Reliability Engineering and System Safety 243 (2024) 109779

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

A data-driven risk assessment of Arctic maritime incidents: Using machine learning to predict incident types and identify risk factors

Rajesh Kandel, Hiba Baroud *

Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN 37235, USA

ARTICLE INFO

Dataset link: https://doi.org/10.18739/A2DJ58J37

Keywords:
Risk analysis
Machine learning
Arctic
Maritime navigation

ABSTRACT

The rise in Arctic temperatures has caused sea ice to melt, making the region more navigable through a new, shorter shipping route. Navigating the Arctic poses high risks due to extreme environmental conditions. This study proposes scalable data-driven predictive models to assess the risk of Arctic navigation under uncertain weather and sea ice conditions. Machine learning is applied to predict the type of Arctic incidents and identify their risk factors. Several models are investigated, and the most accurate model for the two Arctic routes, the North West Passage (NWP) and the Northern Sea Route (NSR), is determined based on several validation metrics. Random forest and Naïve Bayes models provide the best accuracy and F1-score for NWP and NSR, respectively. The wind speed, vessel type, length, and age are important risk factors for NWP while temperature at 2 m above the surface, vessel length, and age are important for NSR. Partial dependence plots are used to investigate the effect of each feature on predicting each incident type. Equipment failures are more common among newer and longer vessels. Collision related incidents are more likely to be predicted for longer vessels, while grounding related incidents are more frequent at higher air temperature.

1. Introduction

Arctic temperatures have increased immensely in recent years, leading to a reduction in sea ice [1]. As a result, accessibility for Arctic maritime navigation has increased. Between 2013 and 2019, the number of ships entering the Arctic increased by 25 percent [2]. The maritime shipping routes through the Arctic region present a unique opportunity to connect Europe, Asia, and North America. Currently, most of the trade between these regions takes place through Southern passages like the Suez Canal and Malacca Channel. Shipping through the Arctic reduces the transit distance between Northeast Asia and Europe by 24% compared to the route along Suez canal [3]. Thus, the Arctic shipping connection not only provides an alternative shipping route for trade between these global trade centers, it can also drastically reduce the transit time and fuel cost given the shorter route. An important consideration of Arctic maritime navigation lies in the risk of operating vessels in remote areas and harsh weather conditions.

While the Arctic maritime activity represents a small fraction of the global maritime traffic and incidents are rare, the impact of such incidents is significant and can pose substantial risk to the environment and Arctic communities. Ships stranded in the Arctic may need to be towed for days to safety, like the cruise ship Ortelius in 2016 with 146 people on board [4]. In addition, these incidents pose risks to coastal communities and ecosystems in the event of oil spills [5]. The future of Arctic shipping relies on the understanding and assessment of the risk of Arctic navigation and incidents. Specifically, there is a major impetus to determine the risk factors (e.g., climate, lack of infrastructure, vessel traffic, ship design) and to predict the economic, social, and environmental consequences of the increase in Arctic shipping activity. This study focuses on the prediction of Arctic incident types and their corresponding risk factors. Research advances in this area have developed approaches that primarily use probabilistic methods (e.g., Bayesian networks) to quantify the risk of Arctic incidents [6–9]. While these approaches are helpful to assess the causes for incidents, they are prone to biases in their prior assumption and present limitations that challenge their generalization for future scenario analysis. Data-driven methods can help overcome these gaps by leveraging data on historical incidents and weather conditions to predict incidents and identify risk factors. Specifically, machine learning models are proposed in this study because their structure can be generalized to explore the risk of incidents across the entire Arctic region and under various environmental conditions. For example, the proposed approach can be adapted to unlock future iterations of data-driven risk prediction for Arctic incidents by incorporating projected climate scenarios and future Arctic infrastructure development in the input variables space.

The potential for increased Arctic maritime navigation exists along two routes; the North West Passage (NWP) along the north Canadian

* Corresponding author.
E-mail address: hiba.baroud@vanderbilt.edu (H. Baroud).

https://doi.org/10.1016/j.ress.2023.109779
Received 16 January 2023; Received in revised form 28 September 2023; Accepted 2 November 2023
Available online 20 November 2023

and Alaskan waters, and the Northern Sea Route (NSR) along northern Russian waters. A study comparing Arctic shipping routes with more traditional routes like the Suez Canal found that NSR has the potential to save as much as 10 days for a single transit activity due to the reduction in sailing distance [10]. As a result, there has been a marked increase in maritime traffic in the region, the majority of which are supply, research, and survey vessels, followed by fishing, cargo, tanker, and passenger vessels [11]. However, the region's remoteness can lead to limited access to navigation tools and potential shortages of emergency response equipment. As a result, the Arctic shipping routes are still relatively unexplored and historical data on Arctic incidents are limited. Prior studies in the risk assessment of Arctic maritime incidents rely heavily on expert opinion [7,12,13] due to the lack of a comprehensive database. To overcome this limitation, a database is developed in this study that merges historical Arctic maritime incidents with environmental conditions to identify specific patterns in incidents and their corresponding risk factors.

The research objectives of this study are to (1) develop and validate classification techniques to predict the likelihood of different types of Arctic incidents, and (2) identify significant risk factors that contribute to different types of incidents. The proposed approach starts with an exploratory data analysis to identify patterns in historical incident and environmental conditions data. The results of the exploratory analysis are used to inform the input features selection, clustering of incident records by routes, and merging of incident classes during the data preparation for machine learning. Then, multiple machine learning models for multi-class classification are investigated based on validation metrics like accuracy, precision, recall, and F-1 score. The best performing model is identified through a comparison of these metrics. Finally, the machine learning output is analyzed using feature importance and partial dependence plots to investigate the relationship between the features and the type of incidents and identify risk factors that influence the occurrence of different types of incidents. The proposed data-driven risk assessment can inform strategic decisions that help improve Arctic navigation safety through risk-informed resource allocation and voyage planning to prevent incidents and reduce the consequences of different incident scenarios.

2. Background

Arctic navigation has been a subject of considerable interest to researchers recently. As there are more databases keeping track of and publishing the voyage records in the Arctic, there have also been more studies on the analysis of shipping patterns in the Arctic. Studies on spatial and temporal variations in Arctic shipping [11,14] have sought to analyze the changes in shipping patterns in the Arctic Region, closely examining the trends in the Northeast (NSR) and Northwest (NWP) passages. In particular, the NSR has been the focus of such trend studies [15–17] owing to surges in dry cargo as well as oil transport between Europe and Kara Sea driven by the development and marine transportation of Russian Arctic natural resources. On the NWP, Pizzolato et al. [18] have explored the spatial relationship between shipping activity and sea ice concentration between 1990 and 2015 to study the influence of sea ice changes on shipping activity in the Canadian Arctic. These studies have helped understand the trends in Arctic navigation by breaking down the traffic information spatially and temporally based on different characteristics like vessel types, vessel tonnage, and flag state of vessels, among others. However, the risk of Arctic navigation (e.g., likelihood and impact of incidents) has been less studied.

In the analysis of Arctic navigation risk, the lack of data on historical incidents in the Arctic has led researchers to focus on qualitative approaches and expert elicitation to conduct risk assessments. For instance, Marchenko et al. [4] have used a risk matrix that incorporates qualitative data collected from industry specialists and researchers to categorize the risks to vessels in different parts of the Atlantic Arctic. More recently, probabilistic risk analysis models for global and Arctic maritime navigation have been developed using Bayesian networks [8,9,12,13,19]. A study analyzing the causes of grounding incidents in the Arctic using a Bayesian network (BN) [8] looked at global incident records that only include 5 incidents in the Arctic waters. Li et al. [7] use Arctic vessel traffic and environmental data to develop a decision support model for Arctic ship navigation using a dynamic BN where the risk factors are identified by experts and not based on historical incident records. Li et al. [9] use global incident records from the International Maritime Organization (IMO) to develop a BN analysis of the risk factors where the human risk factors are extracted from the accident reports of a portion of the records in conjunction with expert opinions. Similarly, a BN model for ship besetting in ice along the NSR [6] leverages expert opinions to develop a list of the risk factors, and tests the model for a single voyage through the NSR. These studies have advanced probabilistic risk analysis of Arctic maritime navigation. However, the incident reports used to inform the BN models in these studies lack homogeneity in terms of the vessel characteristics or weather observations.

