Brandt, Peter; Munim, Ziaul Haque; Chaal, Meriam; Kang, Hooi Siang

Maritime accident risk prediction integrating weather data using machine learning

Published in:
Transportation Research, Part D: Transport and Environment

DOI:
10.1016/j.trd.2024.104388

Published: 01/11/2024

Transportation Research Part D 136 (2024) 104388

Contents lists available at ScienceDirect

Transportation Research Part D

journal homepage: www.elsevier.com/locate/trd

Maritime accident risk prediction integrating weather data using machine learning

Peter Brandt [a], Ziaul Haque Munim [a,*], Meriam Chaal [b,c], Hooi-Siang Kang [d]

[a] Faculty of Technology, Natural and Maritime Sciences, University of South-Eastern Norway, Horten, Norway
[b] Department of Mechanical Engineering, Marine and Arctic Technology, Aalto University, Espoo, Finland
[c] Kotka Maritime Research Centre, Kotka, Finland
[d] Marine Technology Centre (MTC), Institute for Vehicle System & Engineering, Faculty of Mechanical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

ARTICLE INFO

Keywords:
Maritime safety
Risk assessment
Automated machine learning
Accident analysis

ABSTRACT

The study explores the capability of various machine learning (ML) models in maritime accident risk prediction. Data from 1981 to 2021 from the Norwegian Maritime Authorities (NMA) was analysed together with the data of 51 different weather-related variables, which were collected from Visual Crossing for each accident recorded in the NMA dataset. The findings reveal an increased predictive ability of ML models when relevant weather data is introduced. The results show that the *Light Gradient Boosted Trees with Early Stopping* perform the best, with a five-fold cross validation accuracy of 70.23% when weather data was included, compared to 64.86% without. Furthermore, the study revealed that the leading weather variables for accident prediction are *wind*, *sea level pressure*, *visibility*, and *moon phase*. The most effective multi-classification ML algorithm can be deployed for improving maritime safety resilience through vulnerability assessment and preparedness.

1. Introduction

The maritime domain poses numerous risks to the safety of crew members, cargo, and the environment (Luo & Shin, 2019). Maritime accidents can result in loss of life, severe environmental damage, and financial loss (Adland et al., 2021). According to the European Maritime Safety Agency (EMSA, 2022), 2854 ships within the EU states were involved in accidents in 2021, of which 14 vessels were lost, 36 were fatalities, and 58 accidents involved marine pollution. In the period from 2014 to 2021, more than half of all maritime accidents took place in internal waters (port areas and others), and 44 % were "en route", meaning in transit from point "A" to point "B". Whenever vessels go to sea, the environment and the operations occupying them represent numerous risks that are difficult to comprehend. Internal and external conditions, where for example structural defects count as internal, and human contributions define the external, are good examples of how to categorise potential risks (Yildiz et al., 2021). Research on maritime

Abbreviations: AI, Artificial Intelligence; AIS, Automatic Identification System; API, Application Programming Interface; AutoML, Automated Machine Learning; COLREG, Convention on the International Regulations for Preventing Collisions at Sea; CV, Cross Validation; GBM, Gradient Boosted Machine; ML, Machine Learning; NMA, Norwegian Maritime Authority.

* Corresponding author.
E-mail addresses: ptmbrandt@gmail.com (P. Brandt), ziaul.h.munim@usn.no (Z.H. Munim), meriam.chaal@aalto.fi (M. Chaal), kanghs@utm.my (H.-S. Kang).

https://doi.org/10.1016/j.trd.2024.104388
Received 30 November 2023; Received in revised form 25 August 2024; Accepted 25 August 2024
Available online 13 September 2024

accidents has over the last half-century shifted from focusing primarily on naval architecture to human errors, and the modern computational power allows for far better data-processing capabilities, and therefore, possibilities to study accidents in different ways. The shift in research focus has led to an increasing awareness of the complexity of maritime accidents, leading to an increased number of studies using multiple databases (Luo & Shin, 2019). Forthcoming research will likely be multi-disciplinary and use multiple data sources simultaneously (Luo & Shin, 2019). The use of weather data in maritime accidents studies is one way of doing this.

The maritime transport sector is complex, in the sense that many different variables influence the vessel. The maritime domain is inherently risky due to changing weather conditions, fatigue factors influencing human decisions, technical failures, etc. Human actions and human behaviour have been the contributing factor to more than 80 % of all investigated accidents (EMSA, 2022). EMSA points towards physical and mental stress being one of the most influential contributing factors leading up to maritime accidents, but the analysis lacks insight into what type of stressors influence humans' abilities to cause accidents. Contributions by the weather can potentially be one of many stressors or factors that affect humans.

Previous studies found specific weather parameters, such as fog, wind and waves, to pose a risk towards vessels (Liu et al., 2021). By understanding the influence of weather on maritime accidents, one can mitigate risk and prevent future accidents. Among various accident categories, within Chinese coastal waters, collision accidents hold the highest share (Liu et al., 2021). The vessels exposed to the largest risk of collision in these waters are small general cargo ships, as 56 % of all accidents involved ships of 100 m in length or less. Less serious accidents often happen in areas with low traffic. In the context of maritime accidents in the UK and EU, the same issue was pointed out, with collisions and contact types always being the most frequent accident types at sea (Ugurlu, 2022).

Since the COVID-19 pandemic, there has been an influx of studies on maritime transport resilience. Resilience simply refers to the ability of a system to return to normal condition after a disruption. In the maritime literature studies have focused on port resilience (Li et al., 2022), transport network resilience (Wang and Yuen, 2022), maritime supply chain resilience (Gu et al., 2023) etc. However, safety resilience in terms of accident risk in maritime transport networks has not received much attention yet. Pre-disruption preparation and effective resource allocation is the first phase in building a resilient maritime system (Gu and Liu, 2024). Thus, implementing predictive analytics in accident risk analysis can contribute to maritime safety resilience.

Table 1
A summary of relevant literature.

No	Author(s)	Data type/source	Methods	Conclusion
1	Ugurlu & Cicek (2022)	513 ship accidents (UK MAIB, IMO GISIS, EMSA).	Fault Tree, Multiple Correspondence Analysis.	Two most important factors in collision accidents are maneuvering and perception errors.
2	Merrick et al. (2022)	Vessel traffic, environmental, and waterway conditions in New Orleans.	Logistic regression, random forest, k-Nearest neighbor, decision tree, Neural networks, Gradient boosted trees, Stochastic gradient boosted trees.	Poor accident prediction accuracy.
3	Adland et al. (2021)	Weather data (EU-Copernicus), AIS-data for 42,000 voyages in the North Pacific.	LASSO, eXtreme Gradient Boosted Trees.	Weather impacts marine risk.
4	Yildiz, Ugurlu, Wang & Loughney (2021)	51 grounding accidents that occurred in passenger vessels.	Human Factors Analysis and Classification System for Passenger Vessels (HFACS-PV)	The HFACS-PV approach is effective in analysing grounding, contact, sinking and collision accidents.
5	Uyanık et al. (2021)	Data from a local weather station to estimate visibility.	Gradient Boosting method, Bayesian Ridge, AdaBoost, Gradient Boosting, Random Forest, Multiple-Linear, Multi-Layer Perceptron.	The Gradient Boosting is the best performing ML algorithm, and accident risk can be reduced by predicting visibility condition.
6	Rawson et al. (2021)	Vessel traffic, weather data (EU-Copernicus), historical casualty data.	Support Vector Machines, XGBoost, Random Forest, Stochastic Gradient Decent.	Shows modest success at accident prediction with high false positives. SGD-SVM showed high recall.
7	Liu et al. (2021)	Maritime accidents from Chinese coastal waters.	Bayesian Network, comparative analysis.	Bad weather conditions are highly associated with catastrophic accidents.
8	Fan et al. (2020)	Maritime accident reports (MAIB, TSB).	Bayesian network, TOPSIS, Multiple correspondence analysis, hierarchical clustering, classification tree.	Information availability, clear order, and safety culture are effective in accident prevention.
9	Kretschmann (2020)	Maritime accident data, Port State Control data.	Uses machine learning algorithms to quantify risk of operating conditions.	ML is useful in estimating a leading risk indicator.
10	Bye & Aalberg (2018)	Accident data (NMA) and AIS.	Correspondence analysis, F-test, multivariate logistic regression.	Flag of convenience, vessel type, length, and poor visibility increase the likelihood of navigation-related accidents.
11	Zhang & Li (2017)	10-year ship accident dataset from IMO (755 weather-related cases), Wave data.	Analysis of wave height during maritime accidents.	Sea states with co-occurrence of wind wave and swell conditions increase the accident risk of vessels.
12	Rezaee et al. (2016a)	Historical fishing activity levels, incident data and extreme weather data.	Negative Binomial Regression, Zero-Inflated Negative Binomial Regression, Fractional Logit Regression etc.	Strong association between weather indicators and fishing activity. Size of vessel matters. Wind speed relevant for vessels smaller than 45ft.
13	Knapp et al. (2011)	Wave height and wind strength data.	Binary regression model.	Wind strength shows a positive correlation with the probability of serious casualties in maritime accidents.

Machine learning (ML) has been used in previous research to analyse maritime accidents (Rawson et al., 2021). ML models can analyse large datasets with complex relationships and identify patterns that may not be apparent through other, more traditional methods (He et al., 2021). By studying large datasets of maritime accidents with ML models, one can expect to find improved predictions and better understand the factors that contribute to maritime accidents. Automated machine learning (AutoML) is developed to make the process of ML even better, by allowing the user to test several ML algorithms simultaneously, and to optimise the parameters to make the most accurate predictions (He et al., 2021). Therefore, this study explores the use of AutoML in predicting maritime accident risk while integrating both historical accident data and weather variables. The approach is directed towards understanding why accidents occur, by utilizing a type of data that is easily available as predictions in the form of weather forecasts. The motivation behind the study was to develop a deeper understanding of why accidents happen at sea, and the aim was to make a meaningful contribution through the development of predictive ML models, which can be applied to help prevent accidents and to effectively allocate resources in order to increase maritime safety resilience. Therefore, the research questions in the study are:

- What are the most effective ML models in predicting maritime accidents?
- Can weather data improve ML predictions of maritime accidents?
- Which weather variables have the greatest impact on maritime accidents?

By addressing these research questions, this study contributes to developing a better understanding of maritime accidents. In doing so, first, it explores a large dataset spanning over 40 years of historical accident data using machine learning. Second, over 100 ML models are trained using a cloud-AI platform to identify the best performing models for multiclass maritime accident prediction. Third, over 50 weather-related variables are used as input features in the ML model training, the most comprehensive to the authors' knowledge. Fourth, the most relevant features influencing five major categories of maritime accidents are revealed. Finally, a detailed insight into the influencing factors is presented through feature effect figures.

The rest of the manuscript is structured as follows: Section 2 reviews existing literature on maritime accidents, weather variables, and machine learning techniques. Section 3 describes the methodology used in the study, including data collection and pre-processing, and AutoML application. Section 4 presents the results based on the best performing ML model. Finally, Section 5 discusses the implications of the findings, and Section 6 summarizes the findings including further potential applications, limitations, and future research directions.

2. Literature review

A systematic literature search was conducted to identify relevant studies, mainly in the Web of Science (WOS) database in early 2023. The initial search with the term "Maritime accident*" reverted 394 articles, which reduced to 97 when searched with the Boolean expression ("maritime accident*" AND "machine learning"). After manually reviewing the results by title and abstract, 25 articles were identified for a detailed review, which led to a total of 17 relevant articles. Reading these studies in detail further led to the discovery of new relevant research articles and other publications. The literature review matrix in Table 1 provides an overview of the most relevant research, arranged by publication year. The aim of the systematic literature search was to identify the core maritime accident studies that utilized ML models. For a comprehensive literature review of machine learning applications in maritime accident analysis, see Rawson & Brito (2023).

Examining the literature using one database might have incorporated uncertainties on the extent to which the dataset is representative of the body of science. Therefore, the number and types of ML algorithms reported in Table 1 might not be comprehensive. However, WOS is recognized as a major database for the scientific community, which in most cases includes journal articles that are covered by the other databases.

Weather-related variables are directly and objectively measurable. In a maritime context, the weather represents variables that will influence most operations daily. Operational conditions can be external and internal — external conditions being either weather conditions or local conditions, and internal conditions being structural defects of the vessel (Uğurlu et al., 2018). Several studies have analysed the effect of different meteorological variables on maritime accidents, including Fan et al. (2020) who concluded that most maritime accidents were caused by poor visibility. The study consisted of a dataset of 161 maritime accidents from Canada and the UK, which were analysed by incorporating Bayesian network (BN) and Technique for Order of Preference by Similarity to Ideal Solutions (TOPSIS) in a Multi-Criteria Decision-Making system (MCDM). These findings are supported by Liu et al. (2021) who, through the use of Bayesian networks and comparative analysis of maritime accidents in Chinese coastal waters found fog to be a critical factor. Accident probability and visibility were found to be negatively correlated; as visibility decreased, the probability of accidents increased. Fog was present in 35 % of the accidents, rain in 23 %, clouds in 19 % and sun in 19 %. A visibility of less than 2 nm occurred in 61 % of the accidents. Liu et al. (2021) concluded that bad weather conditions often lead to catastrophic accidents.

