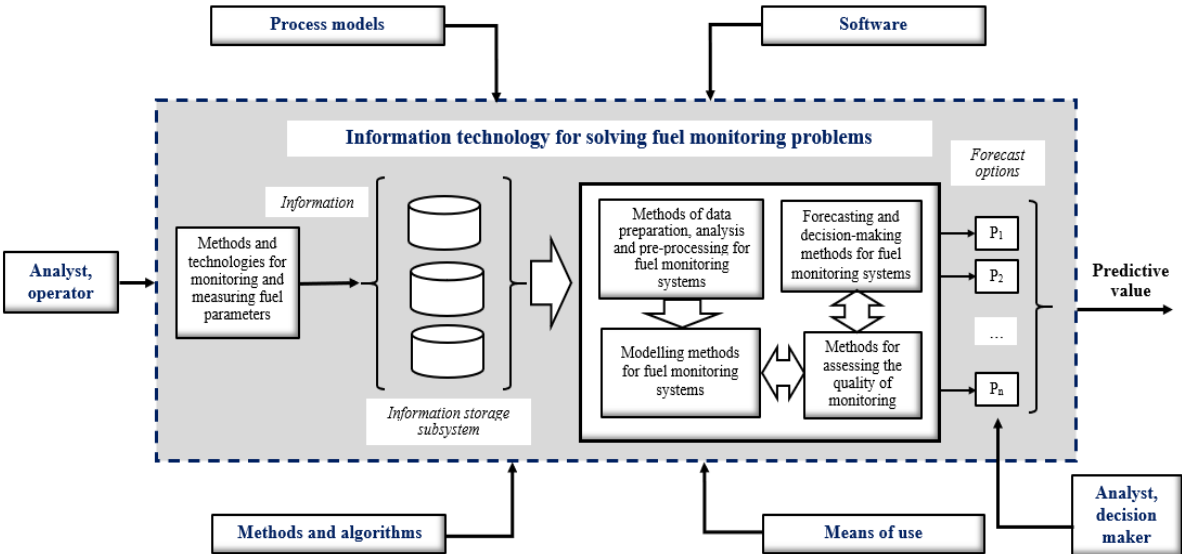

Figure 1: Structure of information processing when solving the monitoring task.

3.2. Information technology for monitoring fuel in storage systems

Information technology for monitoring FM is built on the basis of a system approach and combines groups of methods and methodological approaches, which are grouped by functional purpose [39], [42]. The structural model of information technology for monitoring FM is presented in Fig. 2.

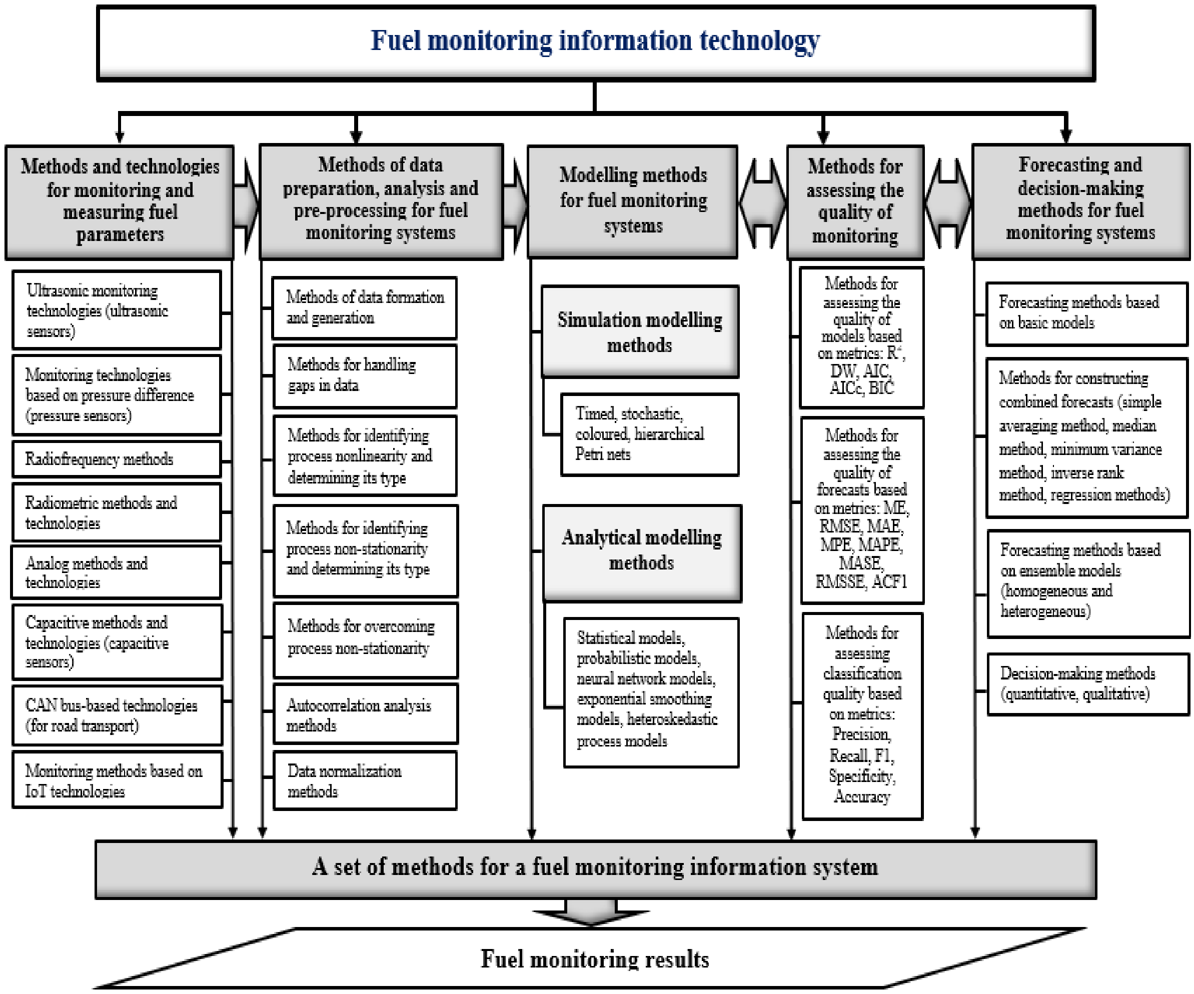

Figure 2: Structural model of information technology for monitoring fuel and oil products.