While Bayesian networks have proven to be valuable for assessing failure scenarios of a particular ship at a micro level by considering the different events that might lead to an incident, they have poor scalability and cannot be applied simultaneously between different voyages, geographic regions, and storm systems. These models are often used as a first step to envisage incidents and make decisions to mitigate the consequences [13]. Probabilistic methods using fault tree analysis [20] and fuzzy fault tree analysis [21,22] have also been used in Arctic navigation risk analysis. The use of Bayesian networks or qualitative risk models have significant uncertainties in deriving the priors and may be subject to bias by the contributing experts [23]. As a result, exploratory studies using machine learning approaches have been conducted to identify conditions associated with navigation incidents that could be used as risk indicators [24]. For example, a study on the global incident data obtained from the International Maritime Organization (IMO) uses machine learning techniques for classification of vessel traffic instances as incidents and non-incidents [25]. Similarly, Wang et al. [26] use logistic regression to analyze the risk factors affecting the severity of global maritime incidents. The use of machine learning can overcome the issue of scalability and adaptability of the models to different voyages and future climate scenarios. However, prior studies have either focused on a small geographic region within the Arctic and limited incident types [24,27] or have used global incident records to analyze the risk of maritime incidents globally and not specifically in the Arctic [25,26,28].

To better understand potential risk factors of maritime navigation across the Arctic, there is a need to develop models that can handle comprehensive data sets of historical Arctic incidents, vessel characteristics, and weather and sea ice conditions to examine relationships between all variables and draw insights on the drivers of different types of incidents. As more records on Arctic maritime incidents become available, risk analysis of Arctic navigation should incorporate machine learning approaches and in-depth exploratory analysis. As such, this study addresses the lack of data-driven methods in risk analysis approaches of Arctic maritime navigation. Specifically, the novel contribution lies in (1) the methodological approach for risk analysis which is founded in machine learning models and employs a comprehensive database covering incidents across the entire Arctic and corresponding weather and sea ice observations, and (2) the identification of Arctic maritime risk factors.

3. Method

The main purpose of this study is to develop machine learning models that leverage historical Arctic maritime incident records to better understand the risks of Arctic maritime navigation. The first step of the process is to gather sufficient information that is representative of spatial and temporal variations as well as environmental conditions in the Arctic. For this study, the incidents dataset from the Protection of Arctic Marine Environment (PAME) [29] is used and merged with historical weather records that correspond to the incident date and location. The final dataset contains incident records with the date and location, vessel characteristics, weather variables (e.g., wind, temperature) and sea ice variables. The response variable considers that an incident has occurred and evaluates the likelihood of a particular incident type.

First, an exploratory data analysis is performed to investigate developing trends in the Arctic incidents (rates by year and month) and help confirm intuitions about those incidents (e.g. number of incidents during different seasons and incidents of different vessel types). Second, different classification techniques are explored and applied to predict the likelihood of incident types. Four supervised machine learning techniques, namely Logistic Regression (LR), Naïve Bayes (NB) method, Support Vector Machine (SVM), and Random Forest (RF) classifier are used to predict the type of Arctic incidents. These specific methods are chosen based on the size of the database used in this research and to reflect varying levels of complexity that consider both linear and non-linear relationships between the variables in the data set. Logistic regression is a simple, fast, and easily interpretable method that uses linear relationships between features. Naive Bayes is another simple method that requires relatively less training data. Support vector classifier is comparatively more complex with a large number of hyperparameters and uses non-linear kernels to model the relationships between features. Random forest classifier is an ensemble method that can handle non-linearity and can deal with complex datasets. These models are chosen to cover a wide range of possible model structure and assumption. This allows an investigation of different classes of techniques to determine the best approach for this classification problem.

The performance of the models is evaluated using different predictive accuracy metrics. The outcomes of the best performing model are then further analyzed to evaluate the influence of different model features on the likelihood that an incident type occurs.

3.1. Models

The machine learning methods used in this study are briefly described in this section.

Logistic Regression (LR) Logistic regression utilizes a logistic function (also called logit function, shown in Eq. (1)) to model the probability (P) of a certain class or event [30]. In the most basic form, it is used to model dependent variables with two possible outcomes. Logistic regression is widely used in classification problems in various fields of study.

$$logit(P) = log(\frac{P}{1-P}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{1}$$

Naïve Bayes (NB) Classifier

Naïve Bayes classifier is a probabilistic classifier based on Bayes' Theorem [31] and assumes independence among the input features given the output class. Eq. (2) gives the class y with maximum probability, given the input values. This classification technique is preferred for datasets with limited number of training data points.

$$y = argmax_y P(y) \prod_{i=1}^{n} P(x_i|y) \tag{2}$$

Support Vector Machine (SVM)

Support vector machine classification uses a hyperplane or set of hyperplanes, linear or non-linear, constructed in the feature space to categorize the datapoints based on input features [32]. The best hyperplane is chosen with the largest separation between the classes of data and minimum error in the classification. Fig. 1 shows a linear hyperplane separating observations in a two-feature space.

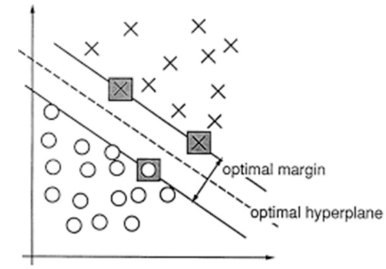


Fig. 1. Representation of a linear hyperplane in Support Vector Machine.

Random Forest (RF)

Random forest classification [33] involves constructing multiple decision trees using the training data and classifying based on the class selected by the most trees. The forest of decision trees is generated using techniques like bagging and bootstrapping. Random forest is one of the newer and most popular classification techniques in machine learning.

The SVM and RF classifiers have hyperparameters that can be chosen to optimize the model performance. Grid search is implemented for these models to determine the hyperparameters to be used. For SVM, multiple values of C (regularization parameter), Gamma, and kernel type are tested. Similarly, a number of options for the maximum depth of tree and number of trees are tested to determine the best RF model. The incident types (classes) have a considerable imbalance in the number of observations with one type of incidents, *equipment failure*, accounting for approximately half of the records. Therefore, Synthetic Minority Oversampling Technique (SMOTE) [34] is used to synthesize observations for minority classes in the training data to improve prediction.

3.2. Model performance

To evaluate the performance of each model in predicting the type of Arctic incident, we use a validation approach. First, the complete incidents dataset is divided into training and test sets using an 80/20 split and the confusion matrix for classification (general representation in Table A.1 in the appendix) of the test set is obtained. The training data is normalized before it is used to train the models. Oversampling is used on the training data to account for the imbalance in the number of observations in each class. Separately normalized test input data is used to obtain classification results from the trained models. The models are tuned with different sets of input variables and hyper-parameters (where applicable) to optimize performance for each modeling technique. A backward selection strategy is used to determine the best performing set of features for each model starting with the model with the full set of input features. The feature with the lowest permutation feature importance is selected at each step and the model performance is compared between models with and without that feature. If removing the feature improves the model performance, it is removed from the set of input features and the step is repeated for the next feature with the least importance. Additionally, a 5-fold cross-validation is performed on the entire normalized dataset and average values of several performance metrics are obtained for each model. The performance metrics are summarized in Table 1. Finally, the feature importance is calculated for each variable in the models to evaluate their significance in predicting the type of incidents.

The models are analyzed for the importance of the input features using permutation feature importance and partial dependence plots. Permutation feature importance gives a measure of how much a model depends on a certain feature by measuring the decrease in model evaluation score from randomly shuffling a feature value. The partial dependence plots give the probability of predicting a particular target

Table 1
Description and formula for the evaluation metrics of classification models.

Metric	Definition	Formula
Accuracy	Fraction of all data that are classified correctly	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision / Positive Predictive Value (PPV)	Fraction of positive classifications that are classified correctly	$\frac{TP}{TP+FP}$
Recall/ Sensitivity	Fraction of positive observations that are classified correctly	$\frac{TP}{TP+FN}$
F-1 score	Harmonic mean of precision and recall	$\frac{2\times Precision\times Recall}{Precision+Recall}$

class at different values of the individual predictors. This gives an idea of how the predictions might change when the values of particular features in the model change, and thus provides insights on the effects of different features for particular outcomes.

Performing the proposed data-driven risk assessment of Arctic maritime incidents requires an integrated data that combines incident types with features that describe the potential risk factors (e.g., ship design, weather, and sea ice). The following section describes the data sources and processing to develop the comprehensive database.