Uyanık et al. (2021) used the Gradient Boosting method, among other ML algorithms, to estimate the visibility in the Strait of Istanbul. They used data from local weather stations to build a predictive visibility model based on wind, humidity, pressure, and time indicators. They found that visibility had a high positive correlation with wind speed and a high negative relationship with humidity. They found that it was possible to predict the visibility in the Strait, and in that sense reduce the risk of maritime accidents. In one study of more than 20.000 vessels in the North Atlantic and Arctic regions, Knapp et al. (2021) found that wind strength and wave height were positively correlated with accident probabilities. As the wind increases, the probability of serious casualty increases. Adland et al. (2021) discovered that winds exceeding 35 kt increased the likelihood of maritime insurance claims in a study of vessels transiting the North Pacific.

Wind speed was found to have a significant effect on the activity level of the Canadian fishing fleet in a study by Rezaee et al. (2016a), especially among the smaller vessels. The critical factors influencing the severity of fishing incidents were wind speed, the presence of ice, temperature, changes in sea level pressure and darkness (Rezaee et al., 2016b). Changes in sea level pressure were also found as one influential factor in predictions of maritime insurance claims, particularly pressure between 1010mb and 1012mb appear to positively impact predictions of an insurance claim to occur (Adland et al., 2021). Moreover, Zhang & Li (2017) found that 52 % of the accidents that occurred had relatively low wave heights, meaning swell provided the dominant wave energy. Most of the cases had less than, or equal to 3 s difference between the swell period and wind wave period. Another important factor generating higher-risk sea states is when the wind direction is less than 60° relative to the swell direction (Zhang & Li, 2017). When this occurs, the wind wave will interact with the swell and generate a more dangerous sea state and threat to shipping.

Building on the findings of the past studies, this study investigates the role of factors related to the external and internal operational conditions in predicting maritime accidents risk. Majority of the past studies that used ML in maritime accident prediction, used only a limited number of models. Several studies used only one model (e.g. Knapp et al., 2011; Kretschmann 2020; Liu et al., 2021), while a maximum seven models (i.e. Uyanık et al., 2021; Merrick et al., 2022) are observed in Table 1. A recent study that implemented AutoML in maritime accident context (Munim et al, 2024) trained 29 ML models. Using AutoML, this study trained a total of 105 ML models, which ensures identifying best-in-class models for real-world applications in improving maritime safety resilience. Although a few of the past studies included weather related variables, comprehensive weather data of 51 variables in total as predictors of maritime accident risks cannot be found. Further, this study explores five major accident categories, while the majority of the past studies either focused on a binary accident variable or only one type of accident. Thus, the multi-classification approach adopted in this study is a novel contribution in maritime safety context.

3. Research methodology

3.1. Data

This study adopted a systematic methodological workflow as depicted in Fig. 1. A dataset of more than 38,000 maritime accidents was collected from the Norwegian Maritime Authorities (NMA) from 1981 until 2021. The open-source dataset contains data on accidents in Norwegian waters, and Norwegian-flagged vessels sailing abroad. Reporting maritime accidents has been mandatory for Norwegian flagged and foreign vessels in Norwegian waters since 2008 (Forskrift om melde- og rapporteringsplikt til sjøs, 2008). Engine failures, fire, collision, hull cracks, marine pollution, flooding, human casualties, vessel sinking, etc shall be reported in a specific form (KS-0197B).[1] This form asks the captain to specify details such as time, position, type of accident, vessel characteristics, speed, course, and a few weather variables. However, the quality of the information reported is often not very detailed regarding weather, and it is left out of the publicly available dataset. Even though reporting was somewhat different in 1981 than in 2021, and safety measures like the International Safety Management Code (ISM) were introduced in the period, each accident in the study was weighted equally.

The dataset was pre-processed in FileMaker Pro, a database software, where accidents with missing relevant data were removed. For accidents where the exact location was unknown, but the general area was included, a new position was assigned based on the average position of other accidents in the same area. An illustrative example could be a *collision with a quay/bridge* without an exact position, but the area is reported to be *Andfjorden*. Other accidents from the area *Andfjorden*, with exact positions, will generate a centre of gravity for the accidents and assign it to our example accident. A total of 10,135 accidents were left for further processing.

After the pre-processing phase, a script was made in Filemaker Pro to collect historical weather data using an API for each accident from Visual Crossing, which is a provider of historical weather data and forecasts. It was possible to retrieve relevant historical data for most of the locations and 57 different weather variables were added to the dataset. The variables include temperature, wind, humidity, precipitation, pressure, cloud cover, sunset/rise, and moon phase. Data for daily and historical minimum, maximum and mean values for wind, temperature, precipitation, cloud cover and humidity, in addition to the data for the hour of the accident were included in the dataset. The objective of the selection of variables was to be able to see the potential of as many different variables as possible. The variables containing historical and daily means relative to the values at the time of the accident were gathered to grasp the potential changes or abnormal weather, and its potential effect on accidents.

The dataset was pre-processed again to remove any accidents with missing relevant weather data. A total of 909 records were deleted because of missing essential weather variables. Personal work-related accidents such as man overboard, poisoning etc. were removed from the dataset. Weather-related variables were taken into the study on the basis that weather is definitely affecting vessels, and personnel onboard with seasickness as well as their mental state. Questions about the weather effects on maritime accidents caused by either constraints to ship motions or human errors started to arise. None of the historical accident records database provides any data on crew fatigue and decision-making under stress at the time of accidents since they were not measured. Most studies focusing on these factors are conducted in a simulated environment, and provide valid insights. Therefore, the complete dataset consisted of 9,226 accidents with numerous weather variables. Table 2 and Fig. 2 present a summary of accident types in the dataset with frequencies and an overview of the accidents' dataset in geographical setting, respectively.

[1] KS-0197B, Sjøfartsdirektoratet. *Rapport om sjøulykke, arbeidsulykke og nestenulykke*.https://www.sdir.no/globalassets/skjemaer/ks-0197-rapport-om-sjoulykke-arbeidsulykke-og-nestenulykke-bm.pdf.

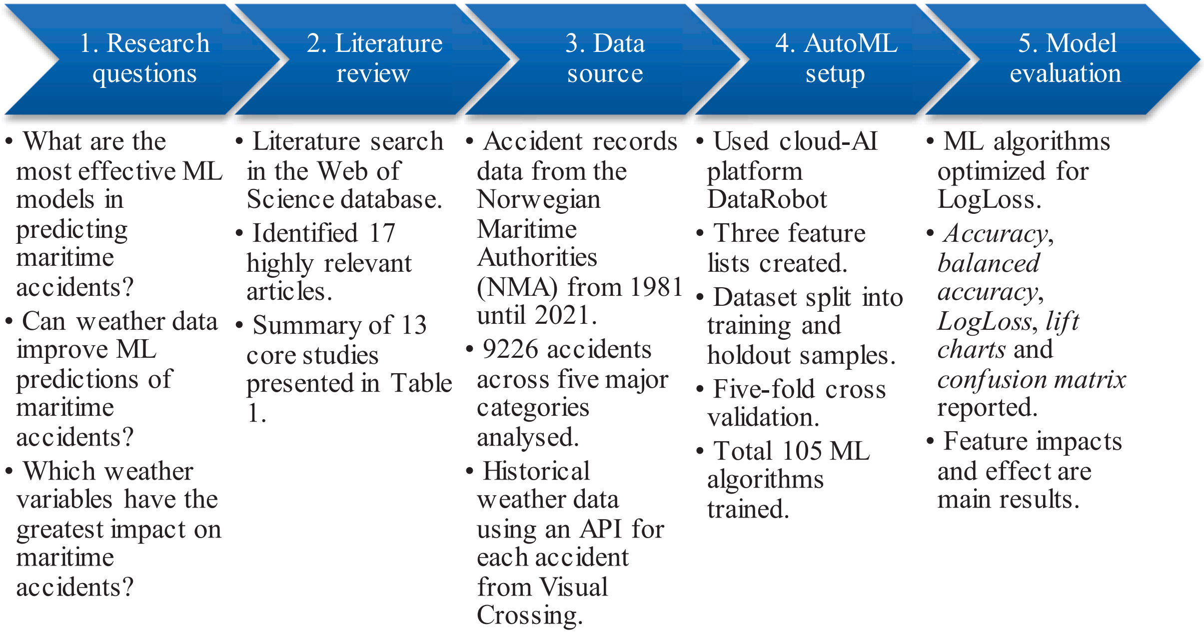
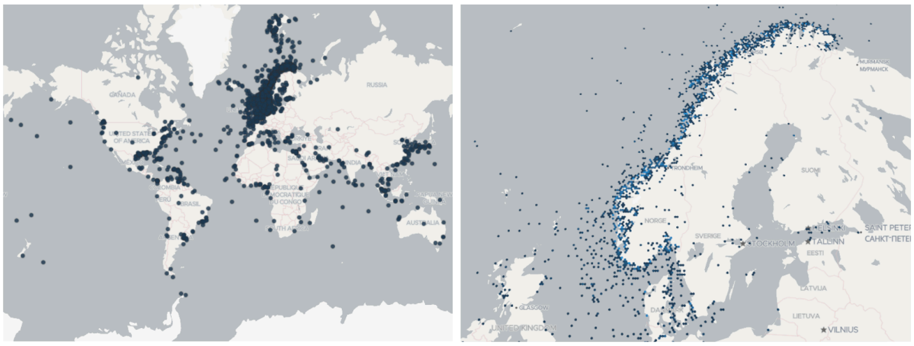


Fig. 1. Methodological workflow.

Table 2
Accident types and their frequencies.

Type of accident	Number of accidents	Percentage
Fire/explosion	946	10.20 %
Grounding	4,915	53.30 %
Heavy weather accidents	146	1.60 %
Collision	1,917	20.80 %
Contact damage (quay, bridge, etc.)	1,302	14.10 %
Total	9,226	100.00 %



(a) Global overview (b) Focused on Norwegian coastal waters

Fig. 2. Accidents in the dataset (Source: Authors, using DataRobot).

3.2. Feature engineered variables

In ML modelling, one of the most challenging tasks is to prepare the dataset. To further enhance the readability of the dataset by ML algorithms, the following variables were coded and added to the dataset:

- *Nationality*: replaced by *reg. code,* a category to distinguish Norwegian and foreign vessels. Norwegian flagged vessels = 1. Other = 2.
- *Day of month:* replaced by *Day of Month Category*. Day 1–10 = 1. Day 11–20 = 2. Day 21–31 = 3

- *Time of accident:* replaced by *Time Category*. 00:00–03:59 = 1. 04:00–07:59 = 2. 08:00–11:59 = 3. 12:00–15:59 = 4. 16:00–19:59 = 5. 20:00–23:59 = 6. This was done to categorise in order of normal bridge watches. For accidents with time 00:00:00, the value is set to *missing*.
- *Sunrise/set*: replaced by the categorical value *Sun UpOrDwn* to distinguish whether an accident occurred while the sun was up or down. 1 for up, 0 for down. For the accident with a time 00:00:00, it's assumed to be a missing time value, and the sun up/down category is therefore set to *missing*.
- *Temp diff C D*: a numerical value, calculated to find the difference between the daily average temperature and the temperature during the accident. The variable was created to be able to see if certain accidents are associated with changes in temperature during the day.
- *Ccover diff C D*: a numerical value, calculated to find the difference between daily average cloud cover and cloud cover during the accident. An increase in cloud cover is associated with an increased probability of precipitation and wind.
- *Press diff C D*: a numerical value to find the difference between daily average sea level pressure and pressure at the time of the accident. A reduction in sea level pressure is associated with deteriorating weather.
- *TempCDewC*: a numerical value to find the difference, or spread, between the temperature and the dew point temperature during the accident. Temperature and dew point temperature are essential in anticipating reduced visibility and can indicate the likelihood of poor visibility in this database. As the spread between temperature and dew point decreases, the relative humidity moves towards 100 %. At 100 % relative humidity, the air is saturated and starts condensing on surfaces such as windows or on particles in the air, and we can experience fog (United States Department of Transportation, 2022, p. 6-4)
- *Vis diff C D*: a numerical value to find the difference in visibility, measured in percentage, from the daily average to the time of the accident. A reduction in visibility is associated with a higher risk of maritime accidents (Fan et al., 2020)
- *Wdir diff C D*: a numerical value to find the difference in the daily average wind direction and the time of the accident. A combination of high wind speed and significant changes in wind direction is associated with a high risk of dangerous waves (Zhang and Li, 2017)
- *WspC OorU D*: a category to distinguish accidents occurring with wind speeds over and under the daily average. The value 1 is given when the wind is greater than average, 0 if under. The variable was created to be able to see if either increasing or decreasing wind speeds from the daily average had any impact on the accidents.
- *Latitude/longitude*: a combined variable, a location, was generated using DataRobot.

3.3. Machine learning modelling

In machine learning, two types of techniques are commonly used: supervised and unsupervised. In unsupervised learning, the models will try to identify hidden patterns or structures in the dataset. The ML algorithm only considers input data, and *clustering* is the most common learning technique. A training partition of the dataset with known input and output data is used to train models in supervised learning. The data must contain input and output data, and the model can, based on the quality and quantity of the training, predict outputs for the given input. Supervised learning can use either *classification* or *regression* techniques to develop predictive models (Murphy, 2012). If the target, or the outcome to be predicted is a number (temperature, price etc.), regression models are used. On the other hand, if it is classes or categories, classification techniques are appropriate. For training models to predict maritime accidents in this study, supervised learning based classification ML, particularly multi-classification ML is used.