Information technology combines the following groups of methods: methods and technologies for monitoring and measuring FM parameters, methods for preparing, analyzing and pre-processing data for FM monitoring systems, modeling methods for FM monitoring systems, methods for assessing the quality of monitoring, methods for forecasting and decision-making for FM monitoring systems.

Methods and technologies for monitoring and measuring FM parameters combine ultrasonic monitoring technologies (ultrasonic sensors); monitoring technologies based on pressure differences (pressure sensors); radio frequency methods; radiometric methods and technologies; analogue methods and technologies; capacitive methods and technologies (capacitive sensors); technologies based on CAN buses (for road transport); monitoring methods based on IoT technologies.

Methods for preparing, analysing and pre-processing data combine methods for forming and generating data; methods for processing gaps in data; methods for identifying process nonlinearity and determining its type; methods for identifying process non-stationarity and determining its type; methods for overcoming process non-stationarity; methods for analysing autocorrelation; methods for normalising data.

Modeling methods for FM monitoring systems combine simulation modeling methods (temporal, stochastic, color, hierarchical Petri nets) and analytical modeling methods (statistical models, probabilistic models, neural network models, exponential smoothing models, heteroscedastic process models).

Monitoring quality assessment methods combine methods for assessing the quality of models based on appropriate metrics, methods for assessing the quality of forecasts based on quality metrics.

Forecasting and decision-making methods for FM monitoring systems combine forecasting methods based on basic models; methods for constructing combined forecasts (simple averaging method, median method, minimum variance method, inverse rank method, regression methods); forecasting methods based on ensemble models (homogeneous and heterogeneous); decision-making methods (quantitative, qualitative).

The result of using methods and methodological approaches of information technology for monitoring FM is a set of methods for the information system for monitoring FM.

3.3. Development of an information and analytical system for fuel monitoring

To solve the problems of FM monitoring in storage systems based on the information technology presented in Fig. 2, the structure of the FM monitoring information and analytical system (Fig. 3) has been developed. The structure of the monitoring system is presented at the subsystem level. Information on monitoring FM parameters is stored in the information collection and storage subsystem. The monitoring process begins with this subsystem, which combines the functions of observation, analysis, forecasting, evaluation and development of recommendations for decision-making. The data preparation subsystem is formed from two modules: a module for forming a data set from stored information and a module for forming and generating time series. The prepared data set is transferred to the data analysis and pre-processing subsystem. This subsystem includes four modules: a module for detecting and processing data gaps, an autocorrelation analysis module, a module for identifying nonlinearity and its types, and a module for identifying non-stationarity and its types. The prepared data are then used in the modelling subsystem [39, 40, 42].

The modeling process begins with the procedure of dividing the data set into two samples: training and test. These samples are used in the modules for developing basic forecast models and in the module for assessing the quality of these models. The best forecast models are used by the forecasting subsystem. First, forecasts are formed based on the basic models. Then, using the module for assessing the quality of forecast values, the forecasts are evaluated to select the best ones. But the module for combining forecast values allows using seven methods of combining forecast values based on the basic models to obtain improved forecasts.

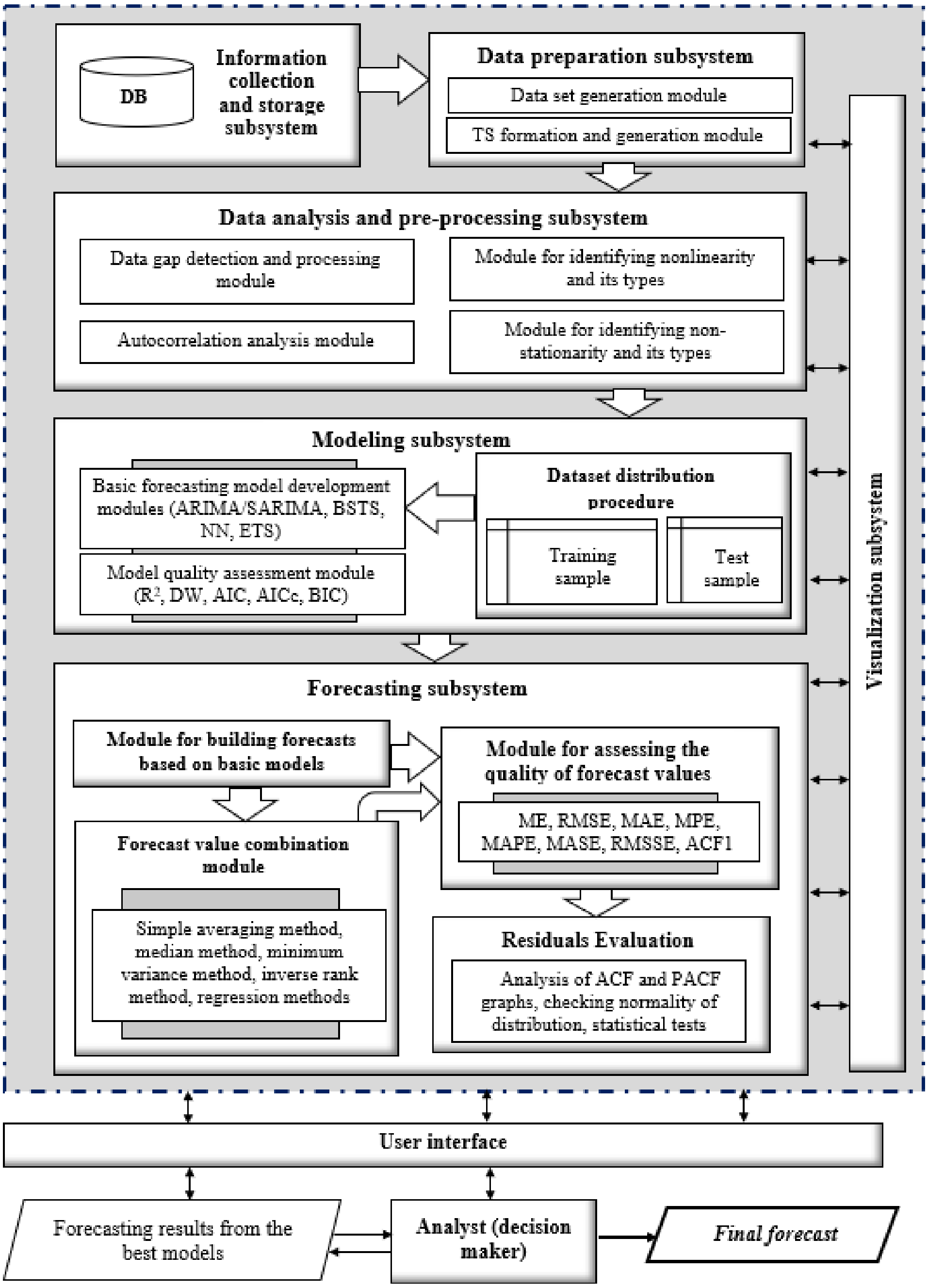

Figure 3: Structure of the information and analytical system for fuel monitoring.