4. Data

4.1. Data sources

Data from two sources are used to develop a new database and address the gap in data-driven risk analysis of Arctic maritime incidents. The first data set contains incident records with information on the date, location, vessel type, incident type, and vessel length between the years 2005 and 2017. The second data set contains weather variables extracted for the same period of analysis (2005–2017) and filtered corresponding to the location and the date of the incidents. The final data set used in the analysis [35] integrates information on incidents records, vessel characteristics, and environmental conditions. Data is collected from different sources and merged across common variables. Each data source is briefly described in this section.

Protection of Arctic Marine Environment (PAME) incident data

The Collection of Arctic Ship Accidents (CASA) dataset obtained from PAME (referred to hereon as PAME incident dataset) is chosen for this analysis because of the coherence of the data and its focus on the Arctic region. The PAME incident dataset contains records on Arctic incidents for a 13-year period in a standard format collected from six arctic states: Canada, Russia, Iceland, Denmark, Norway, and the United States. The PAME incident dataset [29] contains 5004 data points for incidents between January 2005 and December 2017. The original dataset has 24 features describing the incidents. The descriptions of relevant features are provided in Table 2. The data points are classified by PAME based on the location of occurrence as either above or below 58-degree latitude. This distinction helps differentiate Arctic incidents from the other events. There are 2638 incident data points above 58-degree latitude on which exploratory analysis is performed in this study.

Climate data

Data describing the environmental conditions is obtained from the ERA5 global reanalysis data and referred to as climate data in the literature [36]. The climate data contains weather variables describing wind, temperature, and sea ice, among others. Daily observations for six weather variables are extracted for the corresponding locations and dates of the incidents. Since the time of incident is not available in the incident data, the first observation of the weather variables on the day and location of the incidents are considered. These variables are listed in Table 3.

4.2. Exploratory data analysis

4.2.1. Incident data

The PAME incident data contains records of vessel traffic attributes for the incidents. The dataset is filtered for incidents at or above 58 degree latitude to only include incidents in Arctic conditions. An exploratory analysis is performed on the incident data to investigate the emerging patterns in Arctic incidents based on the location, vessel type, incident type, year, and month of incidents. The incident records are broken down into incident counts corresponding to different categories within these features and the findings are presented in this section.

Location

The 56 zones of the incidents classified in the original dataset are classified into seven general regions: Atlantic Area, Central Region, Inland Russia, North Atlantic Ocean, Norwegian Sea, Pacific Area, and Russian Arctic. The Pacific Area has the highest share of the incidents with 1872 data points. Russian Arctic region has 170 incidents and the Central Region that includes the Canadian Arctic has 86 incidents. The Atlantic and North Atlantic areas have a total of 58 incidents in the report.

Vessel Types

The vessels involved in the incidents are classified into 94 types in the original dataset. These classifications are condensed into seven general vessel types. Fishing vessels are the most common type of vessels involved in Arctic incidents, with 1049 occurrences. Passenger ships are involved in 601 incidents, Service vessels in 275, Cargo ships in 218, Recreational crafts in 151, Tankers in 127 and Icebreakers in further 23 incidents. Fig. 2(b) shows the number of incidents for different vessel types.

Incident Type

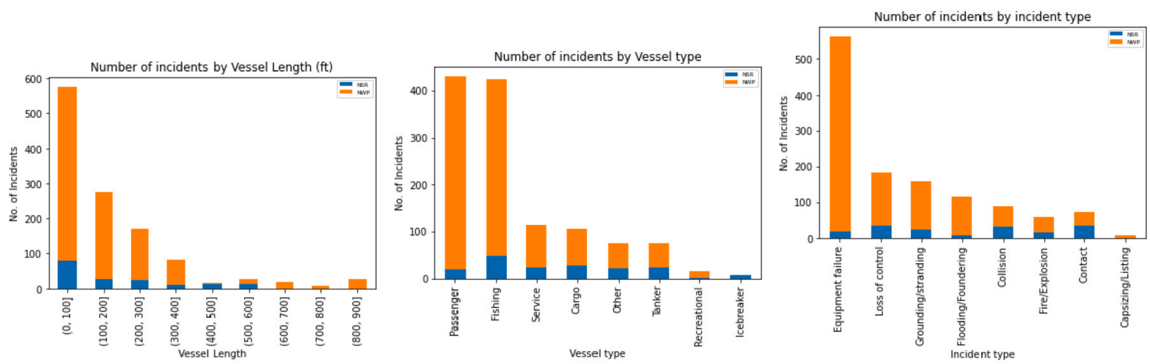
The dataset contains 81 different incident types which are reclassified into 10 general incident types. Equipment failure is the most prevalent type with 620 instances, followed by Loss of control with 334, Grounding/Stranding with 300, Collision with 115, Contact with 110, and Fire/Explosion with 97 occurrences. Other incident types are Damage to ship/equipment, Capsizing/Listing, and Non-incidental event. Fig. 2(c) shows the number of incidents by the type of incident.

Year and Month of Incidents

To understand the trend of incidents over time, we evaluate the rate of incidents for each month over four years between 2013 and 2017. The incident rates are only obtained for these years as the traffic records corresponding to the study period are only available for the years 2013–2017. The incident rates are calculated using the monthly vessel counts in the Arctic obtained from the PAME Automatic Identification System (AIS) database and the corresponding incident records. The AIS data provides a unique ID to each ship traveling each month. To calculate the incident rate, the number of unique ship IDs is considered to be the monthly vessel count. The number of incidents in a given month divided by this number gives the monthly incident rate. For yearly

Table 2
PAME data features.

Feature	Description
Data Source	Country that provided the data
Incident Date	Date of the incident
Flag State	State of the vessel involved in the incident
Incident Country/ Jurisdiction	Country that controls the incident location
General Incident Location	General area of the incident location
Lat58plus	Whether or not the latitude is above 58 degrees
Latitude	The latitude of the incident location
Longitude	The longitude of the incident location
Vessel Type	Type of the ship involved based on its use
Vessel name	Name of the vessel involved
Vessel Tonnage	The cargo carrying capacity of the vessel involved
Vessel Length	Length (in ft) of the vessel involved
Vessel Age	Age (in years) of the vessel involved
Consequence of Incident	General consequence of the incident
Incident Type	Description of the nature of incident



(a) Number of incidents by *Vessel Length* (b) Number of incidents by *Vessel type* (c) Number of incidents by *Incident Type*

Fig. 2. Number of incidents broken down by vessel length, vessel type, and incident type.

Table 3
Climate data features.

Feature	Unit
Wind speed	m/s
Surface pressure	Pa
Air temperature at 2 meters above surface	K
Sea surface temperature	K
Column water vapor amount	kg m-2
Sea ice cover	0–1 (value)

rates, the number of incidents in a year is divided by the sum of all monthly vessel counts over that year.

Over the entire Arctic, the summer months have the highest average rate of incident by month (Fig. 3(a)), peaking in June (0.41%) and declining through December (0.10%). This rate remains between 0.1% and 0.2% during the winter months and picks up from May (0.27%). When broken down by the two shipping routes, NSR and NWP, two different trends are observed. The NWP data follows a similar trend to the overall Arctic data with highest rate in June (1.26%) and lowest in December (0.31%). The rates for NSR are much lower and consistent around 0.20%, with the highest rate also observed in June (0.31%) and lowest in December (0.12%). As the number of vessels are marginally higher in the NSR compared to NWP, the NSR rate tempers the overall trend which agrees with the NWP rates. Even after accounting for the exposure in the rate of incidents by considering the ratio of incidents to the number of unique voyages, the average incident rates are lower for the winter months.

From 2013 to 2017, there is a steady increase (from 0.16% to 0.26%) in the yearly rate of incidents overall (Fig. 3(b)). This trend is also observed for the NSR (from 0.04% to 0.23%) but the rates for the NWP are steady at around the 0.25%. As a result, the overall yearly rates follow the dominant trend of NSR but are tempered by the NWP values. The total yearly vessel counts are obtained by aggregating the number of vessels for each month. The limitation in calculating the yearly rates this way is that there is a possibility of double counting the vessels navigating over multiple months as the ship IDs reset every month and unique vessel IDs are not available.