To evaluate the effect of weather variables in predicting maritime accident risk, three sets of feature lists are created for training ML algorithms. First, predictions of accident type are made based on only the database from NMA. The list of variables (or features) used are reported in Table 6 (Appendix D), where NMA features are referred to as *Standard features*. Second, predictions of accident type are made using only the *Weather features* in Table 7 (Appendix D). Feature statistics were calculated on 80 % of the dataset (training and validation). The holdout part is separate. Lastly, a combination of both sets of variables is used.

3.4. Automated machine learning (AutoML)

Machine learning is ideal to use when you have complex tasks with many variables and a large volume of data. As the literature review in Section 2 suggested that several ML models could prove valuable, a more efficient Automated machine learning (AutoML) approach is adapted in this study. AutoML can be implemented through various cloud AI platforms, using many of the same algorithms. Hutter et al. (2019) describe AutoML as a democratisation of ML. Except for Munim et al., (2024), the application of AutoML in maritime accident prediction is limited. Thus, this study explores the potential of a wide range of models simultaneously by using AutoML. This study uses DataRobot, which offers many different algorithms and evaluation metrics in predictive modelling. DataRobot is an AutoML platform for data analytics, allowing the building and deployment of predictive models accessible to data scientists. DataRobot provides an extensive range of algorithms and efficient pre-processing tools and offers valuable insight into large datasets. Machine learning in general, also the ones trained in this study have certain limitations. ML models can be difficult to understand, and their predictions could vary significantly for unseen new data, which is why several accuracy measures are considered and multiple cross validation folds are used when training the models.

3.5. The AutoML process

The starting point of the AutoML phase in DataRobot is to import the prepared dataset from outside of DataRobot. The next phase is the automatic pre-processing of the dataset in DataRobot, which includes the removal of outliers and anomalies, identifying variable types, formatting the data, and generating statistics (DataRobot, 2022a). The platform recognises different types of variables such as categories, date, time, numerical values and location. After the automatic pre-processing, it was necessary to manually change a few variable types, such as *location* since this variable was uploaded as latitude and longitude in separate columns.

3.5.1. Features

After finalising the pre-processing, it was necessary to create feature lists as three different set of features were to be trained. Additionally, some data, such as identification number and year of the accident, originally a part of the larger dataset, was irrelevant for further modelling. By making different feature lists, it is possible to test different dataset variables separately. For classification problems, it is necessary to choose a feature, that would be predicted, also known as the target feature. The target feature in this study is *accident type*.

3.5.2. Split dataset, train, and test models

The models are made using the Autopilot function in DataRobot. First, the models are trained, using 16 % of the dataset, and using the selected feature list. Thereafter, the top 16 ML models are selected and trained again on 32 % of the dataset. The 8 best performing models are then trained again on 64 % of the dataset (DataRobot, 2023d). After finalising the training, the models are validated using a separate, untouched dataset partition, containing 16 % of the data (i.e. validation sample). Out of these, the best-performing model is tested on the holdout partition, which is set to 20 % of the entire dataset. This partition is similar to the commonly used 60/20/20 split in ML studies.

3.5.3. Cross-validation

Cross-validation (CV) is a default setting in DataRobot where the training and validation dataset is divided into five different partitions. The holdout partition is fixed and does not change, but the training and validation partitions change. The CV accuracy as an evaluation metric is the average of all five CVs. The advantage of validating in this way is to gain a better estimate of the model's performance in predicting new data by minimizing the chance of model overfitting, which is a well-known limitation in ML modelling. A disadvantage of this method is that it is more time-consuming and computationally intensive than only validating once (DataRobot, 2023e). This study uses the five-fold CV method for validation.

3.5.4. Analyse the model outcome

The final step of the modelling process is to analyse and evaluate the model outcome based on a number of relevant evaluation metrics. The final step of the AutoML process using DataRobot would be to deploy the models for predictive measures, but that is outside the scope of this study.

3.6. Evaluation of the prediction results

DataRobot provides several metrics to be used to evaluate the performance of the classification models. *Accuracy*, *balanced accuracy*, *LogLoss*, *lift charts* and *confusion matrix* are the most relevant evaluation tools included in this study. *Feature impact* and *feature effect* offer insight into the importance of the different variables and predicted accident risk under various conditions of those variables, respectively.

Accuracy is the ratio (a value between 0 and 1) of the total correct prediction relative to the total number of all predictions. Correct predictions can be either true positives (TP) or true negatives (TN), while false negatives (FN) and false positives (FP) are incorrect predictions (DataRobot, 2023a). Accuracy can be expressed as in Eq (1). The F1-score is a product of *precision* and *recall* of the model and is another measure of the model's accuracy (DataRobot, 2023c).

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{TP + TN + FP + FN} \tag{1}$$

$$PositivePredictiveValue(PPV) = Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{2}$$

$$TruePositiveRate(TPR) = Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{3}$$

$$F1Score = \frac{2(\mathrm{PPV}^*\mathrm{TPR})}{PPV + TPR} = \frac{2TP}{2TP + FP + FN} \tag{4}$$

Balanced accuracy only looks at the TP and FN predictions per class, and ignores all FP and TN the prediction gives. Balanced accuracy is the sum of the recall values of each class, divided by the number of classes. Balanced accuracy is a value between 0 and 1, measuring how often the model predicts correctly when the class is actual. DataRobot recommends this measure for multiclass classification problems (DataRobot, 2023a). Balanced accuracy values of all the trained models are provided in the supplementary excel file.

LogLoss is an inaccuracy measure of the predicted probabilities. The value increases as the predicted probability diverges from the actual classification. The metric is logarithmic, so the further away from the true class the prediction is, the more rapidly increases the LogLoss. As the prediction gets closer to the actual class, LogLoss moves towards 0 (DataRobot, 2023a). LogLoss values of all the trained models are provided in the supplementary excel file.

The *confusion matrix* is a tool to help evaluate multiclass model performance. It visualises any mislabelling of predictions, and shows the distribution of TP, FP, TN and FN. The matrix can be used to see what inputs are classified as when incorrectly classified (DataRobot, 2023a).

Feature impact shows the modelling importance for each variable, relative to the most important one. In other words, it shows what variables impact the predictions the most. The most important factor has 100 % as a default, and other variables impact is relative to this, also measured in percentage (DataRobot, 2023b).

Feature effects rank variables based on their impact score and show the effect of changes in the value to the prediction. The results are displayed as a graph, illustrating how a specific model uses a particular variable.

4. Findings

DataRobot trained 36 models on the NMA features list, 16 models on the weather features list and 63 models on the combined list. The best-performing model for all three tests on 64 % sample size was the *Light Gradient Boosted Tree Classifier with Early Stopping*. However, the results change slightly when 100 % sample size is considered, see Table 3. The combined use of NMA features and weather features gives the best prediction accuracy of up to 70.23 % at cross-validation, meaning that weather data increases the accuracy of predicting maritime accidents. It is noteworthy that without combination, the standalone NMA dataset or the weather dataset have similar prediction accuracy. Accuracy metrics of top performing models on the combined dataset are reported in Table 4. A list of 30 models is reported in Table 8 in Appendix D, and a full list of 105 model is provided as supplementary material.

The confusion matrix in Fig. 3 provides an overview of performance of the best-fitting model and how the different accidents are predicted. *F1-score*, *recall* and *precision* for each accident type is also presented, with *grounding* achieving the overall best scores. *Recall* is the true positive rate, or how often the model predicts correctly when the class is actual. *Precision* is how often the model is right when a class is predicted. The green circles are correct predictions, and the red are wrong predictions. The larger the circle is, the more correct or wrong predictions are made, thus large and green circles are good. The blue bars on the x- and y-axis represent the quantity of each accident type.

4.1. *Light gradient boosted model*

Light Gradient Boosted Trees Classifier with Early Stopping (LightGBM) was the model recommended by DataRobot. It achieved the highest scores on all relevant model performance evaluation parameters. LightGBM uses a tree-based algorithm and is known to handle large amounts of data efficiently and accurately (DataRobot, 2022b). The models are considered to be among the most versatile and handle missing data in an effective manner, which the analysis in this study benefits from. The models are based on AdaBoost, an algorithm developed by Freund and Schapire (1995), and are related to random forest models. GBM is different from Random Forest in the way that GBM uses the residual errors from all the previous trees combined in the new ones (DataRobot, 2022b). Atak and Arslanoğlu (2022) also achieved good results with LightGBM in their study of container port accidents. They used container handling data, weather data and accident reports from container terminals in Turkey to predict accidents with ML classification techniques. They achieved an accuracy of more than 97 % with LightBGM for *binary* classification. Note this study is a *multi-classification* ML application.

The Early Stopping function of the model is a design element to determine the number of trees to use. 10 % of the training data is saved for early testing, to find the breaking point where additional data no longer provide more accuracy. By finding this point, the model will prevent overfitting and save time (DataRobot, 2022b). Fig. 4 shows the blueprint of the LightGBM demonstrating the working mechanism with the different variable types.

4.2. *Learning curves*

The learning curve is a visualisation of how the models learn from the data. Fig. 5 shows the learning curves during the training period. Only the combined feature list with a sample size of 64 %, is shown in the figure. The accuracy is at its highest as the sample size reaches 64 %, and the learning rate for LightGBM is quite similar in all validations.

4.3. *Feature importance ranking by accident type*

Based on the LightGBM model from the combined feature list, the feature impacts for each accident type and on an aggregate level is reported in Fig. 6. The 25 most important features are listed for each accident type, with the most important feature at the top at 100 %. The results show that *Phase of Operation*, *Waters* and *Geographical area* are the most influential variables aggregated, with all three represented among the top four features for all accident types. The variables from the NMA score relatively high compared to the weather variables. Among the weather variables, *cloud cover* and *wind speed* features have the highest impact. *Wind speed* features has its greatest impact on *heavy weather* accidents. Feature impacts on the aggregate level based on only weather and standard feature lists are reported respectively in Fig. 10 and 11 in Appendix B.

Table 3
Models for all feature lists.

Feature list	Best performing model	No. of models trained	Sample size	Accuracy (Validation)	Accuracy (Cross-validation)	Accuracy (Holdout)	Balanced accuracy (Cross-validation)	Log Loss (Cross-validation)
Only NMA	Light Gradient Boosted Tree Classifier with Early Stopping (SoftMax Loss) (64 leaves)	36	100 %	0.6506	0.6486	0.6352	0.4466	0.9013
Only weather	RandomForest Classifier (Gini)	16	100 %	0.6310	0.6354	0.6428	0.3261	1.0347
Combined	Light Gradient Boosted Tree Classifier with Early Stopping (SoftMax Loss) (64 leaves)	63	100 %	**0.7143**	**0.7023**	**0.6878**	**0.5055**	**0.8161**

Table 4
Best performing models, combined feature list.

Feature list	Best performing models	Sample size	Accuracy (Validation)	Accuracy (Cross-validation)	Accuracy (Holdout)	Balanced accuracy (Cross-validation)	Log Loss (Cross-validation)
Comb.	Light Gradient Boosted Tree Classifier with Early Stopping (SoftMax Loss) (64 leaves)	100 %	**0.7143**	**0.7023**	**0.6878**	**0.5055**	**0.8161**
Comb.	Light Gradient Boosted Tree Classifier with Early Stopping (SoftMax Loss) (64 leaves)	80 %	0.6859	0.6815	0.6884	0.4740	0.8519
Comb.	Light Gradient Boosted Tree Classifier with Early Stopping (SoftMax Loss) (64 leaves)	64 %	0.6859	0.6769	0.6715	0.4621	0.8553
Comb.	RandomForest Classifier (Gini)	100 %	0.6831	0.6759	0.6667	0.4593	0.8592
Comb.	eXtreme Gradient Boosted Trees Classifier with Early Stopping	100 %	0.6886	0.6903	0.6873	0.4947	0.8397
Comb.	Gradient Boosted Trees Classifier with Early Stopping	100 %	0.6933	0.6805	0.6770	0.4816	0.8471

4.4. Feature effects for accident types

The feature effects result of a selection of the most important features for all accident types are presented in this section. Fig. 7 illustrates the *partial dependence feature effect* for *Phase of operation*, *waters*, *CcoverNM*, *WspC*, *MphaseC*, *PressD* and *VisD*. The partial dependence score indicates influence of a feature's values or conditions on the target, considering the effect of all other features included in the model. *Phase of operation* and *waters* are among the most important features from NMA for all accident types, and the rest of the selection are important weather features for most accident types. The values presented in Fig. 7 are on a scale between 0 and 1, where 1 indicates the highest effect. Feature effect values indicate how the predictions are affected by changes in the underlying feature.

5. Discussion

In this study, 105 ML algorithms are trained and evaluated for predicting maritime accident risk. The training and testing datasets comprised historical maritime accident records and a comprehensive list of weather data as input variables. The results indicate that weather variables can help improve the prediction accuracy of maritime accidents. Moreover, the results provide new insights about the differences in predicting various types of maritime accidents.

The best performing model in this study reached 70.23 % accuracy on five-fold cross-validation. Previous studies reported somewhat similar accuracy levels. For instance, Theofilatos et al., (2019) reported total accuracy of 68.95 % when predicting real-time crash utilizing ML models. Some studies exist reporting higher levels of accuracy, such as Yang et al., (2022) reports a range of 85–91 % accuracy when predicting maritime traffic accidents. The accuracy of prediction models in this context relies on several factors. From a ML application point of view, these are the target variable properties (e.g. grounding or not binary target), the sample size and its context, and the ML training properties such as number of cross validation folds influence the accuracy. This study is predicting multiclass accidents (i.e. the target), which has five classes. Further, it uses fivefold cross validation, which increases generalizability of the model but reduces accuracy.