The module for assessing the quality of forecast values allows confirming the improvement of forecasts after combining. Using the residuals assessment module, residuals are diagnosed for the models selected for forecasting. As a result of the operation of the monitoring information system, the analyst is provided with the results of forecasting for the best models to select the final forecast for the selected horizon value.

4. Experimental part

4.1. Data preparation

The first stage of monitoring is the process of data preparation. The source of data in the experimental part of the work was real ship logs, which were formed on the basis of reports on the provision of fuel and lubricants to service companies. An example of one of such reports is presented in Fig. 4.

	DATE:	01.07.2024		TIME:		08:00:00			
Draft:	For'd	6,65	Aft	7,95	Trim	1,30			
MARINE FUEL OIL									
TANK №	Sounding cm	Gross Observed Volume CU.M	DENSITY at 15 C (vac)	Observed TEMP 'C	TABLE 54B VCF	Gross Standard Volume CU.M	TABLE 56 WCF	WEIGHT MT (in air)	85% Tank Capacity in cu.m
	1	2	3	4	5	6=2x5	7	8=3x6x7	(will not be printed)
Heavy fuel oil tank №7 HFOT P/S	540	67,44	0,9682	29	0,9900	66,766	0,99885	64,57	120
Heavy fuel oil tank №9 HFOT Stbd/S	624	93,63	0,9682	29	0,9900	92,694	0,99885	89,64	129
Settling tank №5 HFOT P/S	197	11,07	0,9682	70	0,9605	10,633	0,99885	10,28	11,33
Service tank №6 HFOT P/S	52	3,74	0,9682	80	0,9532	3,565	0,99885	3,45	14,69
TOTAL: IFO		175,88						167,94	275,02

01.07.2024 | 02.07.2024 | 03.07.2024 | 04.07.2024 | 05.07.2024 | 06.07.2024

Figure 4: Fragment of the report on regular data collection on the level and quantity of fuel in tanks as of 07/01/2024.

Reports of regular data collection on the level and quantity of fuel in tanks provide daily information on the filling level of five types of diesel fuel tanks and four types of fuel oil tanks. For each tank, the maximum fuel filling value is provided, which is equal to 85% of the tank capacity. The set of reports corresponds to the observation period from 01.07.2024 to 01.11.2024.

To implement the data preparation subsystem of the monitoring information system, a flowchart was developed, which is shown in Figure 5. According to the structure of the information system, two modules are presented in the flowchart: a module for forming a data set and a module for forming and generating time series. To form the initial data set, data on filling tanks are converted into daily data on the amount of fuel for refueling and summarized in a *.xlsx file. The data is collected in a table in which each observation is uniquely identified by a timestamp (date) and a grouping variable (type). Fig. 6 visualizes a dataset that represents a multidimensional time series.

4.2. Data analysis and pre-processing

According to the algorithm (Fig. 5), the information system allows you to build models and predict the volumes of possible fuel filling for any tank, as well as predict the total indicators for tanks that correspond to one type of fuel (diesel fuel or fuel oil). The work of the subsystem of analysis and pre-processing of data is presented on the example of preparing data on the capacity of a tank with diesel fuel for refueling. As a result of checking, no missing values were found in the dataset diesel_oil_tank_1. The result of the decomposition of the time series corresponding to the amount of diesel fuel for refueling in tank No. 1 is presented in Figure 7.

STL decomposition shows that the general trend exhibits a gradual decline, and the trend component together with random factors forms the general variance of the data. Seasonality varies noticeably in time, which is repeated every week, but the variance of the seasonality values is quite low. Autocorrelation analysis of the data revealed a moderate relationship between the values of the time series at all nine shifts. The magnitude of the relationship is significantly greater for the first 1-3 shifts, then the degree of relationship gradually decreases. The degree of relationship between the time series and its shifted copies was quantitatively analyzed using ACF and PACF plots. ACF values gradually decrease with increasing lag, but the decay occurs slowly. This indicates the presence of a trend in the time series. It can be linear, quadratic or other type of trend.

There are slight weekly fluctuations that indicate the presence of a seasonal component of the series. At the first lag, the PACF has a high value (0,978), and at subsequent lags the values decrease and fluctuate around zero. Since the PACF has a significant value only at the first lag, this confirms that AR(1) can be a good model to describe the process. Statistical tests for the presence of autocorrelation (Durbin-Watson and Broisch-Godfrey tests) confirmed its existence.