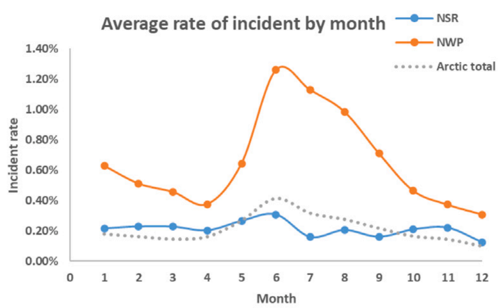
4.2.2. Comparison between NWP and NSR incidents

The two existing routes in the Arctic, NWP and NSR, present different trends and patterns of incidents. The NWP has 86% of the total incidents in the period of study between 2005 and 2017 while NSR has just 14% of the recorded incidents over the same period. This is in contrast with the number of voyages in these regions with NSR accounting for 56% of the voyages between 2013 and 2017.

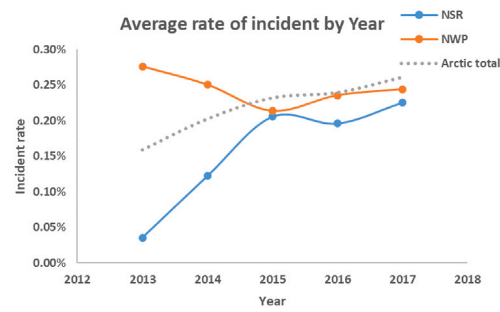
For the NWP, passenger vessel incidents are the most frequent, accounting for 38% of the records. Incidents involving fishing vessels are common for both NWP and NSR at approximately 35% and 28% of recorded incidents respectively. Incidents with service, cargo, and tanker vessels follow for the NWP. For the NSR incidents, additional frequent vessel types include service, cargo, and passenger vessels.

The most common type of incident in the NWP is Equipment failure, followed by Grounding, Loss of control, and Foundering. In comparison, Loss of control is the most frequent incident type in the NSR. Grounding is also the second most common type for NSR, followed by contact, collision and fire. While grounding incidents are common in both the Arctic routes, the order of other incident type frequencies are different.

This comparative analysis examines the difference between the trends in incidents in each of the Arctic routes and provides insights to explain the findings from the classification models. Figs. A.1 through A.5 in the Appendix provide a detailed illustration of how incidents are distributed between NWP and NSR. Based on this analysis, the modeling approach is applied to each route separately.



(a) Rate of incidents by *Month*



(b) Rate of incidents by *Year*

Fig. 3. Average incident rates broken down by month and year of incidents.

4.3. Data preparation for machine learning

First, the incident dataset is filtered to only include incidents in Arctic conditions. Following this, records that indicate an incident type named 'Consequences' are excluded from the analysis. This is done because that incident type does not provide information on the actual type of incident and cannot be appropriately predicted using the models. The resulting incident data is combined with the climate data to create a dataset that contains both traffic and weather variables. Pearson correlation coefficients and covariance matrix for the combined dataset are used to filter the variables to be included in the machine learning models. These variables are used as input features to train and evaluate the classification models. The traffic features in the PAME data include vessel location (Latitude and Longitude), vessel type, vessel length, vessel age, and month of incident. The weather features are wind speed (m/s), surface pressure (Pa), Air temperature at 2 meters above surface, Sea surface temperature, Column water vapor amount, and Sea ice concentration. The models are developed separately for NWP and NSR given the varying conditions, vessel activities, and incident types along each route. As such, the data is split based on the geographical location of the incident. The six most frequent incident types on record are included in the final data: Collision, Contact, Equipment failure, Flooding, Grounding, and Loss of control. However, there is imbalance in the number of observations between classes which can hamper the model performances. To address this issue, some classes are merged resulting in a total of three classes for the response variable. *Equipment failure* is kept as a separate class because of its high number of observations. The remaining classes are merged so that incident types that are similar in nature are grouped together and the number of incidents within each new class is fairly balanced. *Collision* and *contact* between ships are similar in nature and are often accompanied by *fire*, as reported in multiple cases of maritime collisions [37–39], and [40]. Therefore, they constitute the second class. Similarly, *loss of control*, *grounding*, and *flooding* are also known to occur in a sequence [41]. Loss of vessel control causes numerous accidental *grounding* [42,43]. Groundings may lead to hull breaches and result in *flooding* and listing of vessels [44,45]. Therefore, these three types are merged into a third class. Finally, the models are also applied to a subset of the data that includes incidents that have occurred in the open-sea (i.e., far from a port) to investigate the influence of weather conditions on the likelihood of different incident types.

5. Results

This section presents the performance metrics of the machine learning classification models and the subsequent analysis for feature importance and partial dependence of the models. The complete dataset used in the machine learning models contains 1249 incident records after data preparation considering all the features of interest. Given the differences in the patterns between the two routes that were observed in the exploratory analysis, the models are developed separately for the two regions using the NWP and NSR subsets of the complete dataset. The selected list of features for each model's best performance is provided in the Appendix (Tables A.2, A.3, and A.4).

The NWP dataset used in the machine learning models has 1075 incident records. The model performances for the NWP data (Table 4) also show that the random forest model is best suited to predict the incident types with an average accuracy and F-1 score of 0.58 and 0.450, respectively. The next best model is naive Bayes which utilizes three input features: Length of the vessel, age of the vessel and the latitude of the incident. The SVM model has a better accuracy but a lower F-1 score compared to the logistic regression model.

There are 174 incident records in the NSR dataset used for machine learning. The naive Bayes model performs the best for this dataset with average accuracy and F-1 score of 0.581 and 0.512, respectively (Table 5). The random forest model is the next best method, followed by the logistic regression and SVM models. The F-1 score of the random forest model is the lowest of all methods.

To better understand the influence of environmental conditions (i.e., weather and sea ice) on the occurrence of different incident types, we conduct the analysis for a subset of the data containing only incidents that have occurred in the open sea (i.e., far from a port). This subset contains 452 records for the entire Arctic. Further, we exclude features that are indicative of the location and time of the incident to investigate weather and sea ice features and their impact on the predictive accuracy. As such, this subset of the data is not divided between NSR and NWP as the incident location is excluded from the list of model input features. The random forest model is the best performing method with an accuracy of 0.488 and an F-1 score of 0.362, both of which are lower than models that include location-based features (Table 6). However, these models help examine the importance of the features describing environmental conditions.

Permutation feature importance is obtained for the models to assess the importance of the input features in the model predictions, as shown in Figs. 4, 5, and 6. For the random forest models using the NWP dataset, the most important features are the vessel type, age, month, longitude of incident location, wind speed, and vessel length. Among the weather variables, the wind speed is observed to be twice as important as the temperature at 2 m above surface. For the same model using NSR data, the importance of the temperature variable is much higher than that of wind speed. For the open sea data (which is a smaller subset of the data including incidents from both NSR and NWP), the best performing models are random forest and naive Bayes. Both these models have wind speed as the most important feature, followed by vessel length and temperature at 2 m above surface. Note that these results match closely those of NWP because a higher number of open sea incidents occur in the NWP region.

Table 4
Model performance based on cross-validation of all classification methods for the NWP incidents.

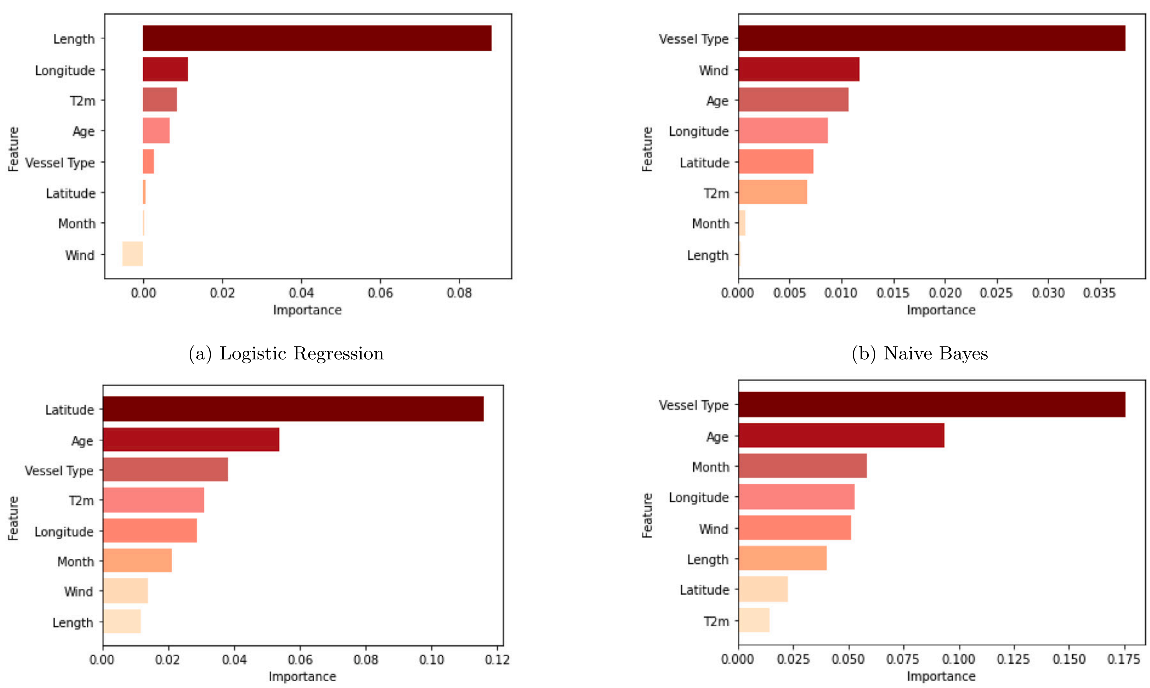
Classification model	Accuracy	Precision (macro)	Recall (macro)	F-1 score (macro)
Logistic Regression	0.511	0.405	0.419	0.398
Naïve Bayes	0.574	0.372	0.426	0.395
Support Vector Machine	0.516	0.390	0.414	0.364
Random Forest	**0.580**	**0.476**	**0.457**	**0.450**