5.1. Impact of standard features

Results from feature impact indicate that *Phase of Operation*, *Waters* and *Geographical area* are the overall most important features in predicting maritime accidents (Fig. 6). *Waters* and *Geographical area* are highly related since they both are variables describing the area of accident and its surroundings. *Waters* will in some cases limit the potential outcome, accident type. A grounding, for instance, is unlikely to occur in open seas as the ship is navigating far from the ground, and it has therefore achieved a low score of 0.35 in *feature effect*. On the opposite end of the scale, we have *narrow coastal waters* with 0.70, as the feature with the highest effect for groundings (Fig. 7b). Surprisingly, *open seas* have a much higher effect on *collisions* than *narrow coastal waters* (Fig. 7b), where the overall concentration of vessels and accidents are significantly higher. The reason for this could potentially be that seafarers are more vigilant navigating narrow waters, and therefore avoid accidents. *Phase of operation* is, similarly to *Waters* and *Geographical area*, related to where a vessel is operating. Some operations, for example *arriving in port*, have a lower effect on collisions than if the vessel is doing *sport exercises*, where one can assume that the vessel is more concerned with that specific activity rather than navigational safety.

5.2. Impact of weather-related features

The most important weather variables are *wind speed*, *sea level pressure*, *visibility*, *cloud cover, and moon phase*.

5.2.1. Wind (WspC)

Wind speed at the time of the accident (WspC) has a large impact on predictions of *collisions* and *contact* accidents. Zhang and Li (2017) came to the same conclusion that heavy sea states were the primary risk factor. Waves are mainly generated by wind over time, but they are restricted in narrow waters by land areas. The largest wind-generated waves are therefore found in open sea areas where waves formation is unrestricted. As wind has its most significant relative effect on *collisions*, and *open seas* have a relatively high feature effect on *collisions*, it could therefore be that waves, as a result of wind, increase the probability of collisions.

Furthermore, wind speed's relative feature effect on heavy weather damages increases rapidly as wind speed exceeds 38 knots (Fig. 7d). The feature effect scores are pretty low, which is likely due to dataset imbalance. Heavy weather damage scores relatively high on outer coastal areas and open seas, typical areas where wind over time has the potential to generate larger waves.

Wind speeds have an increasingly significant effect on contact damage accidents. At around 15 knots of wind, one can see a sudden incline in its feature effect (Fig. 7d). However, one can find a decline in feature effect for collisions at approximately the same levels. Uğurlu et al. (2018) found strong winds to be involved in 12 of 30 wind and sea related collision and contact accidents. To better understand the differences between collision and contact accidents, we need to analyse them separately.

The feature effect indicates that wind is affecting the two accident types quite differently. For contact type, which typically occurs when vessels manoeuvre in the vicinity of quays and bridges, wind has a relatively linear incline in feature effect. Additionally, *phase of operation* has the most significant feature impact on contact accidents, which can be related to the feature effect of *contact damage* to *phase of operation*, where arriving port has the greatest effect. This could indicate that wind is linearly, positively correlated with the probability of contact damages with vessels arriving in port.

The situation for collisions, however, which theoretically could occur anywhere, is quite different. Wind's feature effect on collisions decreases almost linearly towards around 35 knots, when it suddenly increases rapidly. Since open water has the most significant feature effect on collisions, one can argue that winds over 35 knots in open waters increase the probability of accidents. If we consider the effect of wind sea, the results can indicate that waves generated by winds exceeding 35 knots influence the safety of navigation negatively in some way. The study of meteorological data in marine risk assessment on voyages in the North Pacific by Adland et al. (2021), found an increased likelihood of insurance claims as wind speeds exceeded 35 knots. This is in line with the findings by Knapp et al. (2011), where increasing wind speeds led to a higher probability of serious casualty. Interestingly, there is a drop at around 35 knots in feature effect for grounding, where stronger winds have a decreasing effect on the probability of the accident being a grounding. This could perhaps indicate that accidents with such high wind speeds are categorised as *heavy weather damages,* or that fewer vessels navigate in shallow waters when weather conditions are severe enough. A third possible explanation could be found in the study by Adland et al. (2021), where a positive relationship between wave height and predictions of assurance claims was tapering off at above three meters in wave height. The study was done on vessels transiting the North Pacific, where groundings were likely to be underrepresented, so one should interpret the relationship with caution.

5.2.2. Sea level pressure (PressD)

In Table 7 in Appendix D, we can find the mean sea level pressure of all accidents to be 969mb at the time of the accident and 1010mb for the day. The daily mean falls within the lower part of the 1010-1020 mb range, which Adland et al. (2021) identified as having the most significant impact on maritime accident predictions. Their observations vary from this study, and one can therefore assume that the sea level pressure does not necessarily operate in the same range. Another explanation, as Adland et al. (2021) discuss, is that high and low-pressure systems are quite stable, while sudden drops in sea level pressure are associated with bad weather conditions. The resolution in weather data in their study is three hours, while for our study it is one hour, which could indicate that the lower *PressC* compared to *PressD* in our study captures the effect of sudden drops in sea level pressure more accurately.

In Fig. 7f, the feature effect of sea level pressure for collisions, contact and heavy weather damages follows the same negatively correlated pattern. It has its most significant effect at low values and decreases as the sea level pressure increases. For grounding the trend is the opposite, as the effect of changes in pressure increase as it rises. The literature does not provide any explanation as to why that is, but good weather will likely stimulate activities and vessel densities inshore, which is where groundings occur.

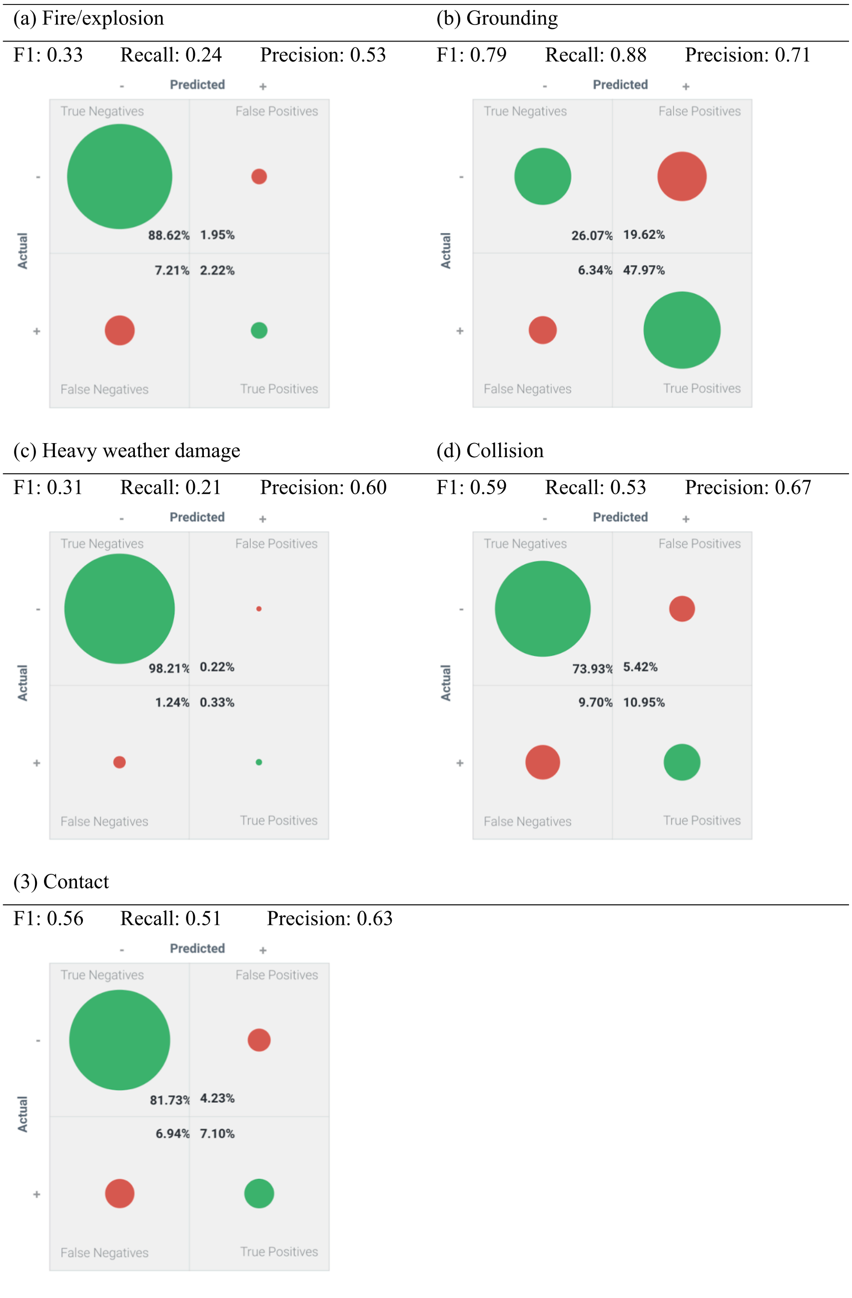


Fig. 3. Confusion matrix and evaluation scores.

5.2.3. *Visibility (VisD)*

In some previous studies, it was found that the number of navigational accidents increases when visibility decreases, with navigational accidents being Collisions and Groundings (Bye & Aalberg, 2018). For example, when the visibility decreased to less than 2 nm, the share of navigational accidents increased to 77–88 % (Bye & Aalberg, 2018). However, in our study, the feature effect for groundings and collisions indicates that they follow different patterns. The partial dependence for grounding is at its lowest when the visibility is at its lowest (Fig. 7g); they are positively correlated. Collisions however, similar to contact and heavy weather damage, are negatively correlated with visibility, as visibility increases, the partial dependence decreases. The confusion matrix in Fig. 3 shows that when the model predicts collisions, it most commonly confuses it with groundings. The high ratio of false negatives to true positives provides a different perspective on this; even though the precision is good, the model has a high accuracy in correctly classifying non-collisions.

Table 7 in Appendix D reveals that the mean visibility of all accidents combined is below the daily mean, which could support the

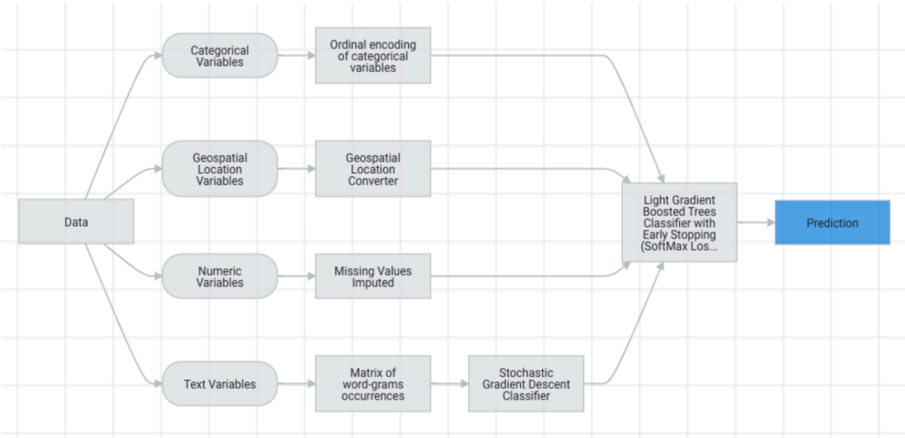


Fig. 4. Blueprint of light GBM classifier (Source: DataRobot).

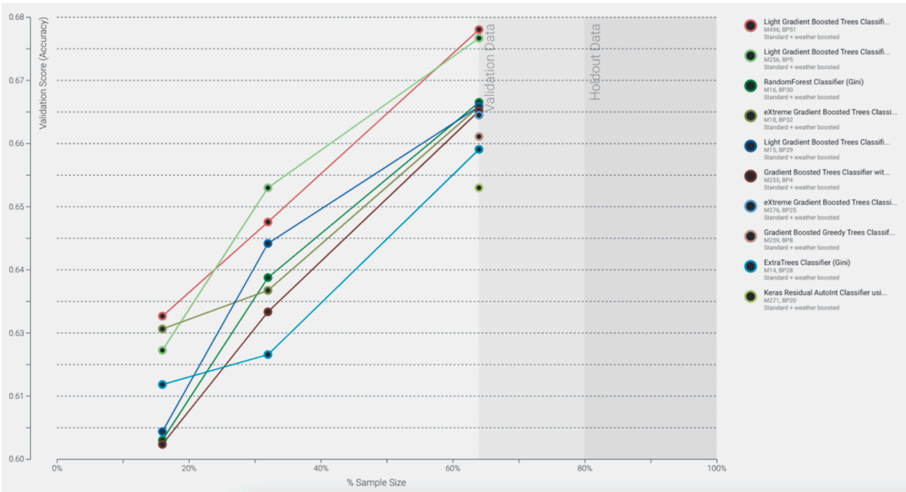


Fig. 5. Learning curves, combined feature list.

findings of Fan et al. (2020), who found fog and poor visibility to be among the most common causes of maritime accidents. Uyanık et al. (2021) argued that visibility was more critical in shallow waters due to its increased risk to navigational safety. They further revealed that estimations of visibility by ML methods could reduce the risk of maritime accidents, which the results from this study also support. Even though Uğurlu et al. (2018) found violations of COLREG Rule 5 (improper lookout) in many of the analysed accidents, our study does not point towards a higher risk of groundings due to reduced visibility. Visibility, as interpreted from Fig. 6, seems to have a higher impact on collisions than groundings, and other accident types.