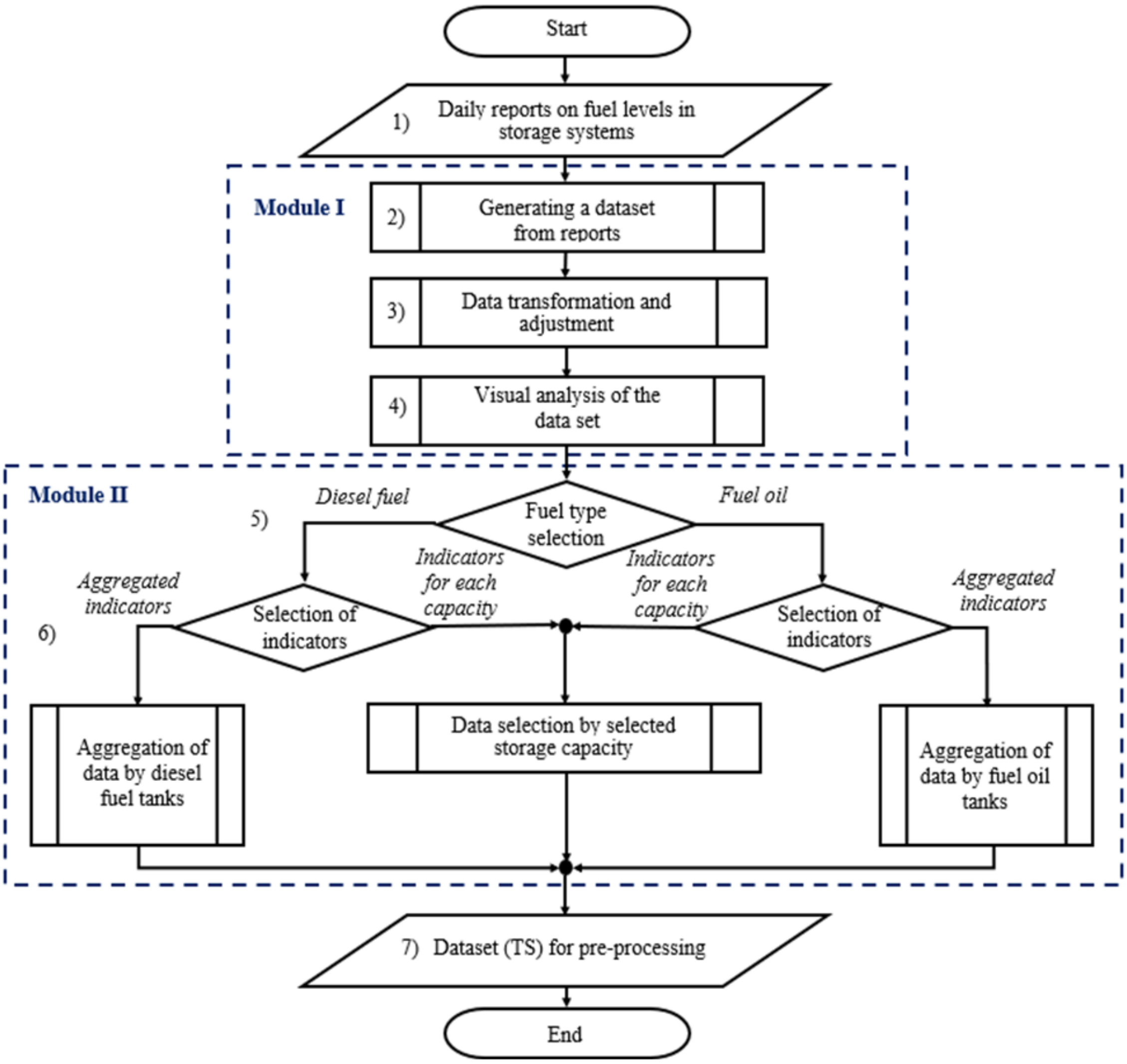

Figure 5: Flowchart of the data preparation subsystem.

Figure 6: Dynamics of the amount of fuel for refuelling in the ninth tanks.

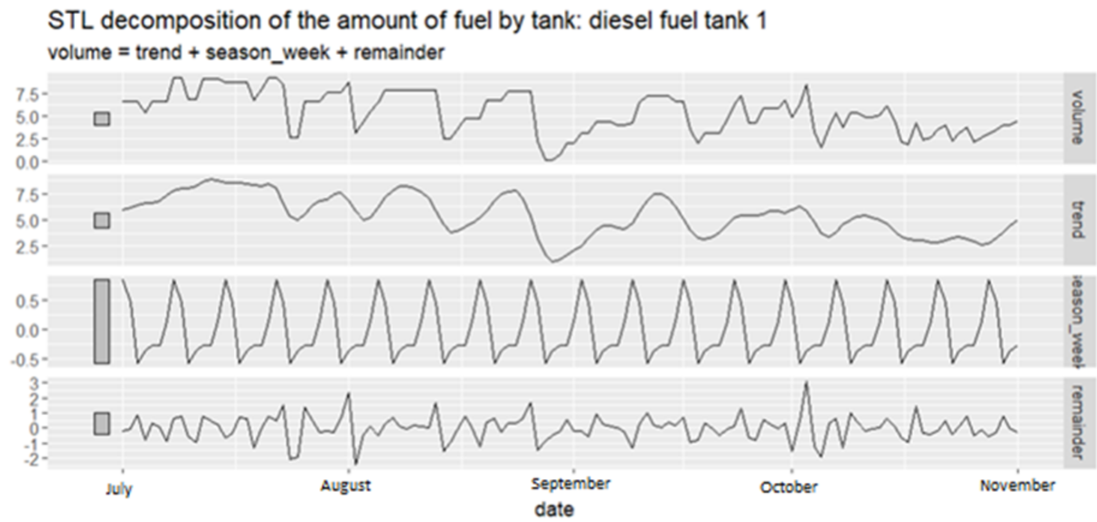

Figure 7: The result of the decomposition of the time series corresponding to the amount of diesel fuel for refuelling in tank No. 1.

In addition to the visual analysis, statistical tests confirmed the presence of nonlinearity in the process under study. The Dickey-Fuller tests, KPSS and Phillips-Perron unit root tests confirmed the presence of non-stationarity.

4.3. Modeling and forecasting

The dataset prepared for modeling was divided in a ratio of 80:20. Thus, 99 observations were selected as a training sample, which is prepared for training models and analyzing their adequacy. And 25 observations were left as a test sample, intended for checking the quality of the basic predictive models. This approach allows you to avoid overtraining when the model shows high results on training data, but works poorly on new, unknown data.

ARIMA/SARIMA models, exponential smoothing models, regression neural network models and Bayesian structural time series models were considered as the main forecasting models. For each model, the structure was selected and the parameters were found at which the models had the best prediction quality indicators on the test dataset. A graphical representation of the results of modeling and forecasting using basic predictive models is presented in Figure 7. For each model, the graphs demonstrate the forecasting of only the test part of the time series. Table 3 presents the values of the forecast quality metrics after training and testing each of the basic predictive models.

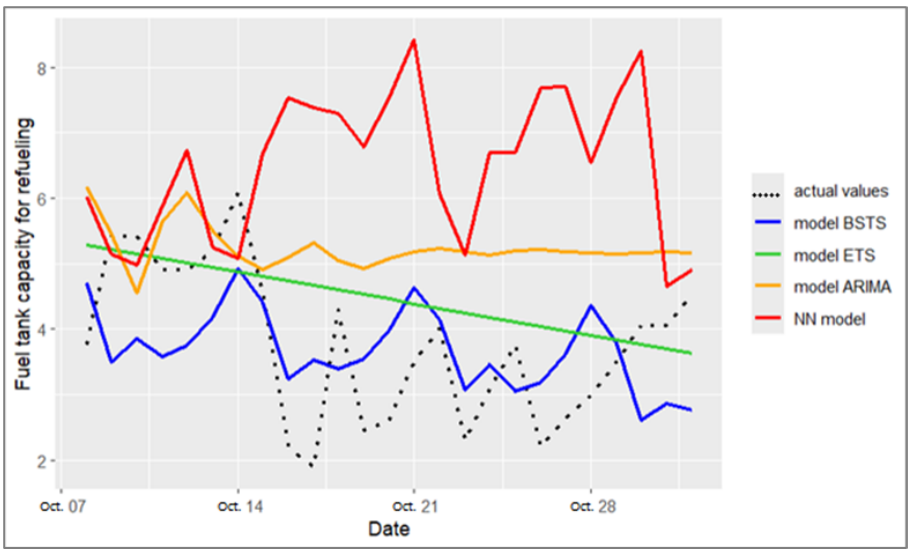

Figure 8: Graphs of actual volumes of possible diesel fuel refuelling for tank No. 1 and forecasts using the best forecasting models.