Table 5
Model performance based on cross-validation of all classification methods for the NSR incidents.

Classification model	Accuracy	Precision (macro)	Recall (macro)	F-1 score (macro)
Logistic Regression	0.477	0.466	0.505	0.436
Naïve Bayes	**0.598**	**0.559**	**0.517**	**0.514**
Support Vector Machine	0.442	0.445	0.494	0.412
Random Forest	0.505	0.376	0.398	0.366

Table 6
Model performance based on cross-validation of all classification methods for the open-sea incidents for features excluding latitude, longitude, and month.

Classification model	Accuracy	Precision (macro)	Recall (macro)	F-1 score (macro)
Logistic Regression	0.385	0.377	0.417	0.331
Naïve Bayes	0.406	0.431	0.340	0.267
Support Vector Machine	0.347	0.352	0.403	0.308
Random Forest	**0.488**	**0.395**	**0.374**	**0.362**



(c) Support Vector Machine (d) Random Forest

Fig. 4. Feature importance for all models using NWP data.

Finally, partial dependence plots (Figs. 7 and 8) are obtained from the models for each class of the target variable. It is observed that for the random forest model using the NWP data, the probability of predicting equipment failure is higher for longer vessels and higher degree of longitude, newer vessels, lower degree of latitude, and lower temperature at 2 m above surface. The probability of predicting the collision related incidents is greater for higher degree of latitude. The probability for predicting grounding related incidents is greater for older vessels, higher degree of latitude, later months, higher temperature at 2 m above surface, shorter vessel length and lower degree of longitude. For the random forest model with the NSR data, the probability of predicting equipment failure is greater for higher degree

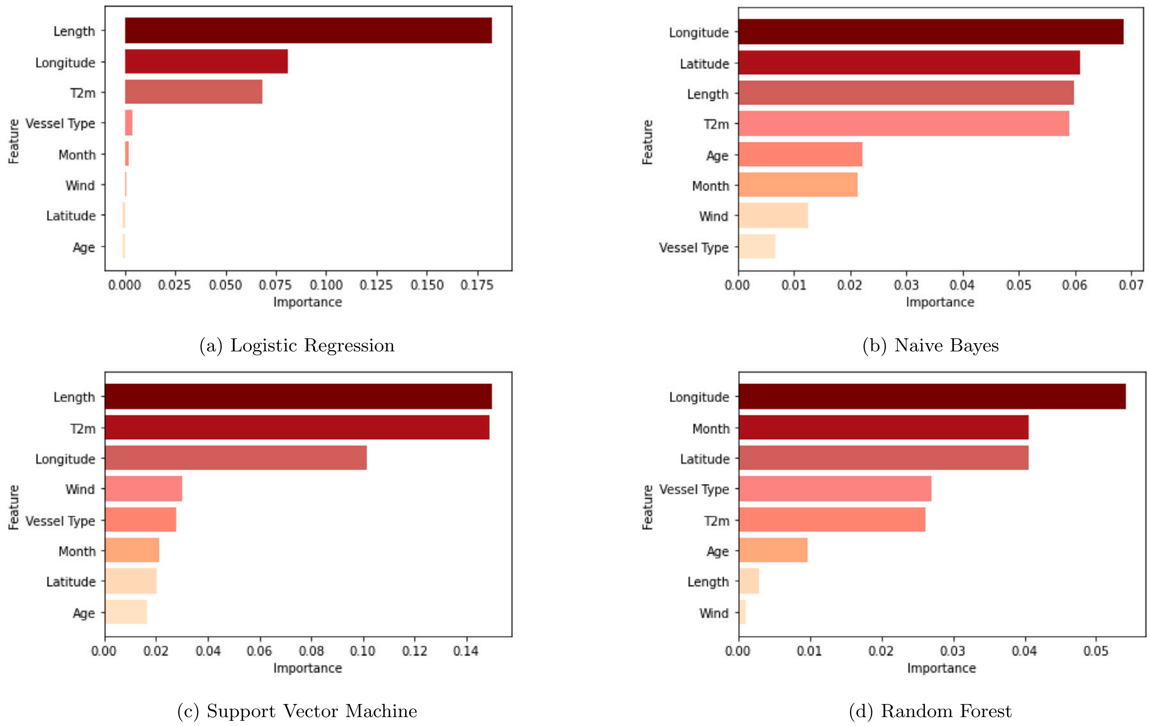


Fig. 5. Feature importance for all models using NSR data.