5.2.4. *Cloud cover (CcoverMN)*

Cloud cover mean normal has achieved the highest aggregated impact of all weather variables (Fig. 10). It has the highest impact on groundings and collisions, two of the three most common accident types. The reason this variable has such a high impact is likely not because of cloud cover, but rather another locational variable since all *normal* values are generated over time, and do not describe the weather at the time of the accident.

Clouds are generated as humid hot air is cooled down to the dew point temperature. Land is heated faster than the ocean, and the hot air will therefore rise. As the air rises, it will be cooled down, generating clouds. The feature effect indicates a positive correlation for groundings, and a negative correlation for collisions. As grounding has its highest partial dependence on extensive cloud cover, one

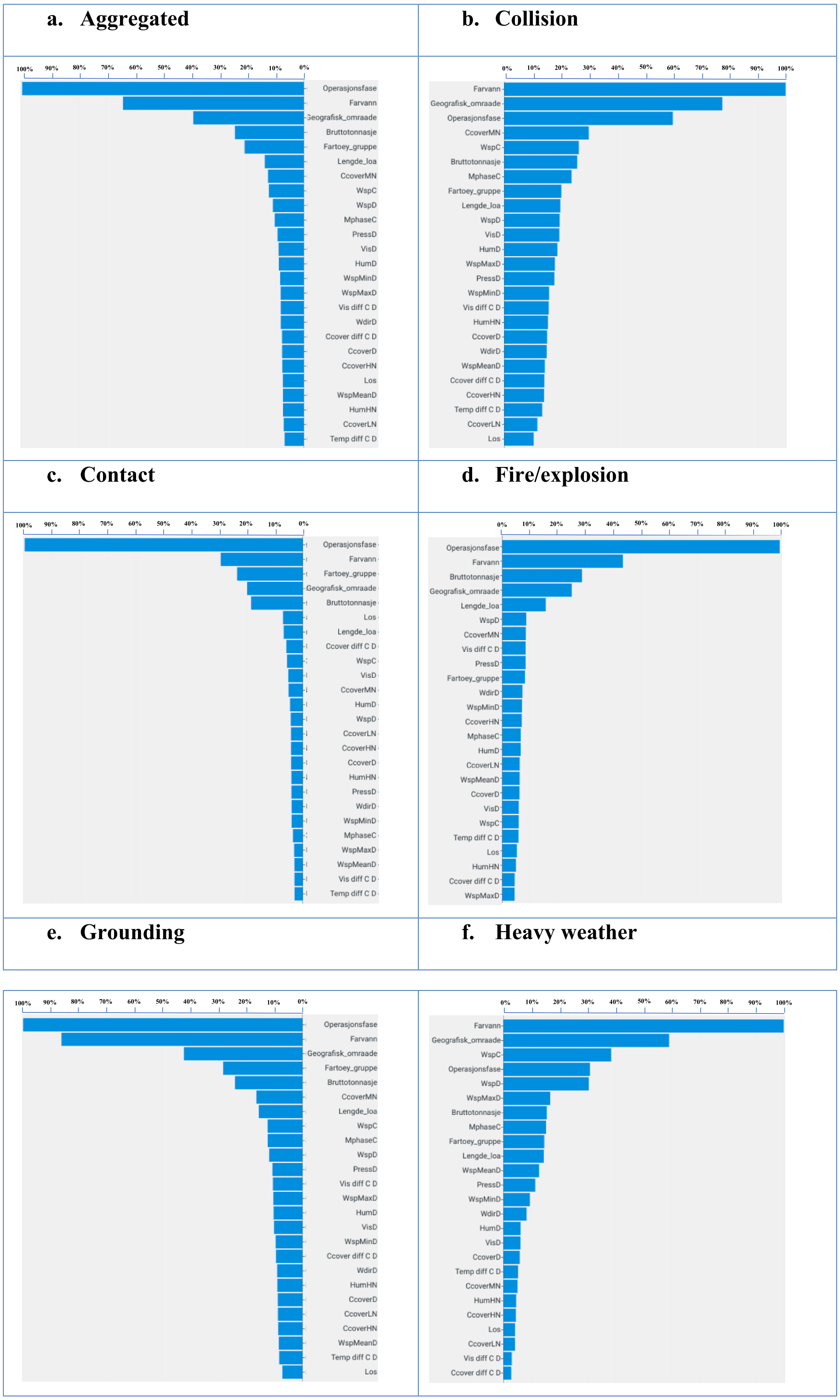


Fig. 6. Feature impacts (see English form of the features in Table 5 in Appendix C).

Fig. 7. Feature effects. (Y-axis indicates the feature effect, with values ranging between 0 and 1. X-axis refers to the feature values or conditions; see English translation in Appendix C).

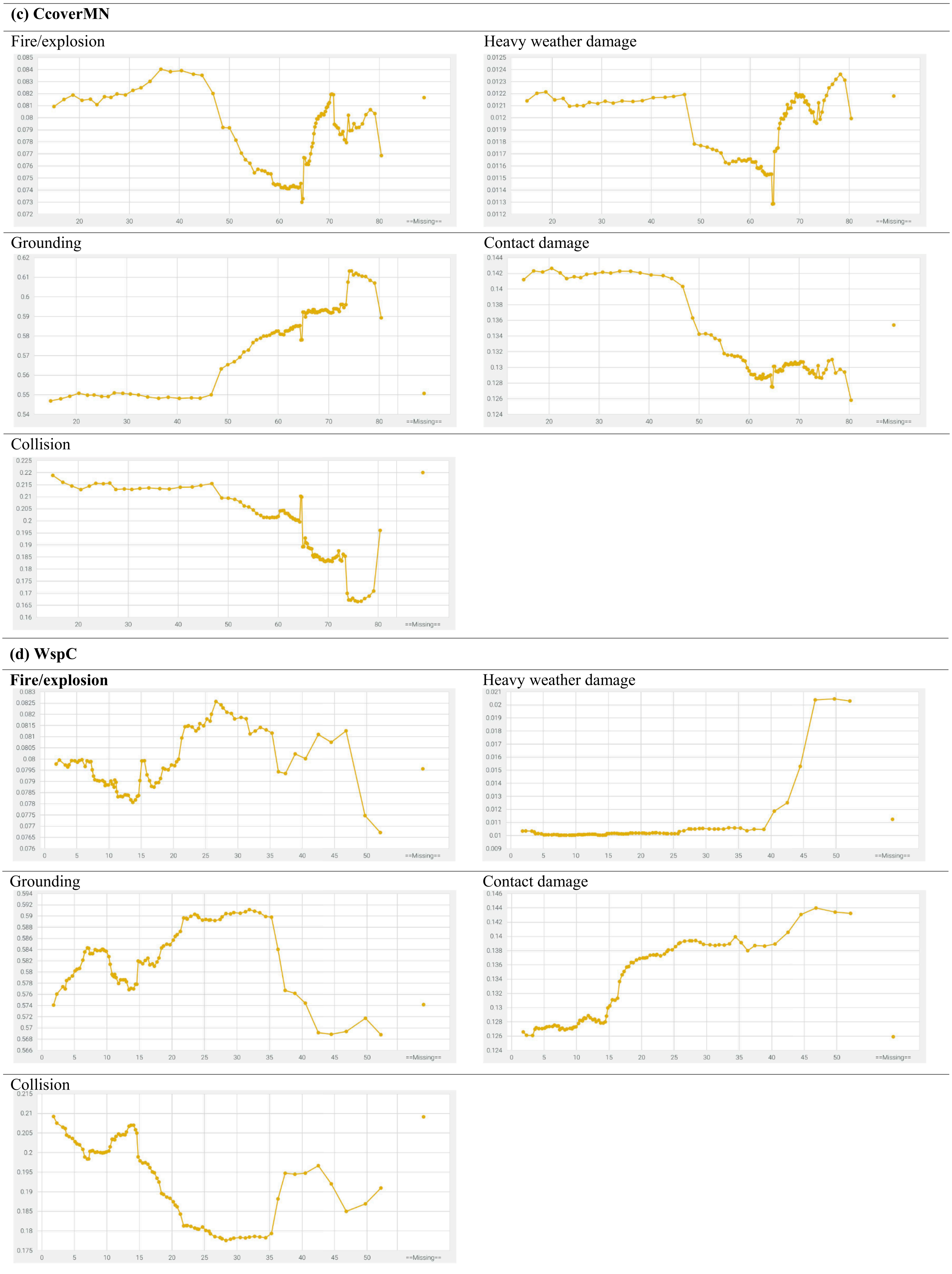


Fig. 7. (*continued*).

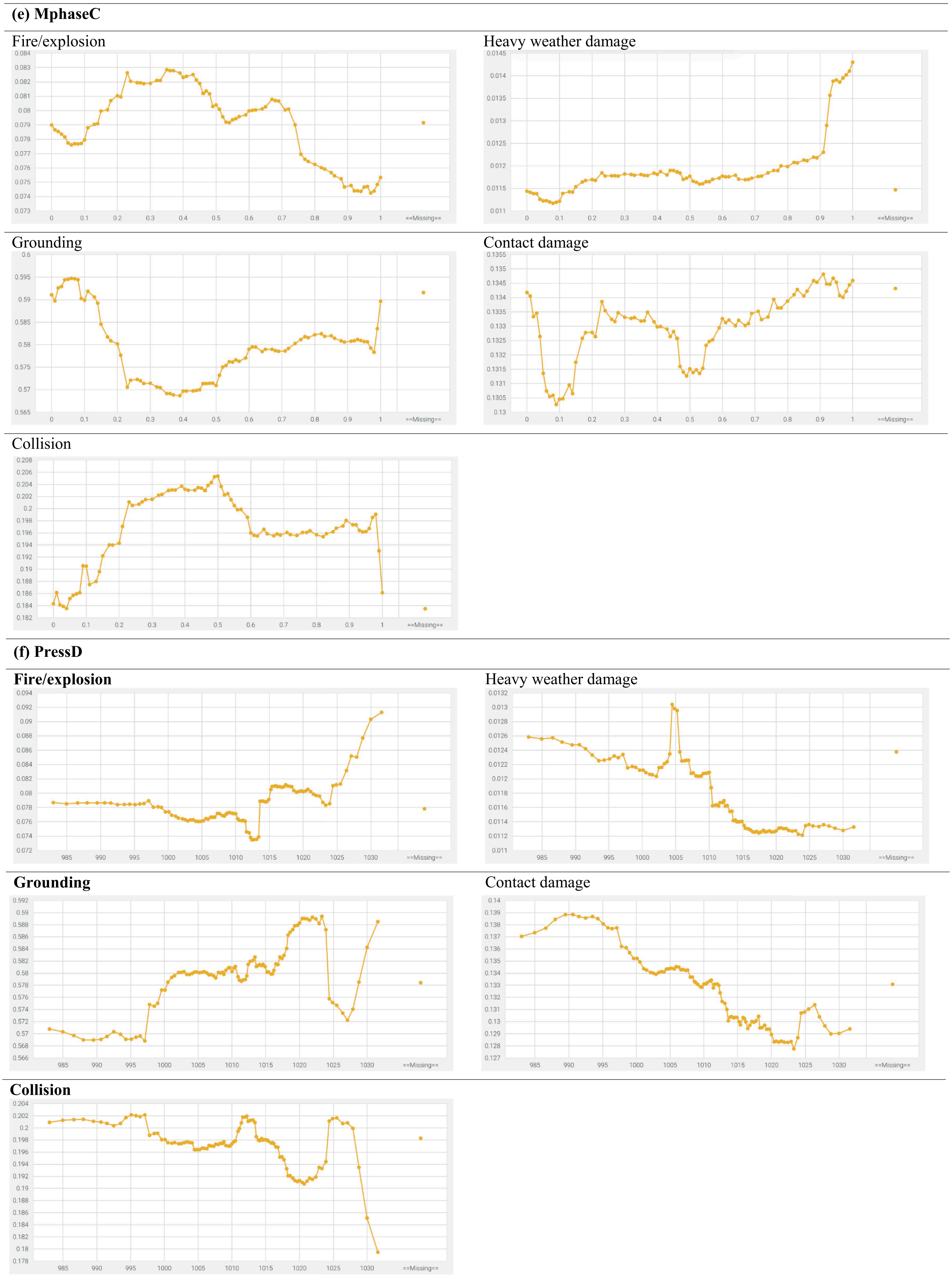


Fig. 7. (*continued*).