Table 1
Table of forecast quality assessments for the test sample based on the best models

№	Type of model	MSE	RMSE	MAE
1	BSTS model (Local level + trend + weekly seasonality)	1,3343	1,1551	1,0594
2	Exponential smoothing model (Holt model)	1,5666	1,2517	0,9637
3	ARIMA (0,1,4)(1,0,0)[7]	3,5024	1,8715	1,6171
4	NNAR (25,25,k)[7], Max_it=1500	11,4945	3,3904	2,8733

The best indicators of quality metrics were obtained by BSTS models, this can be seen visually in Fig. 8 and Table 1. But due to the complexity of the process under study, there was a need to use approaches to improve the quality of forecasts. Therefore, 7 methods of combining forecast values were used: the method of simple averaging, the median method, the method of minimum variance, the inverse rank method, the method of constructing a regression model with coefficients selected by the least squares method, the method of constructing a regression model with coefficients selected by the least absolute deviation method and the method of combining several regression models. Table 2 presents the values of the quality metrics of forecasts after combining the forecast values of the basic predictive models.

As a result of comparing the forecast quality metrics from Tables 1 and 2, it can be seen that three of the seven combination methods demonstrate an improvement in forecast quality compared to the results of the BSTS model. These are the results of using the inverse rank method, the method of using a regression model with coefficients selected by the method of least absolute deviation, and the method of combining multiple regression models.

The residual plot and diagnostic testing based on the forecast combination model further confirm the quality of the model because the model is tested for autocorrelation in the residuals using the Box-Ljung test and heteroscedasticity. The normality of the residuals distribution is confirmed using the Shapiro-Wilk test.

Table 2
Estimates of forecast accuracy by combining forecast values for the best baseline models

№	Conventional designation of models for combining forecast values	MSE	RMSE	MAE
1	Simple averaging model	4,4347	2,1059	1,7710
2	Median model	3,3955	1,8427	0,9637
3	Minimum Variance Method Model	2,4205	1,5558	1,6171
4	Inverse rank method model	0,8535	0,9238	0,6858
5	Regression model with coefficients fitted by the least square’s method	3,0357	1,7423	1,4338
6	Regression model with coefficients selected by the least absolute deviation method	0,7938	0,8910	0,7434
7	Model based on a combination of multiple regression models	0,8250	0,9083	0,7858

The described procedures are used to predict the possible filling volumes of other tanks in the ship's fuel storage system. The prediction is carried out for each tank based on the results of its monitoring.

4.4. Obtaining forecasts for aggregated indicators

To increase the efficiency of tank operation in fuel storage systems, an assessment of fuel needs is used based on aggregated indicators, separately by fuel type (diesel fuel and fuel oil). Figures 9,10 shows the dynamics of diesel fuel needs a) and fuel oil b) for four months.

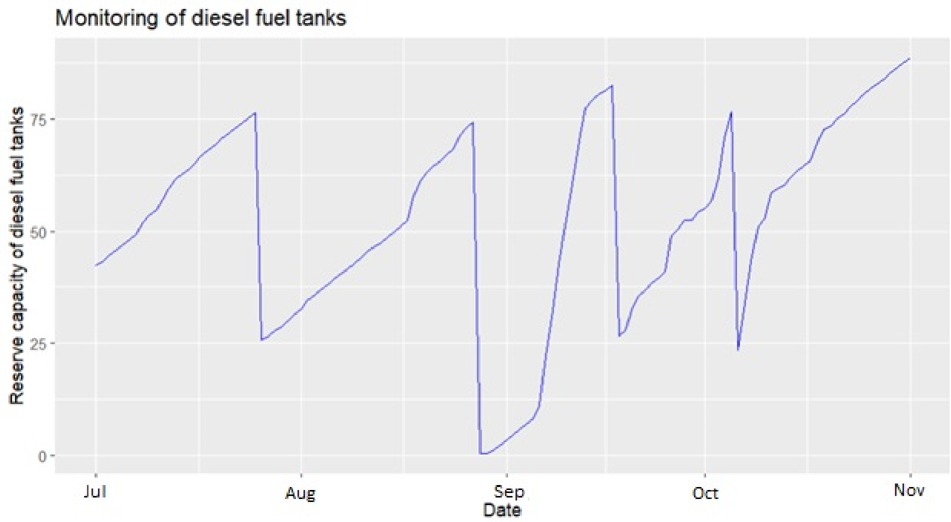

Figure 9: Schedule of possible fueling volumes for all diesel fuel tanks.

Based on the structure of the information system (Fig. 3), an analysis of these aggregated processes was conducted and the appropriate parameters were selected for four basic forecasting models (ARIMA/SARIMA model, exponential smoothing model, regression neural network model and BSTS model). The models were tested on training and test samples.

Figure 11 graphically presents the test sample of data on aggregated indicators for diesel fuel and the best of the basic forecasting models were determined. Table 3 presents the values of quality metrics for the basic forecasting models on the test sample. The best forecasting result is demonstrated by the neural network model. To improve the forecasting results, a combination of forecasts obtained on the basis of basic models was used.

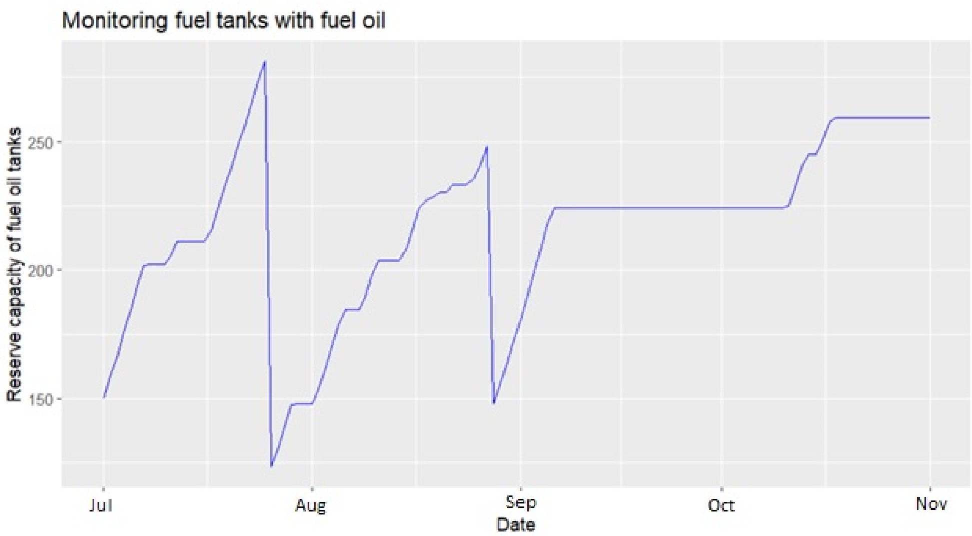

Figure 10: Schedule of possible fuel filling volumes for all fuel oil tanks.