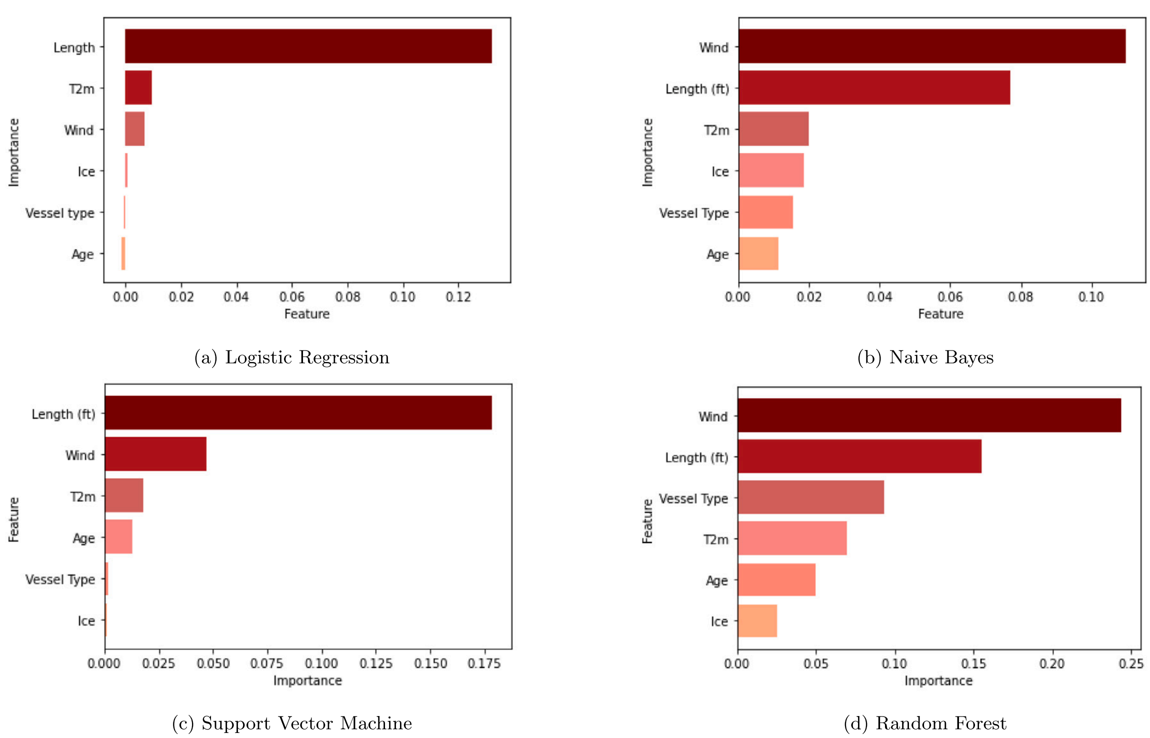
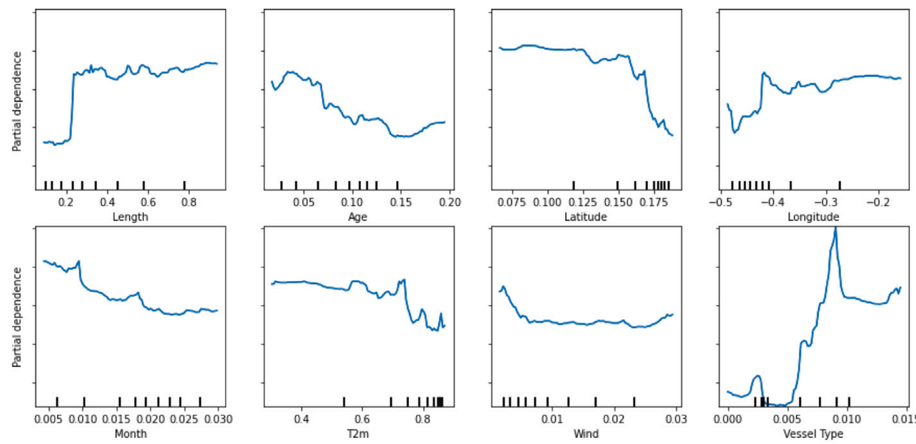
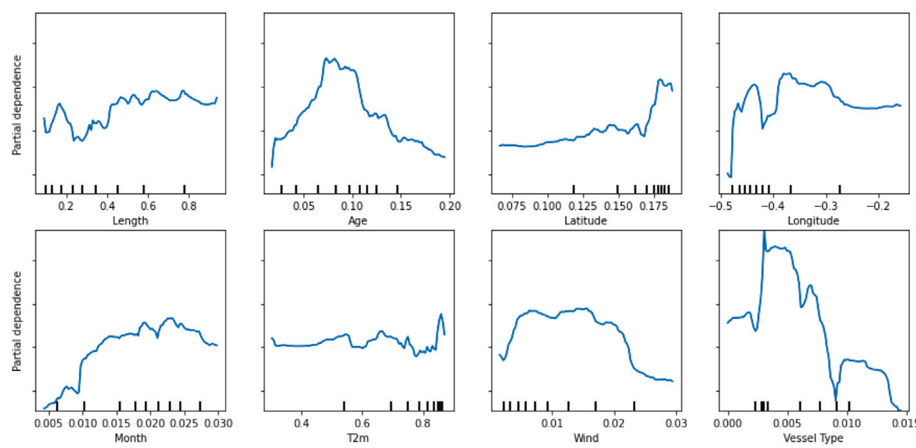


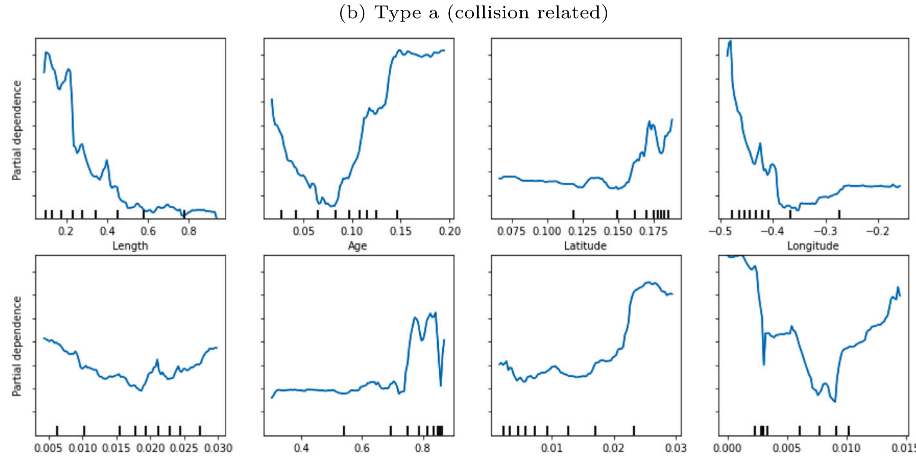
Fig. 6. Feature importance for all models using open sea data (excluding input features Latitude, Longitude, and Month.



(a) Equipment failure



(b) Type a (collision related)



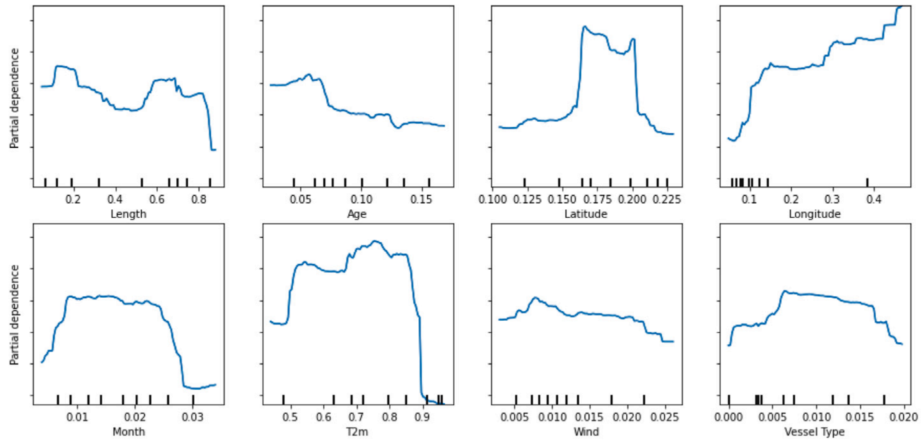
(c) Type b (grounding related)

Fig. 7. Partial dependence plots for Random Forest model using NWP data.

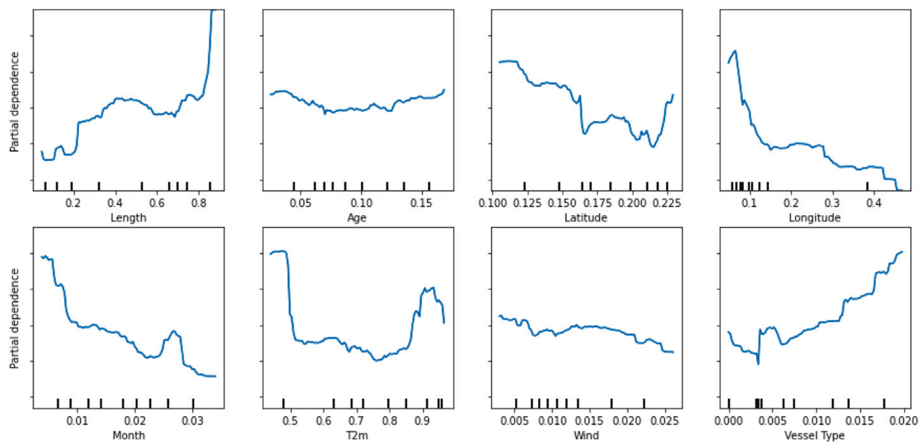
of longitude, shorter vessels, newer vessels, and lower wind speed. The probability of predicting collision related incidents is higher for longer vessels, lower degree of latitude, longitude, earlier months of the year, and lower wind speed. There is a higher probability of predicting grounding related incidents for later months in the year, higher temperature at the surface, wind speed, older vessels, lower degree of longitude and shorter vessels.

6. Discussion

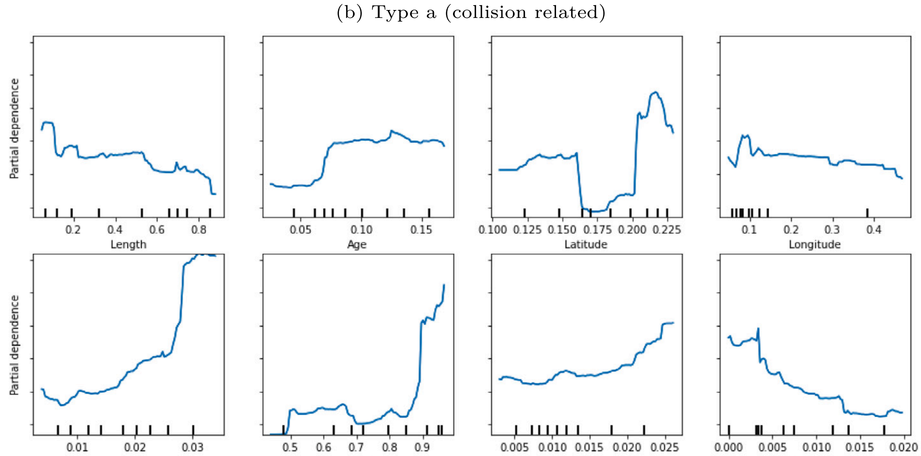
The results from the machine learning models provide insights on model predictive performance, feature importance, and partial dependence for the input features. This section discusses some of the steps taken in the modeling process to arrive at the final models and their impacts on model performance, insights gained from the results section, and limitations associated with this study.