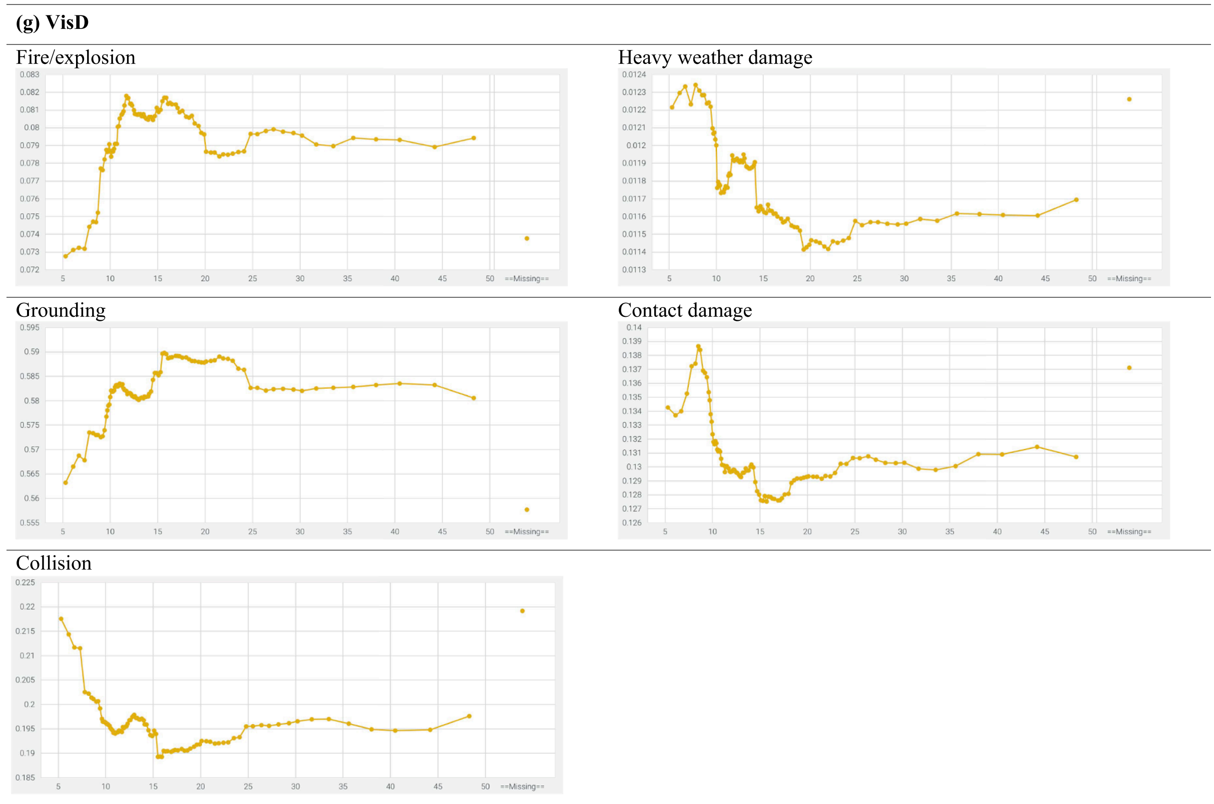


Fig. 7. (*continued*).

can assume most of the accidents occur in shallow waters, and that shallow waters have a higher mean *normal* cloud cover. Collision, which has its feature effect score in open seas, has its highest partial dependence when *CcoverMN* is the lowest (Fig. 7c). Contact damage follows the same pattern as collision, even though the scores are lower. *Oil fields* and *alongside quay* are the categories in contact damage where waters have the largest effect. Quays are typically inshore, and oilfields are offshore, which points in different directions. To understand why collision and contact damage follow the same trend, further studies must be done.

5.2.5. *Moon phase*

The effect of the moon on maritime accidents has to the author's knowledge not been mentioned in previous literature. Fig. 8 illustrates the different moon phases. A well-defined distribution of moon phases is observed in Fig. 9 in Appendix A, with high concentrations around new-moon and full-moon. Tides are generated as the earth rotates around its axis, but the variations in tides depend on how the moon is positioned relative to the sun. When the moon is positioned in line with the sun and earth, being either a new- or full moon, the gravitational forces combine and generate spring tide, which means tidal variations larger than the rest of the cycle. The results show more accidents occurring during these moon phases. If we combine the three largest out of the 30 bins in Fig. 9 (0.0–0.02, 0.50–0.52, 0.98–1.0), which would nearly represent the three most dangerous days in each moon cycle, it will represent 28.6 % of all accidents.

The feature effect on grounding in Fig. 7e shows high partial dependence around new-moon, and lower at full-moon. We are likely to see darker nights when there is a new moon, as there will be less illumination from the moon. The findings could therefore indicate that lack of illumination from the moon increases the risk of groundings. As for collisions, the situation is the opposite as we see the lowest partial dependence towards new-moon, and the highest around full-moon.

The impact of the moon phase is more critical to collisions (ca. 22 %) than for grounding (12 %), as we can see from Fig. 6. At spring tide, as the tidal variations are most significant, the current generated will also be the strongest. Currents like this could, if the geographical area allows, fall into the category of *conditions preventing vessel motion*, from Uğurlu et al. (2018). Navigators who fail to stay vigilant, or experience technical issues in the wrong places, will be penalised far more these few days.

6. Conclusion

This study has utilised AutoML to explore if weather data can provide improved predictions of maritime accident risks. An extensive range of weather variables has been integrated into a dataset collected from the NMA. Three sets of ML models were explored: (1) for NMA data alone as benchmark, (2) for weather variables alone, and then (3) with both combined to find if the weather data provided improvements to the predictions. The results indicate that weather data could improve predictions of most of the accident types. Fire and explosions, groundings, heavy weather damage and collisions all achieved improved prediction accuracy when weather data was included. The prediction accuracy for contact damages was the only accident type which did not experience improvement with the weather data.

The overall strategy behind collecting weather data was to gather as many variables as possible, and then add them to the feature list for training ML models. This revealed the leading weather variables, some were new to the authors' knowledge. The *phase of operation*, *waters* and *geographical area* are the three most essential features on the aggregated level, provided in the original database from NMA. *Wind* at the time of the accident, *sea level pressure*, *visibility*, *statistically mean cloud cover* for the time and location of the accident, and *moon phase* were the five most essential weather features provided to the combined feature list. Their impact and importance vary between the various accident types. This study found that the Light Gradient Boosted Trees (LightGBT) classifier performs best on the combined feature list. The accuracy for the best model was 71.43 % in validation sample and 70.23 % in five-fold cross validation sample. The learning rate of the models indicates that additional data might continue to improve the model's performance.

6.1. Further implications of the study

The best-performing model can be implemented as an AI software package to ships to facilitate safety resilience. The AI package can indicate real-time accident risk based on data from multiple sensors and APIs from weather information providers. For insurance companies, the results could be utilized to optimize premiums by rewarding risk-averse ship owners, and increasing premiums for risk-seeking ship owners who otherwise would capitalise on sailing shorter, more weather-risk-influenced routes. Real-time monitoring of vessel motion and current and forecasted weather, combined with the AI package could work as an economical incentive for ship owners and captains to reduce risk. The models could also be utilized by Vessel Traffic Services, Rescue Coordination Centres and Coast Guards to announce specific dangerous areas, divert traffic, and optimize the use of Search and Rescue resources. The previously discussed wind speed levels of around 35 knots, combined with the findings regarding new and full moon should be a relevant starting point for optimization for all stakeholders.

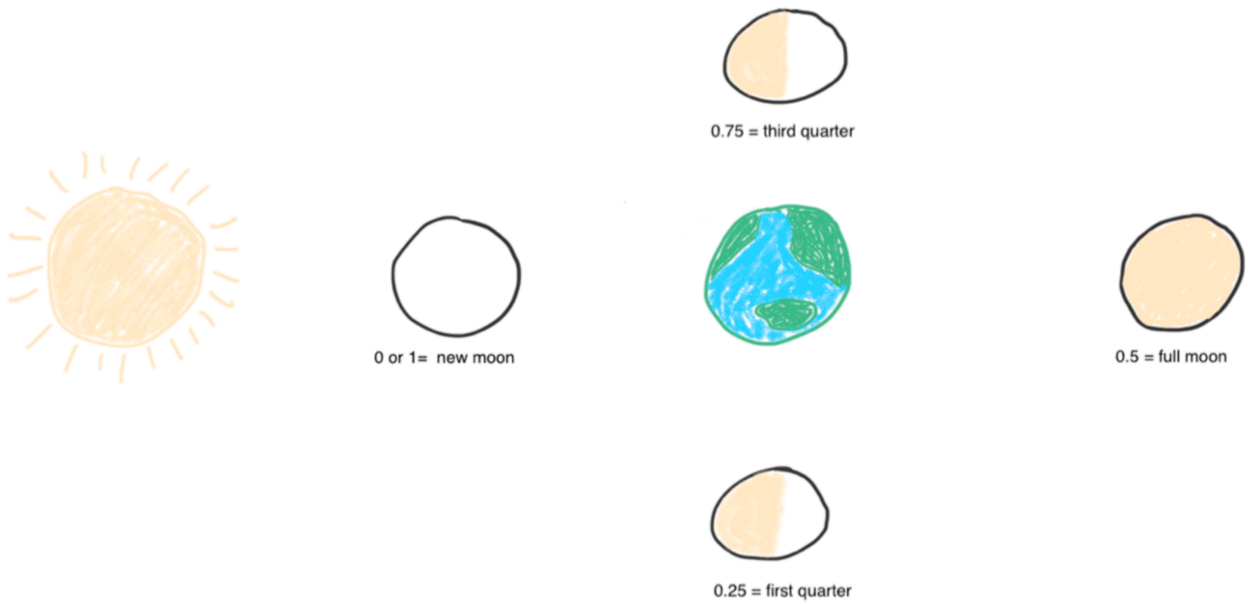


Fig. 8. Illustrations of moon phases.

6.2. *Limitations and future research directions*

One major advantage of AutoML is its ability to analyse large quantities of data in a short amount of time. In this study, a lot of different weather variables have been introduced to various models, where some have proven valuable, and others have not. Further research should pursue the proven weather variables and seek to optimise them. The implementation of AIS data is also recommended, and necessary, to be able to provide valuable insight into the actual risk presented to the vessels. Since locational variables have proven valuable to the models and are of great importance to certain predictions, it is recommended to study limited geographical areas separately. One example could be to split inshore and offshore accidents. Further research on the moon's effect on accidents could also be worth exploring.

The ML models have revealed that weather data can provide accuracy to maritime accident predictions. However, the dataset is imbalanced as there are far more groundings than other types of accidents. To control for this, this study adopted the five-fold cross-validation. Future studies should explore other possibilities such as Synthetic Minority Over-sampling Technique (SMOTE), data augmentation, and adjusting the decision threshold for the classifier to increase the sensitivity of the minority class. Human factors can provide enhancements to the model's predictive accuracy by evaluating OOW's decision-making abilities under different weather conditions. Future studies can explore seafareŕs mental abilities and ít́s effect on decision-making, and if external factors such as bad weather adversely affect their performances.

CRediT authorship contribution statement

Peter Brandt: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization, Writing – review & editing. **Ziaul Haque Munim:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Conceptualization. **Meriam Chaal:** Writing – review & editing, Validation. **Hooi-Siang Kang:** Writing – review & editing, Validation.

Data availability

Data will be made available on request.

Acknowledgements

Ziaul Haque Munim and Hooi-Siang Kang acknowledge funding support from the Norwegian Research Council under the NORGLOBAL2-Norge - global partner (project no.: 322410).

Appendix A. Distribution of data

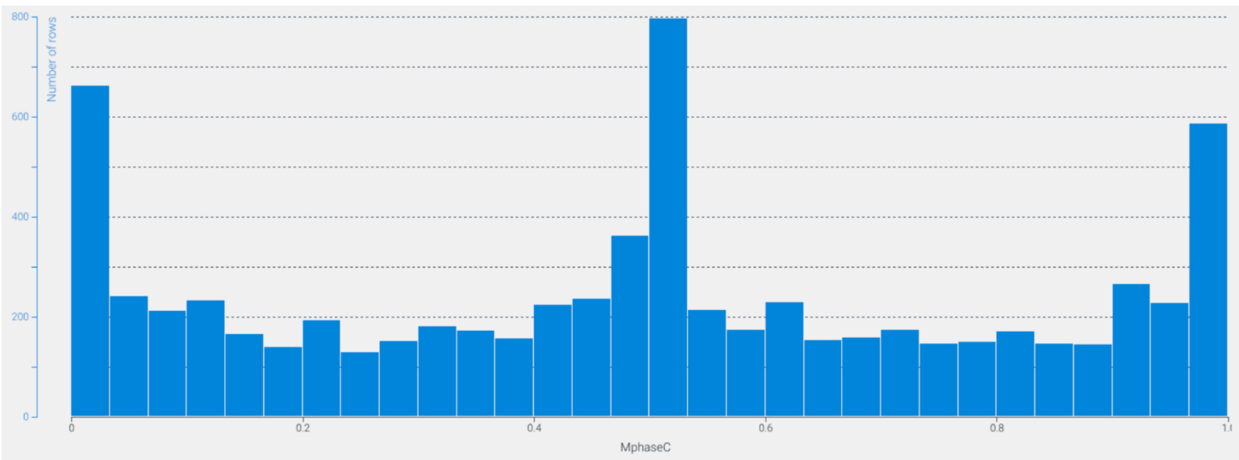


Fig. 9. Moon phase distribution

Appendix B. Feature impact

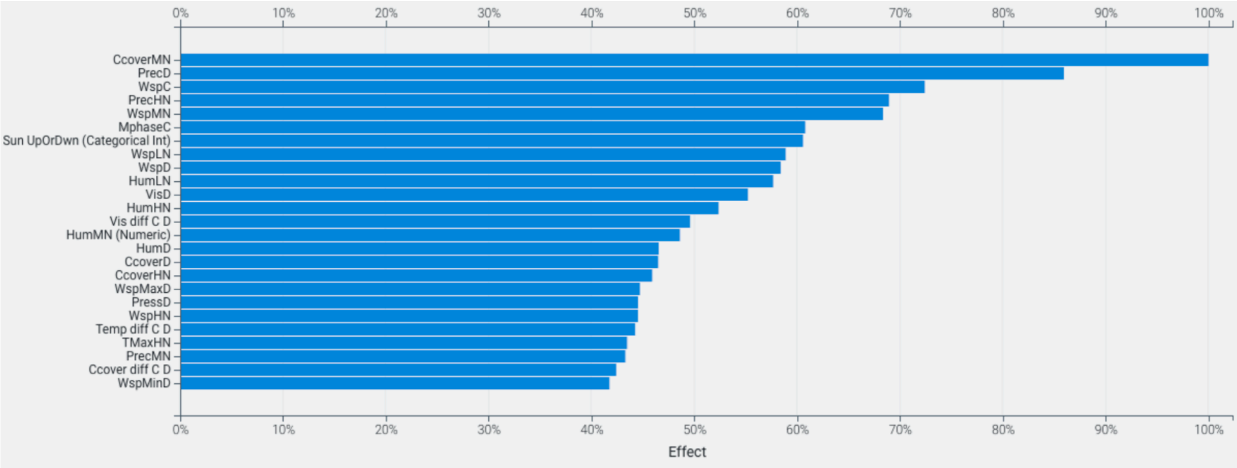


Fig. 10. Feature impact for weather features only.