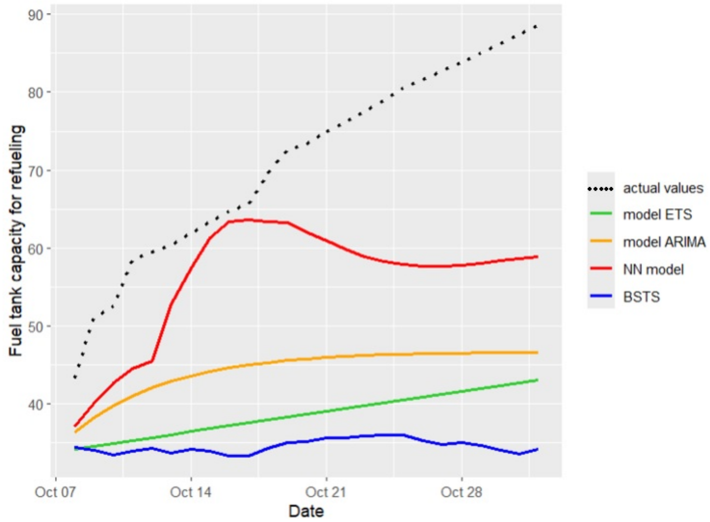

Figure 11: Graphs of actual possible refuelling volumes for all diesel tanks and forecasts using the best forecasting models.

Table 3
Table of forecast quality scores by aggregated indicators for the test sample based on the best models

№	Type of model	MSE	RMSE	MAE
1	BSTS model (Local level)	1485,9090	38,5475	36,6398
2	Exponential smoothing model (Holt model)	1153,0901	33,9572	32,5421
3	ARIMA (0,0,1)	811,6551	28,4896	26,7587
4	NNAR (25,25,k), Max_it=1500	313,1018	17,6947	15,1968

Table 4 shows the quality indicators of forecasts for seven combination methods. The best results were obtained when using the inverse rank method, the method based on a regression model with coefficients selected by the method of the least absolute deviation and the method of combining regression models.

Table 4
Table of forecast quality scores by aggregated indicators for the test sample based on the best models

№	Conventional designation of models for combining forecast values	MSE	RMSE	MAE
1	Simple averaging model	867,3234	29,4504	27,7843
2	Median model	973,4324	31,1999	29,6473
3	Minimum Variance Method Model	591,2766	24,3161	22,5245
4	Inverse rank method model	0,7899	0,8837	0,5073
5	Regression model with coefficients fitted by the least square's method	628,8077	25,0760	23,3200
6	Regression model with coefficients selected by the least absolute deviation method	0,6268	0,7917	0,6194
7	Model based on a combination of multiple regression models	1,9797	1, 4070	1,1513

Figure 12 graphically presents a test sample of data on aggregated indicators for fuel oil and identifies the best of the basic forecasting models. The values of quality metrics for the basic models for predicting the volumes of possible fuel oil filling of ship tanks are presented in Table 5 for the test sample. The best forecasting result is demonstrated by the exponential smoothing model. Table 6 presents the quality metrics of forecasts after using seven methods of combining forecast values. In three out of seven cases, an improvement in forecast quality indicators was obtained.

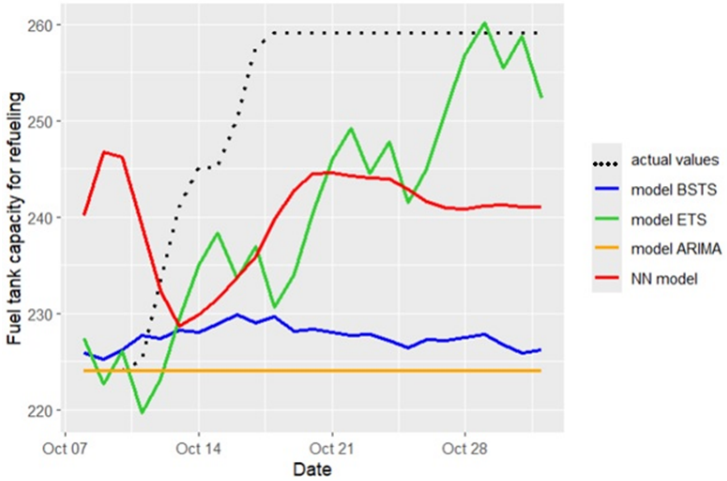

Figure 12: Graphs of actual possible refuelling volumes for all fuel oil tanks and forecasts using the best forecasting models.

Table 5
Table of forecast quality scores by aggregated indicators for the test sample based on the best models

№	Type of model	MSE	RMSE	MAE
1	BSTS model (Local level + trend)	683,4274	26,1424	23,3809
2	Exponential smoothing model (Holt model)	121,5666	12,9411	10,5336
3	ARIMA (1,1,0)(1,0,0)[7]	2472,6278	206,0889	205,8184
4	NNAR (50,50,k)[7], Max_it=5000	283,0057	16,8228	16,3297

Table 6
Forecast accuracy estimates by combining forecast values by aggregated indicators for the best baseline models

№	Conventional designation of models for combining forecast values	MSE	RMSE	MAE
1	Simple averaging model	26012,9583	161,2853	160,8564
2	Median model	43675,2339	208,9862	208,7071
3	Minimum Variance Method Model	372,93047	19,3114	18,8372
4	Inverse rank method model	14,3240	3,7847	2,0151
5	Regression model with coefficients fitted by the least square's method	13159,0772	114,7130	114,0835
6	Regression model with coefficients selected by the least absolute deviation method	13,2489	3,6399	2,3423
7	Model based on a combination of multiple regression models	15,9962	3,9995	2,7117

As a result of the operation of the information system, it is possible to qualitatively determine the volumes of possible filling of tanks in the fuel storage system on board the vessel.