(a) Equipment failure



(b) Type a (collision related)



(c) Type b (grounding related)

Fig. 8. Partial dependence plots for Random Forest model using NSR data.

Models were initially setup to predict six target classes of incident types. The number of output classes was reduced to three for reasons discussed earlier in the data section. An improvement in predictive accuracy is observed in all model performances as a result of reducing the number of target classes from six to three. Following the reduction in the number of classes, the naive Bayes model emerges as a suitable method to predict incident types in the NSR region with the best accuracy and F-1 scores from cross-validation. The NSR portion of the data has a smaller sample size with 174 data points compared to the NWP with 1075 incidents and the Naive Bayes technique is known to perform relatively better than other methods when there are lesser number of observations. The random forest model still has the best performance on the classification of the test data for all three datasets (Tables A.5, A.6, and A.7).

While the incident types are grouped into three classes to address the sample size in each class, the incident classes still remain imbalanced under the new classification, especially in the NSR data. This issue is further addressed using an oversampling technique called

Table 7
Performance of null models for different subsets of the incident data.

Null model	Accuracy	Precision (macro)	Recall (macro)	F-1 score (macro)
NWP	0.325	0.108	0.333	0.156
NSR	0.420	0.140	0.333	0.196

Synthetic Minority Oversampling TEchnique (SMOTE). The equipment failure class has notably higher number of observations compared to the other classes and thus oversampling generates synthetic observations to increase the number of minority observations up to the level of equipment failures. The use of this technique does not improve the model accuracy as expected but increases the recall and F-1 score. Therefore, it is still implemented to ensure training of models using a balanced dataset and better model performance overall.

Although the accuracy and F-1 scores of the best models are not very high, the best models for NSR and NWP provide a significant improvement in comparison to the accuracy of base (null) model, despite there being a majority class in the response variable that accounts for approximately 45% of the observations. The null model yields an accuracy of 32.5% and 42% for the NWP and NSR data, respectively, with a poor precision. The evaluation metrics obtained for a null model used to predict incident classes (Table 7) reflects this observation. On average, the proposed models in this study improve the accuracy and the precision by 68% and 280% respectively for the NWP, and 21% and 230% respectively for the NSR. Furthermore, the evaluation metrics are consistent with the precision (0.4–0.6) and recall (0.4–0.5) values reported in prior work that uses three-class prediction model for maritime accident warning system [28]. The improvement in the prediction of incident types using these data-driven models exhibits their value in improving Arctic navigation preparedness, especially with further improvements in the model accuracy and precision.

The latitude and longitude are both important predictors that present a high value for the feature importance in each of the machine learning models predicting the incident types. The locations of the observed incidents, therefore, are observed to be important in determining the type of risks the vessels are most prone to at a given point in the voyage. Although, there are many other factors (e.g., weather conditions, vessel properties, sea ice concentration) that may collectively explain the risks to the vessels at a certain location. To explore the effects of these factors on model performances, an analysis is performed on a subset of the dataset with instances that occurred in open ocean where there is a higher influence of the weather conditions. The model performances without features describing location (latitude and longitude) and time of year (month), while not as good as with these features, are still valuable in terms of the accuracy and F-1 scores. Fig. 6 shows the feature importance for the models using open-sea data. This suggests that the vessel properties and weather conditions by themselves are as important in the prediction of the Arctic incident types. Among the weather variables, the wind speed is more important than the temperature at 2 m for the two best performing open-sea models. This observation is consistent with the NWP models as a high number of open-sea incidents have occurred in the NWP region.

The relative feature importance of some input features for the same modeling technique are seen to be different from each other for different datasets. Comparing the feature importance for random forest models reveals this tendency with the weather variables (Figs. 4(d), 5(d), 6(d)). The wind speed is almost twice as important as the temperature variable for the NWP data while the importance of temperature is higher for the NSR data. An explanation for this observation might lie in the difference in the change in sea ice cover between the two regions in recent years. The sea surface is rougher when there is a higher ice cover. As a result, the ice-free areas in the Arctic are more exposed to the direct effect of strong winds [46]. The change in the ice cover over the last decade is lower for the NSR as compared to NWP. This points to a faster increase in ice-free water surface in the NWP and therefore a higher variation in wind speeds in the region , making the wind speed a more important feature. The temperature ranges for the NSR region are also more spread out compared to the NWP, making it more important to the models. A covariance analysis between the input variables shows that the covariance between the wind speed and the temperature at 2 m from surface is −0.07. Since this value is close to 0, it suggests that the wind speed does not change positively or too negatively when there is a change in the temperature and vice-versa. Therefore, the relationship between the observations of these two variables is not strong.

The observations from the partial dependence plots reveal the relationship between the predicted class of incidents and the input features for the random forest model. Collision based incidents are more likely to be predicted for longer vessels in both the NWP and NSR. This is intuitive as a longer vessel might be less responsive to its rudder [47] and thus could have difficulties avoiding obstacles. Another intriguing relationship observed is that there is a higher probability of predicting grounding related incidents in higher latitude and higher temperatures at 2 m above surface. Higher temperatures generally mean lesser sea-ice and more navigable conditions. However, for the Arctic, it also means the ships could be navigating in newly open regions that were unexplored before [48]. This can lead to more grounding incidents if shallow conditions are not detected on time. Similarly, the higher latitudes are less explored/ mapped compared to the lower latitudes which could also lead to conditions that aid grounding related incidents.

There are some limitations in assessing the incidents in the Arctic based on a single dataset aggregated from numerous data sources, as is the case with the PAME incident dataset. Roughly a third of the recorded incidents in the dataset are originally labeled as type ‘Consequences’ which gives no information about the actual incident type. These records are omitted from this analysis because of the incompleteness of the information provided. Similarly, it is observed that the type ‘Equipment failure’ is mainly clustered in the NWP region and the type ‘Loss of control’ in the NSR region. This is either because these incident types are typical to the corresponding regions or because there are different record keeping norms in the two regions of the Arctic. In addition, the Arctic incident records do not contain detailed information on human and organizational factors pertaining to these incidents. The inclusion of human risk factors would require the elicitation of expert opinion from a diverse group of stakeholders [49] and the use of probabilistic and simulation approaches to examine the influence of different behavioral factors [9].

7. Conclusion

This study analyzes a comprehensive dataset of Arctic maritime navigation incidents and applies machine learning techniques to predict the incident type based on historical data. The study uses the Arctic incident records between the years 2005 and 2017 with the corresponding weather conditions to perform an exploratory analysis and develop prediction models.

The exploratory data analysis of Arctic maritime shipping incidents data provides a breakdown of the incidents by vessel type, incident type and temporal trends (month, year) of incidents. Fishing vessels and passenger ships are the most frequent vessel types involved in incidents with service and cargo vessels also accounting for high number of incidents. The most frequent incident types are Equipment failure, Loss of control, Grounding, Foundering, Contact, and Collision. The rate of Arctic incidents every year is observed to be on the rise over the period of record. Additionally, the summer months are found to have greater rates of incidents compared to the winter months, particularly in the NWP. Incidents in areas in the Arctic along the NSR and the

NWP notably differ in their characteristics based on the features of their incident records.

Machine learning classification methods are used to classify incident types using the recorded features of the incidents and corresponding weather variables. The machine learning techniques predict the potential incident types with up to 60% accuracy and 57% precision based on these input features. A comparison between the performance of the models shows that the random forest model is best suited for this classification. The observations from the partial dependence analysis of the models can be utilized to improve preparedness towards Arctic maritime incidents. For instance, service stations to handle equipment failures could be stationed strategically to support longer vessels and in areas with lower longitudes (western Arctic) and lower temperature along the NWP. Similarly, precautions against collisions and fire could be taken for longer vessels in the western part of the NSR. Grounding, flooding, and loss of control are predicted more in warmer temperatures and for shorter vessels in general. Rescue ships can be assigned in greater numbers for fishing areas for those high-risk conditions.