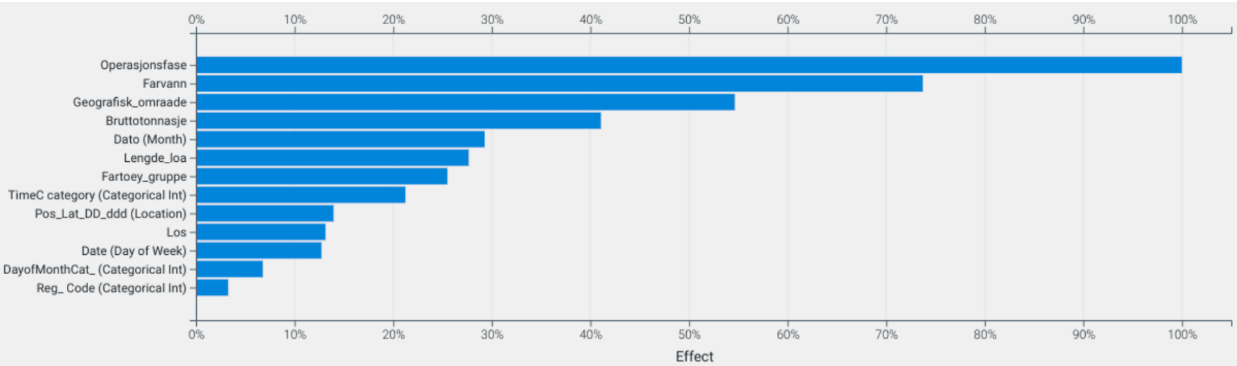


Fig. 11. Feature impact for standard features only.

Appendix C. Translations

Table 5
English form of Norwegian features.

Norwegian	English
Waters	
Annet	Other
Arkipelagisk	Archipelagic
I havneområde	Harbour area
Innsjøer	Lakes
Kanal	Canal, river
Langs kai	At quay (Moored in port)
Oljefelt	Oil Field
Separasjonsområde	Separation zone
Trangt kystfarvann	Narrow coastal waters
Ukjent	Unknown
Ytre kystfarvann	Outer coastal area
Åpent havområde	Open sea
Phase of operation	
I boreoperasjon	At drilling position
I drift	Drifting
I opplag	Laid up
Langs kai	At quay

(*continued on next page*)

Table 5 (*continued*)

Norwegian	English
Oppankret	Anchored
På dynamisk posisjonering (DP)	Dynamic positioning
Sikkerhetstilstand	Security condition
Sportsutøvelse	Sports exercise
Under avgang havn	Departing (port)
Under fiske	Fishing
Under slep	Under towing
Underveis	Underway
Ved ankomst havn	Arrival at port
Ved installasjon	At offshore installation
Ved lastebøye	At cargo bouy
Ved verksted	Shipyard

Appendix D. Feature lists

Table 6
Standard feature list.

Features	Description	Var type	Value	Frequencies and statistics	%
Water	Where the vessel was at the time of accident	Categorical	Narrow coastal waters	2726	36,9%
			Harbour area	1827	24,8%
			Outer coastal area	1482	20,1%
			Open sea	548	7,4%
			Canal, river	387	5,2%
			Missing	155	2,1%
			At quay	140	1,9%
			Oilfield	64	0,9%
			Separation zone	17	0,2%
			Lake	15	0,2%
			Other	10	0,1%
			Unknown	9	0,1%
			Archipelagic	1	0,0%
Phase of operation	What the vessel was doing at the time of accident	Categorical	Underway	4343	58,8%
			Arrival at port	1104	15,0%
			Departing	388	5,3%
			Fishing	295	4,0%
			At quay	190	2,6%
			Drifting	133	1,8%
			Anchored	102	1,4%
			Other	91	1,2%
			At offshore installation	73	1,0%
			Unknown	47	0,6%
			Under towing	25	0,3%
			Dynamic positioning	19	0,3%
			Laid up	10	0,1%
			At drill position	6	0,1%
			At cargo bouy	5	0,1%
			Shipyard	5	0,1%
			Sports exercise	1	0,0%
			Missing	544	7,4%
Day of week		Categorical	0: Monday	1061	14,4%
			1: Tuesday	1114	15,1%
			2: Wednesday	1059	14,3%
			3: Thursday	1161	15,7%
			4: Friday	1171	15,9%
			5: Saturday	936	12,7%
			6: Sunday	879	11,9%
Month		Categorical	1: January	866	9,4%
			2: February	808	8,8%
			3: March	830	9,0%
			4: April	651	7,1%
			5: May	630	6,8%
			6: June	719	7,8%

(*continued on next page*)

Table 6 (*continued*)

Features	Description	Var type	Value	Frequencies and statistics	%
			7: July	718	7,8%
			8: August	745	8,1%
			9: September	760	8,2%
			10: October	831	9,0%
			11: November	926	10,0%
			12: December	742	8,0%
Day of month	A categorization of days in the month	Categorical	1: 1-10	2509	34,0%
			2: 11-20	2415	32,7%
			3: 21-31	2457	33,3%
Vessel group	Type of vessel	Categorical	Cargo	3200	43,4%
			Fishing	2174	29,5%
			Passenger	1862	25,2%
			Pleasure	93	1,3%
			Mobile facility	24	0,3%
			Unknown	16	0,2%
			Missing	12	0,2%
Pilot	If the vessel had a pilot onboard at the time of accident	Categorical	Unknown	4215	57,1%
			Pilot onboard	537	7,3%
			Not required	1008	13,7%
			Pilot exemption certificate	69	0,9%
			Missing	1552	21,0%
TimeC category	A categorization of the time of accident according to standard bridge watches	Categorical	1: 00:00–03:59	720	9,8%
			2: 04:00–07:59	1152	15,6%
			3: 08:00–11:59	1066	14,4%
			4: 12:00–15:59	1065	14,4%
			5: 16:00–19:59	1366	18,5%
			6: 20:00–23:59	1097	14,9%
			Missing	915	12,4%
Reg code	To categorise Norwegian and foreign flagged vessels	Categorical	1: Norwegian	6811	
			2: Other	570	
Length LOA	Length of vessel in meters	Numeric	Missing	98	
			Mean	60.68	
			SD	56.10	
			Median	46	
			Min	1	
			Max	399	
Gross tonnage		Numeric	Missing	159	
			Mean	4494	
			SD	14,085	
			Median	429	
			Min	0	
			Max	185,398	
Geographical area	A textual geographical description of the area of the accident. Example: North Sea, West Coast Norway.	Text	126 unique descriptions		
Location	A composition of latitude and longitude	Location			

Table 7
Weather-related feature list.

Feature Name	Description	Var Type	Missing	Mean	SD	Median	Min	Max
	Temperature (Celsius)							
TempC	Current	Numeric	662	6.91	7.77	6.40	−24.70	37.10
TempMinD	Minimum Day	Numeric	1	4.12	7.73	3.90	−33.90	31.90
TempD	Daily	Numeric	192	6.97	7.69	6.40	−25.30	33.10
TempMaxD	Maximum Day	Numeric	1	9.29	8.12	8.20	−21.90	40.70
TMinLN	Min. Low Normal	Numeric	329	−1.29	8.81	−1.30	−39	29.70
TMinMN	Min. Mean Normal	Numeric	330	4.95	6.65	4.20	−15.60	31.70
TMinHN	Min. High Normal	Numeric	329	10.38	6.03	9.30	−6.40	42.80
TMaxLN	Max. Low Normal	Numeric	330	3.63	8.22	3.10	−31.30	38.30

(continued on next page)

Table 7 (*continued*)

Feature Name	Description	Var Type	Missing	Mean	SD	Median	Min	Max
TMaxMN	Max. Mean Normal	Numeric	329	9.51	7.06	8.40	−9.20	42.50
TMaxHN	Max. High Normal	Numeric	329	15.83	7.83	14.10	−1.60	48.10
Temp diff C D	Difference C and D	Numeric	660	0.00	1.92	0	−11.10	15
TempCDewC	Difference C and DewC	Numeric	675	3.80	3.24	3	−3.10	40.20
	Dew point temperature (Celsius)							
DewC	Current	Numeric	675	3.10	7.62	3.10	−41	27.60
DewD	Daily	Numeric	202	3.13	7.58	3	−30.10	26.80
	Wind speed (knots)							
WspC	Current	Numeric	674	19.04	13.87	15.90	0	130
WspD	Daily (max. hourly)	Numeric	194	31.93	16.04	28.60	0	137
WspMinD	Minimum Day	Numeric	0	7.9	7.6	5.40	0	85.6
WspMeanD	Mean Day	Numeric	4	15.08	11.08	12.70	0	109
WspMaxD	Max Day	Numeric	0	30.65	16.87	27.70	0	137
WspLN	Low Normal	Numeric	337	12.73	6.16	12.60	0	83.20
WspMN	Mean Normal	Numeric	337	30.44	9.38	29.50	4	83.20
WspHN	High Normal	Numeric	337	57.59	20.78	55.40	5.40	289
	Wind direction (degrees)							
WdirC	Current	Numeric	808	184	95.75	185	0	360
WdirD	Daily	Numeric	198	187	93.72	191	0	360
Wdir diff C D	Difference C and D	Numeric	807	33.48	37.69	19.20	0	180
	Sea level pressure (millibars)							
PressC	Current	Numeric	2651	969	199	1010	955	1066
PressD	Daily	Numeric	606	1010	12.61	1011	960	1046
Press diff C D	Difference C and D	Numeric	2841	−0.18	3.18	0	−27.50	61.60
	Relative humidity (%)							
HumC	Current	Numeric	675	78.42	15.11	81.10	0	100
HumD	Daily	Numeric	202	78.26	11.37	79.90	18.50	100
HumLN	Low Normal	Numeric	341	55.26	11.32	55.80	0.20	96.60
HumMN	Mean Normal	Numeric	341	77.18	5.17	77.30	4.90	97.60
HumHN	High Normal	Numeric	335	92.38	5.28	92.70	0.35	100
	Cloud cover (%)							
CcoverC	Current	Numeric	1253	65.24	30.62	77.60	0	100
CcoverD	Daily	Numeric	233	64.72	22.90	68.90	0	100
CcoverLN	Low Normal	Numeric	381	13.91	15.93	8	0	100
CcoverMN	Mean Normal	Numeric	381	58.64	17.95	64.40	0	100
CcoverHN	High Normal	Numeric	381	91.35	10.86	92.60	0	100
Ccover diff C D	Difference C and D	Numeric	1254	−0.17	21.31	1.50	−99.60	82.10
	Visibility (kilometres)							
VisC	Current	Numeric	1445	15.94	14.11	11	0	75
VisD	Daily	Numeric	348	17.11	11.15	13.60	0	75
Vis diff C D	Difference C and D	Numeric	1446	−1	56.56	−6.78	−100	468
	Precipitation (mm)							
PrecC	Current	Numeric	2390	0.15	0.98	0	0	23.40
PrecD	Daily	Numeric	1428	2.51	7.79	0	0	198
PrecLN	Low Normal	Numeric	1962	0.07	4.75	0	0	350
PrecMN	Mean Normal	Numeric	1963	3.65	3.38	2.70	0	28
PrecHN	High Normal	Numeric	1962	24.70	27.13	17.30	0	487
	Wind speed							
				Over	Under			
WspC OorU D	Current over or under daily	Categorical	826	4450	2105			
	Moon phase							
MphaseC	0 = new moon 0.5 = full moon 1 = next new moon	Numeric	241	0.49	0.31	0.50	0	1
	Sun							
			Missing	Up	Down			

(*continued on next page*)

Table 7 (*continued*)

Feature Name	Description	Var Type	Missing	Mean	SD	Median	Min	Max				
Sun UpOrDwn	Up or down during accident Icon	Categorical	909	3279	3193							
IconC	A fixed summary of current weather	Categorical	Missing	Partly cloudy day	Partly cloudy night	Cloudy (>90 % c.cover)	Rain	Clear night	Clear day	Wind	Snow	Fog (<1 km vis.)
			662	1992	1703	861	647	617	573	224	68	34

C=Current. Indicates the time of accident.
D=Daily. Daily average values.
N=Normal. Statistical values for the time and location. TMinLN referes to the lowest minimum temperature of the statistical period. TMaxMN referes to the statistically mean maximum temperature of the day.

Appendix E. Details of top 30 model

Table 8
A list of 30 trained ML models (sorted by Sample%, full list of 105 models provided as supplementary material).