5. Conclusions

The article considers forecasting based on machine learning methods for shipboard fuel monitoring systems. Monitoring tasks were solved on the basis of information technology monitoring in the information system. The information system consists of the following subsystems: information collection and storage subsystem, data preparation subsystem, data analysis and preprocessing subsystem, modelling subsystem and forecasting subsystem. The following monitoring tasks were solved with the help of the system: data analysis and evaluation, model building and forecasting values for decision-making. In the forecasting subsystem of the information system, a module for building forecasts based on basic alternative models was implemented. The following models were used as basic ones: ARIMA/SARIMA models, exponential smoothing models, regression neural network models and Bayesian structural time series models. The quality assessment of the obtained forecast values was carried out using the following quality metrics: MAE, MSE, RMSE.

In the experimental part, the task of predicting the volume of possible fuel filling of tanks in the ship's storage systems was considered. The data source was reports on the level and amount of fuel in the tanks on the ship. The developed information system allowed obtaining high-quality forecasts for each of the tanks on the ship and forecast values for aggregated indicators based on different types of fuel. As a result of the experiment, it was proven that when using three methods of combining forecasts (the inverse rank method, the method of constructing a regression model with coefficients selected by the method of least absolute deviation and the method of combining several regression models), the quality of forecasts was improved. The monitoring information system allows obtaining high-quality forecasts of the amount of fuel on board for tanks of various types.

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly in order to: Grammar and spelling check. After using these tools/services, the authors reviewed and edited the content as needed and takes full responsibility for the publication’s content.

References

[1] P. Kalsi, H. Singh, Sensors based automotive vehicle for the detection of fuel level, Int. J. Innov. Technol. Explor. Eng. 8 (11) (2019) 1001–1004. doi:10.35940/ijitee.I7624.0981119.

[2] N. R. Rai, P. D. Rajguru, A. N. Jagtap, A. S. Shinde, S. D. Martande, A review on: digital fuel level and battery life indicator, IERJ 3 (4) (2020) 6174-6177. doi:10.13140/RG.2.2.30670.00325.

[3] J. Hüffmeier, M. Johanson, State-of-the-art methods to improve energy efficiency of ships, J. Mar. Sci. Eng. 9 (4) (2021). doi:10.3390/jmse9040447.

[4] Sh. Sondkar, O. Patil, P. Bhosale, P. Nyahalde, Diesel level monitoring system, IJRASET 12 (5) (2024) 4301–4306. doi:10.22214/ijraset.2024.62590.

[5] M. F. Pradana, Enhancing ship fuel efficiency in the archipelago area through a combination of speed optimization and flettner rotor implementation, D.Sc. Thesis, Universität Duisburg-Essen, Essen, Germany, 2024, doi:10.17185/duepublico/82503.

[6] R. Krishnasamy, R. Aathi, B. Jayapalan, K. Karthikeyen, M. Nowfal, Automatic fuel monitoring system, Int. J. Recent Technol. Eng. 8 (8) (2019) 348–352. doi:10.35940/ijrte.D1078.1284S219.

[7] G. D. Obikoya, Design, construction, and implementation of a remote fuel-level monitoring system, J. Wireless. Com. Network. 2014 (76) (2014). doi:10.1186/1687-1499-2014-76.

[8] R. Patel, H. Pungalia, S. Mahajan, Flow meter and Arduino based fuel gauge for automotive vehicles, IOSR-JMCE 13 (5) (2016) 85-92. doi:10.9790/1684-1305078592.

[9] G. L. Reynolds, The reduction of GHG emissions from shipping – A key challenge for the industry, in: Proceedings of the World Maritime Technology Conference, WMTC ’2009, WMRC, Southampton, UK, 2009, pp. 21–24.
[10] A.-R. Kim, Y-J. Seo, The reduction of SOx emissions in the shipping industry: The case of Korean companies, Marine Policy 100, (2019) 98-106. doi:10.1016/j.marpol.2018.11.024.
[11] A. A. Safaei, H. Ghassemi, M. Ghiasi, VLCC’s fuel consumption prediction modeling based on noon report and automatic identification system, Cogent Engineering 6 (2019). doi:10.1080/23311916.2019.1595292.
[12] N. Bialystocki, D. Konovessis, On the estimation of ship's fuel consumption and speed curve: A statistical approach, J. Ocean Eng. Sci. 1 (2) (2016) 157-166. doi:10.1016/j.joes.2016.02.001.
[13] K.-K. Kee, B.-Y. L. Simon, Cloud-based IoT solution for predictive modeling of ship fuel consumption, in: Proceedings of the 2019 8th International Conference on Software and Computer Applications, ICSCA '19, Association for Computing Machinery, New York, NY, 2019, pp. 44–49. doi:10.1145/3316615.3316710.
[14] J. Yuan, V. Nian, Ship energy consumption prediction with Gaussian process metamodel, Energy Procedia 152 (2018) 655-660. doi:10.1016/j.egypro.2018.09.226.
[15] K. Young-Rong, J. Min, P. Jun-Bum, Development of a fuel consumption prediction model based on machine learning using ship in-service data, J. Mar. Sci. Eng. 9 (2) (2021). doi:10.3390/jmse9020137.
[16] Z. Eddaoudi, Z. Aarab, K. Boudmen, A. Elghazi, M. D. Rahmani, A brief review of energy consumption forecasting using machine learning models, Procedia Computer Science 236 (2024) 33–40. doi:10.1016/j.procs.2024.05.001.
[17] R. H. Zwart, J. Bogaard, A. A. Kana, A Grey-box model approach using noon report data for trim optimization, International Shipbuilding Progress 70 (1) (2023) 41–63. doi:10.3233/ISP-2200.
[18] M. I. Rodríguez-García, J. González-Enrique, J. A. Moscoso-López, J. J. Ruiz-Aguilar, J. C. Rodríguez-López, I. J. Turias, Comparison of maritime transport influence of SO2 levels in Algeciras and Alcornocales Park (Spain), Transportation Research Procedia 58 (2021) 591–598. doi:10.1016/j.trpro.2021.11.078.
[19] A. Kılıç, M. Yolcu, F. Kılıç, L. Bilgili, Assessment of ship emissions through cold ironing method for Iskenderun Port of Turkey, Environmental Research and Technology 3 (4) (2020) 193–201. https://doi.org/10.35208/ert.794595.
[20] E. Czermanski, G. T. Cirella, A. Oniszczuk-Jastrzabek, B. Pawłowska, T. Notteboom, An energy consumption approach to estimate air emission reductions in container shipping, Energies 14 (2) (2021). doi:10.3390/en14020278.
[21] X. Wu, Y. Zhang, Y. Chen, A dynamic programming model for joint optimization of electric drayage truck operations and charging stations planning at ports, IEEE Transactions on Intelligent Transportation Systems 24 (11) (2023) 11710–11719. doi:10.1109/TITS.2023.3285668.
[22] J. J. Corbett, J. J. Winebrake, E. H. Green, P. Kasibhatla, V. Eyring, A. Lauer, Mortality from ship emissions: A global assessment, Environ. Sci. Technol. 41 (24) (2008) 8512–8518. doi:10.1021/es071686z.
[23] L. Schrooten, I. De Vlieger, L. I. Panis, C. Chiffi, E. Pastori, Emissions of maritime transport: A European reference system, Sci. Total Environ. 408 (2) (2009) 318–323. doi:10.1016/j.scitotenv.2009.07.037.
[24] R. Zaccone, E, Ottaviani, M. Altosole, Ship voyage optimization for safe and energy-efficient navigation: A dynamic programming approach, Ocean Engineering 153 (1) (2018) 215–224. doi:10.1016/j.oceaneng.2018.01.100.
[25] X. Lang, H. Wang, W. Mao, N. Osawa, Impact of ship operations aided by voyage optimization on a ship’s fatigue assessment, J. Mar. Sci. Technol. 26,(2021) 750–771. doi:10.1007/s00773-020-00769-8.