No.	Model Type	Feature List	Sample Size	Sample %	Holdout Size	Accuracy (1)	Accuracy (2)
1	Light Gradient Boosted Trees Classifier with Early Stopping (SoftMax Loss) (16 leaves)	standard var., ex. weather	9226	100.00	1845	65.06 %	64.86 %
2	RandomForest Classifier (Gini)	standard + weather boosted	9226	100.00	1845	68.31 %	67.59 %
3	RandomForest Classifier (Gini)	Only weather, boosted	9226	100.00	1845	63.10 %	63.54 %
4	**Light Gradient Boosted Trees Classifier with Early Stopping (SoftMax Loss) (64 leaves)**	**standard + weather boosted**	**9226**	**100.00**	**1845**	**71.43 %**	**70.23 %**
5	Light Gradient Boosted Trees Classifier with Early Stopping (SoftMax Loss) (64 leaves)	standard + weather boosted	7381	80.00	1845	68.59 %	68.15 %
6	Light Gradient Boosted Trees Classifier with Early Stopping (SoftMax Loss) (16 leaves)	standard var., ex. weather	7381	80.00	1845	65.40 %	65.00 %
7	RandomForest Classifier (Gini)	standard + weather boosted	7380	79.99	1845	66.35 %	66.62 %
8	Light Gradient Boosted Trees Classifier with Early Stopping (SoftMax Loss) (16 leaves)	standard + weather boosted	5904	63.99	1845	66.76 %	66.43 %
9	Keras Deep Residual Neural Network Classifier using Training Schedule (3 Layers: 512, 64, 64 Units)	Only weather, boosted	5904	63.99	1845	53.83 %	54.00 %
10	Majority Class Classifier	Only weather, boosted	5904	63.99	1845	51.86 %	
11	Gradient Boosted Trees Classifier with Early Stopping	standard var., ex. weather	5904	63.99	1845	64.05 %	64.10 %
12	eXtreme Gradient Boosted Trees Classifier with Early Stopping	Only weather, boosted	5904	63.99	1845	57.01 %	57.66 %
13	Keras Residual Neural Factorization Machine Classifier using Training Schedule (2 Layers: 96, 96 Units)	Only weather, boosted	5904	63.99	1845	52.20 %	
14	eXtreme Gradient Boosted Trees Classifier with Early Stopping	standard + weather boosted	5904	63.99	1845	67.10 %	66.93 %
15	Keras Residual Cross Network Classifier using Training Schedule (3 Cross Layers, 4 Layers: 96, 96, 72, 72 Units)	Only weather, boosted	5904	63.99	1845	52.74 %	53.57 %
16	Stochastic Gradient Descent Classifier	Only weather, boosted	5904	63.99	1845	51.52 %	
17	Light Gradient Boosted Trees Classifier with Early Stopping (SoftMax Loss) (64 leaves)	Only weather, boosted	5904	63.99	1845	61.07 %	60.94 %
18	Stochastic Gradient Descent Classifier	Only weather, boosted	5904	63.99	1845	53.42 %	53.42 %
19	RandomForest Classifier (Gini)	Only weather, boosted	5904	63.99	1845	61.95 %	61.88 %
20	Stochastic Gradient Descent Classifier	standard var., ex. weather	5904	63.99	1845	64.05 %	63.57 %
21	Stochastic Gradient Descent Classifier	Only weather, boosted	5904	63.99	1845	53.35 %	
22	Keras Residual AutoInt Classifier using Training Schedule (2 Attention Layers with 2 Heads, 2 Layers: 96, 96 Units)	Only weather, boosted	5904	63.99	1845	53.69 %	53.22 %

(continued on next page)

Table 8 (*continued*)

No.	Model Type	Feature List	Sample Size	Sample %	Holdout Size	Accuracy (1)	Accuracy (2)
23	RandomForest Classifier (Gini)	standard + weather boosted	5904	63.99	1845	67.57 %	66.75 %
24	Keras Deep Residual Neural Network Classifier using Training Schedule (2 Layers: 512, 512 Units)	Only weather, boosted	5904	63.99	1845	55.92 %	
25	Decision Tree Classifier (Gini)	Only weather, boosted	5904	63.99	1845	52.54 %	
26	Keras Deep Self-Normalizing Residual Neural Network Classifier using Training Schedule (3 Layers: 256, 128, 64 Units)	Only weather, boosted	5904	63.99	1845	53.49 %	53.33 %
27	Gradient Boosted Trees Classifier with Early Stopping	Only weather, boosted	5904	63.99	1845	56.80 %	57.01 %
28	Gradient Boosted Greedy Trees Classifier with Early Stopping	Only weather, boosted	5904	63.99	1845	56.20 %	55.94 %
29	Gradient Boosted Trees Classifier with Early Stopping	standard + weather boosted	5904	63.99	1845	66.89 %	66.05 %
30	RandomForest Classifier (Gini)	standard var., ex. weather	5904	63.99	1845	64.18 %	64.61 %

Accuracy(1) and Accuracy(2) percentages indicate the range of accuracy across validation, cross validation, and holdout samples.

Appendix F. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trd.2024.104388.

References

Adland, R., Jia, H., Lode, T., Skontorp, J., 2021. The value of meteorological data in marine risk assessment. Reliab. Eng. Syst. Saf. 209, 107480 https://doi.org/10.1016/j.ress.2021.107480.

Atak, Ü., Arslanoğlu, Y., 2022. Machine learning methods for predicting marine port accidents: a case study in container terminal. Ships Off. Struct. 17 (11), 2480–2487. https://doi.org/10.1080/17445302.2021.2003067.

Bye, R.J., Aalberg, A.L., 2018. Maritime navigation accidents and risk indicators: an exploratory statistical analysis using AIS data and accident reports. Reliab. Eng. Syst. Saf. 176, 174–186. https://doi.org/10.1016/j.ress.2018.03.033.

DataRobot, 2022a. Exploratory Data Analysis: DataRobot Docs. https://app.eu.datarobot.com/docs/data/analyze-data/eda-explained.html.

DataRobot, 2022b. Light Gradient Boosted Trees Classifier with Early Stopping Documentation: DataRobot Model Documentation. https://app.eu.datarobot.com/model-docs/tasks/ESLGBMTC-Light-Gradient-Boosted-Trees-Classifier-with-Early-Stopping.html.

DataRobot, 2023a. Optimization metrics: DataRobot Docs. https://app.eu.datarobot.com/docs/modeling/analyze-models/evaluate/multiclass.html.

DataRobot, 2023c. Metrics explained: DataRobot Docs. https://app.eu.datarobot.com/docs/modeling/analyze-models/evaluate/roc-curve-tab/metrics.html.

DataRobot, 2023d. Basic model workflow: DataRobot Docs. https://app.eu.datarobot.com/docs/modeling/build-models/build-basic/model-data.html.

DataRobot, 2023e. Data partitioning and validation: DataRobot Docs. https://app.eu.datarobot.com/docs/modeling/reference/model-detail/data-partitioning.html.

DataRobot, 2023b. Feature impact: DataRobot Docs. https://app.eu.datarobot.com/docs/modeling/analyze-models/understand/feature-impact.html.

EMSA, 2022. Annual overview of marine casualties and incidents. European Maritime Safety Agency. https://www.emsa.europa.eu/newsroom/latest-news/item/4867-annual-overview-of-marine-casualties-and-incidents-2021.html.

Fan, S., Zhang, J., Blanco-Davis, E., Yang, Z., Yan, X., 2020. Maritime accident prevention strategy formulation from a human factor perspective using Bayesian Networks and TOPSIS. Ocean Eng. 210, 107544 https://doi.org/10.1016/j.oceaneng.2020.107544.

Forskrift om melde- og rapporteringsplikt til sjøs. (2008). *Forskrift om melde- og rapporteringsplikt ved sjøulykker og andre hendelser til sjøs* (FOR-2008-06-27-744). Lovdata. https://lovdata.no/dokument/SF/forskrift/2008-06-27-744.

Freund, Y., Schapire, R.E., 1995. A desicion-theoretic generalization of on-line learning and an application to boosting. In: Vitányi, P. (eds) Computational Learning Theory. EuroCOLT 1995. Lecture Notes in Computer Science, vol 904. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-59119-2_166.

Gu, B., Liu, J., 2024. A systematic review of resilience in the maritime transport. Int J Log Res Appl. https://doi.org/10.1080/13675567.2023.2165051.

Gu, B., Liu, J., Chen, J., 2023. Scenario-based strategies evaluation for the maritime supply chain resilience. Transp. Res. Part D: Transp. Environ. 124, 103948.

He, X., Zhao, K., Chu, X., 2021. AutoML: A survey of the state-of-the-art. Knowl.-Based Syst. 212, 106622 https://doi.org/10.1016/j.knosys.2020.106622.

Hutter, F., Kotthoff, L., Vanschoren, J., 2019. Automated machine learning: methods, systems, challenges. Springer Nature, p. 219. https://doi.org/10.1007/978-3-030-05318-5.

Knapp, S., Kumar, S., Sakurada, Y., Shen, J., 2011. Econometric analysis of the changing effects in wind strength and significant wave height on the probability of casualty in shipping. Accid. Anal. Prev. 43 (3), 1252–1266. https://doi.org/10.1016/j.aap.2011.01.008.

Kretschmann, L., 2020. Leading indicators and maritime safety: predicting future risk with a machine learning approach. J. Ship. Trade 5 (1), 19. https://doi.org/10.1186/s41072-020-00071-1.

Li, W., Asadabadi, A., Miller-Hooks, E., 2022. Enhancing resilience through port coalitions in maritime freight networks. Transp. Res. A Policy Pract. 157, 1–23.

Liu, K., Yu, Q., Yuan, Z., Yang, Z., Shu, Y., 2021. A systematic analysis for maritime accidents causation in Chinese coastal waters using machine learning approaches. Ocean Coast. Manage. 213, 105859 https://doi.org/10.1016/j.ocecoaman.2021.105859.

Luo, M., Shin, S.H., 2019. Half-century research developments in maritime accidents: Future directions. Accid. Anal. Prev. 123, 448–460. https://doi.org/10.1016/j.aap.2016.04.010.

Merrick, J.R., Dorsey, C.A., Wang, B., Grabowski, M., Harrald, J.R., 2022. Measuring prediction accuracy in a maritime accident warning system. Prod. Oper. Manage. 31 (2), 819–827. https://doi.org/10.1111/poms.13581.

Munim, Z.H., Sørli, M.A., Kim, H., Alon, I., 2024. Predicting maritime accident risk using automated machine learning. Reliab. Eng. Syst. Saf. 248, 110148 https://doi.org/10.1016/j.ress.2024.110148.

Murphy, K.P., 2012. Machine learning: A probabilistic perspective. MIT press. Cambridge, Massachusetts.

Rawson, A., Brito, M., Sabeur, Z., Tran-Thanh, L., 2021. A machine learning approach for monitoring ship safety in extreme weather events. Saf. Sci. 141, 105336 https://doi.org/10.1016/j.ssci.2021.105336.

Rawson, A., Brito, M., 2023. A survey of the opportunities and challenges of supervised machine learning in maritime risk analysis. Transp. Rev. 43 (1), 108–130.

Rezaee, S., Pelot, R., Finnis, J., 2016a. The effect of extratropical cyclone weather conditions on fishing vessel incidents' severity level in Atlantic Canada. Saf. Sci. 85, 33–40. https://doi.org/10.1016/j.ssci.2015.12.006.

Rezaee, S., Pelot, R., Ghasemi, A., 2016b. The effect of extreme weather conditions on commercial fishing activities and vessel incidents in Atlantic Canada. Ocean Coast. Manage. 130, 115–127. https://doi.org/10.1016/j.ocecoaman.2016.05.011.

Theofilatos, A., Chen, C., Antoniou, C., 2019. Comparing machine learning and deep learning methods for real-time crash prediction. Transp. Res. Rec. 2673 (8), 169–178.

Ugurlu, H., Cicek, I., 2022. Analysis and assessment of ship collision accidents using fault tree and multiple correspondence analysis. Ocean Eng. 245, 110514 https://doi.org/10.1016/j.oceaneng.2021.110514.

Uğurlu, Ö., Yıldız, S., Loughney, S., Wang, J., 2018. Modified human factor analysis and classification system for passenger vessel accidents (HFACS-PV). Ocean Eng. 161, 47–61. https://doi.org/10.1016/j.oceaneng.2018.04.086.

United States Department of Transportation. (2022). Aviation Weather Handbook. https://www.faa.gov/sites/faa.gov/files/FAA-H-8083-28_FAA_Web.pdf.

Uyanık, T., Karatuğ, Ç., Arslanoğlu, Y., 2021. Machine learning based visibility estimation to ensure safer navigation in strait of Istanbul. Appl. Ocean Res. 112, 102693 https://doi.org/10.1016/j.apor.2021.102693.

Wang, N., Yuen, K.F., 2022. Resilience assessment of waterway transportation systems: combining system performance and recovery cost. Reliab. Eng. Syst. Saf. 226, 108673.

Yang, Y., Shao, Z., Hu, Y., Mei, Q., Pan, J., Song, R., Wang, P., 2022. Geographical spatial analysis and risk prediction based on machine learning for maritime traffic accidents: a case study of Fujian sea area. Ocean Eng. 266, 113106.

Yildiz, S., Uğurlu, Ö., Wang, J., Loughney, S., 2021. Application of the HFACS-PV approach for identification of human and organizational factors (HOFs) influencing marine accidents. Reliab. Eng. Syst. Saf. 208, 107395 https://doi.org/10.1016/j.ress.2020.107395.

Zhang, Z., Li, X.M., 2017. Global ship accidents and ocean swell-related sea states. Nat. Hazards Earth Syst. Sci. 17 (11), 2041–2051. https://doi.org/10.5194/nhess-17-2041-2017.