[26] H. Wang, X. Lang, W. Mao, Voyage optimization combining genetic algorithm and dynamic programming for fuel/emissions reduction, Transp. Res. D: Transport and Environment 90 (2021). doi:10.1016/j.trd.2020.102670.
[27] H. M. Perez, R. Chang, R. Billings, T. L. Kosub, Automatic identification systems (AIS) data use in marine vessel emission estimation, in: Proceedings of the 18th Annual International Emission Inventory Conference, IEIC ’2009, U.S. Environmental Protection Agency, Washington, DC, 2009, pp. 1–17.
[28] Y. Tian, L. Ren, H. Wang, T. Li, Y. Yuan, Y. Zhang, Impact of AIS data thinning on ship air pollutant emissions inventories, Atmosphere 13 (7) (2022). doi:10.3390/atmos13071135.
[29] K. Sheng-Long, C. Wu-Hsun, C. Chao-Wei, AIS-based scenario simulation for the control and improvement of ship emissions in Ports, J. Mar. Sci. Eng. 10 (2) (2022). doi:10.3390/jmse10020129.
[30] M.-R. Mehregan, A. Kazemi, H. Shakouri, A fuzzy linear programming model for allocation of oil and gas resources in iran with the aim of reducing the greenhouse gases, Environmental Progress & Sustainable Energy 201332 (3) (2013) 854–860. doi:10.1002/ep.11692.
[31] H. Shi, K. Miao, X. Ren, Short-term load forecasting based on CNN-BiLSTM with Bayesian optimization and attention mechanism, Concurr. Comput. Pract. Exp. 35 (17) (2023). doi:10.1002/cpe.6676.
[32] Y. Wang, P. Guo, N. Ma, G. Liu, Robust wavelet transform neural-network-based short-term load forecasting for power distribution networks, Sustainability 15 (1) (2023). doi:10.3390/su15010296.
[33] P. Prousaloglou, M.-Ch. Kyriakopoulou-Roussou, P. J. Stavroulakis, V. Tsioumas, S. Papadimitriou, Artificial intelligence in the service of sustainable shipping, J. Ocean Eng. Mar. Energy 11 (2025) 621–653. doi:10.1007/s40722-025-00390-0.
[34] M. R. Mehregan, M. Taghizadeh-Yazdi, G. H. Shakouri, M. B. Menhaj, A. Kazemi, Design of a multi-level fuzzy linear regression model for forecasting transport energy demand: A case study of Iran, in: Proceedings of the International Conference on Computers & Industrial Engineering, ICCIE '2009, IEEE, New York, NY, 2009, pp. 1757–1762. doi:10.1109/ICCIE.2009.5223766.
[35] R. Torkzadeh, A. Mirzaei, M. M. Mirjalili, A. S. Anaraki, M. R. Sehhati, F. Behdad, Medium term load forecasting in distribution systems based on multi linear regression & principal component analysis: A novel approach, in: Proceedings of the 2014 19th Conference on Electrical Power Distribution Networks, EPDC ’2014, IEEE, New York, NY, 2014, pp. 66–70. doi:10.1109/EPDC.2014.6867500.
[36] X. Sun, Z. Ouyang, D. Yue, Short-term load forecasting based on multivariate linear regression, In: Proceedings of the 2017 IEEE Conference on Energy Internet and Energy System Integration, EI2 ’2017, IEEE, New York, NY, 2017, pp. 26–28. doi:10.1109/EI2.2017.8245401.
[37] A. Nielsen, Practical Time Series Analysis: Prediction with Statistics and Machine Learning, O’Reilly Media, Sebastopol, CA, 2019.
[38] J. D. Kelleher, B. Mac Namee, A. D’Arcy, Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies, 2nd ed., MIT Press, Cambridge, MA, 2020.
[39] P. Bidyuk, I. Kalinina, A. Gozhyj, I. Pikh, V. Chorna, V. Gozhyi, A systematic approach to modeling and forecasting based on real data in machine learning tasks, in: M. Zgurovsky, N. Pankratova (Eds.), System Analysis and Data Mining, Springer, Cham, Switzerland, 2025, pp. 71–87, doi:10.1007/978-3-031-97529-5_5.
[40] I. Kalinina, A. Gozhyj, Forecasting electricity demand in Ukraine using machine learning methods, in: Proceedings of the Computational Intelligence Application Workshop, CIAW ’2024, CEUR Workshop Proceedings, Aachen, Germany, 2024, pp. 42–56.
[41] A. Gozhyj, V. Nechakhin, I. Kalinina, Solar power control system based on machine learning methods, in: Proceedings of the 2020 IEEE 15th International Conference on Computer

Sciences and Information Technologies, CSIT '2020, IEEE, New York, NY, 2020, pp. 24–27. doi:10.1109/CSIT49958.2020.9321953.

[42] P. Bidyuk, I. Kalinina, A. Gozhyj, V. Gozhyi, S. Shiyan, An approach to combining forecasts when solving machine learning problems, in: Proceedings of the Modern Machine Learning Technologies Workshop, MoMLeT '2025, CEUR Workshop Proceedings, Aachen, Germany, 2025, pp. 12–24.