This study complements probabilistic models in identifying the risk factors and predicting the type of Arctic incident. The probabilistic models using Bayesian networks for risk analysis of Arctic shipping assume the probability associated with the risk factors from prior knowledge based on expert opinion in cases where historical records are scarce. Through the use of machine learning methods, this research utilizes the recent increase in the availability of Arctic incident records and seeks to overcome issues of scalability and uncertainties in deriving priors for probabilistic methods. The paper expands on this contribution by incorporating the weather observations corresponding to the Arctic incidents and analyzing their importance in the prediction models.

However, expansions need to be made in the modeling process to obtain a comprehensive risk analysis framework for Arctic maritime navigation. As the models are trained on just the incident records, they predict the most likely incident class given there is an incident. Analyzing the incident records and the incident type predictions is an important step towards understanding the circumstances in which the incidents occur. Still, it is equally important to assess the general maritime ship traffic conditions as the basis of the analysis for a more complete picture. To achieve this, further analysis should be performed using the Arctic ship traffic to identify the likelihood of incidents based on the common attributes of the traffic and incident data. Additionally, machine learning models that predict the likelihood of incidents could be combined with the model outputs in this paper to provide real-time risk updates for the trajectories of the vessels navigating the Arctic. As more climate data is obtained from various climate models, the uncertainty in the model predictions stemming from the climate input variables could be evaluated to develop a comprehensive risk analysis framework for Arctic maritime navigation.

CRediT authorship contribution statement

Rajesh Kandel: Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation. **Hiba Baroud:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

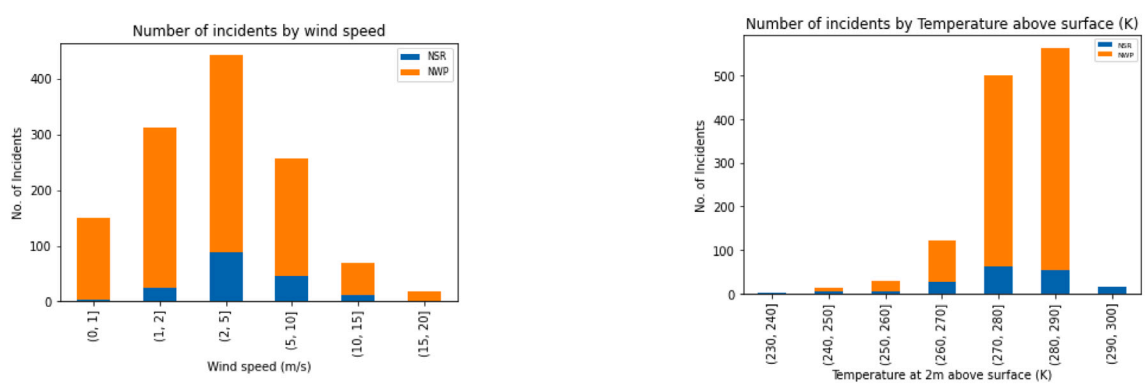
Data is available at https://doi.org/10.18739/A2DJ58J37.

Acknowledgments

This material is based upon work supported by the National Science Foundation, USA under Grant 1928112.

Appendix

See Figs. A.1–A.5 and Tables A.1–A.7.



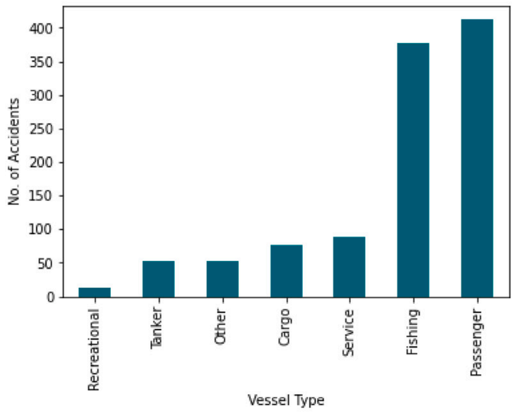
(a) Number of incidents by *Wind speed (m/s)*

(b) Number of incidents by *Temperature at 2m above surface*

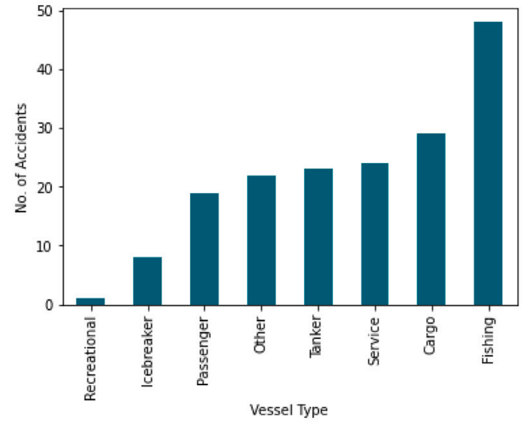
Fig. A.1. Number of incidents broken down by wind speed and temperature.

Table A.1
Standard confusion matrix.

		Actual Class	
		Positive (P)	Negative (N)
Predicted Class	Positive (P)	True Positive (TP)	False Positive (FP)
	Negative (N)	False Negative (FN)	True Negative (TN)

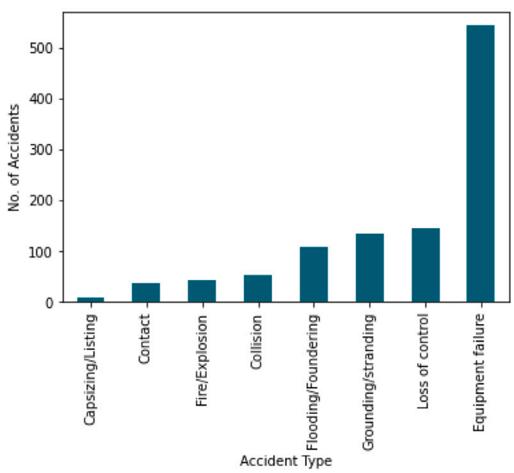


(a) Number of incidents by *Vessel type* in NWP

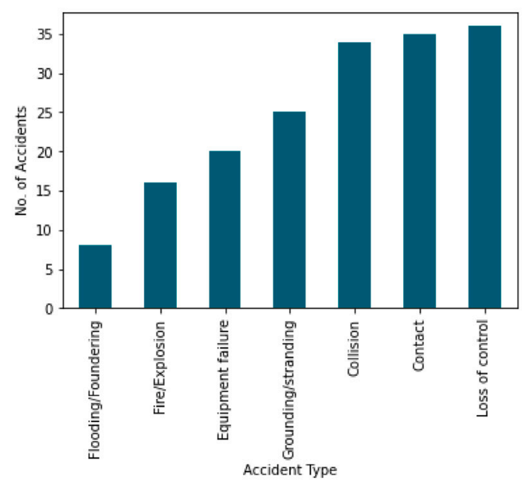


(b) Number of incidents by *Vessel type* in NSR

Fig. A.2. Comparison of incidents in NWP and NSR by vessel type.

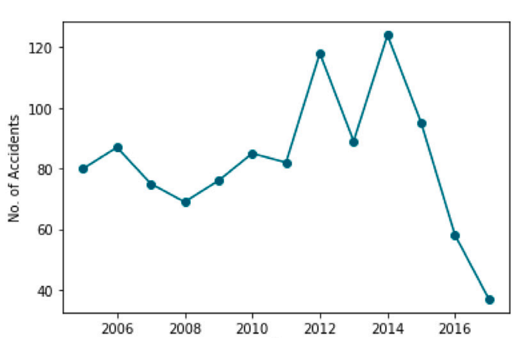


(a) Number of incidents by *Incident type* in NWP

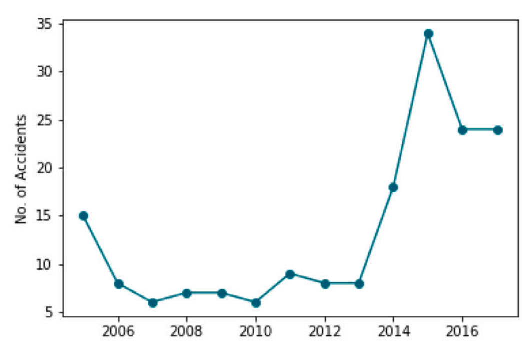


(b) Number of incidents by *Incident type* in NSR

Fig. A.3. Comparison of incidents in NWP and NSR by incident type.

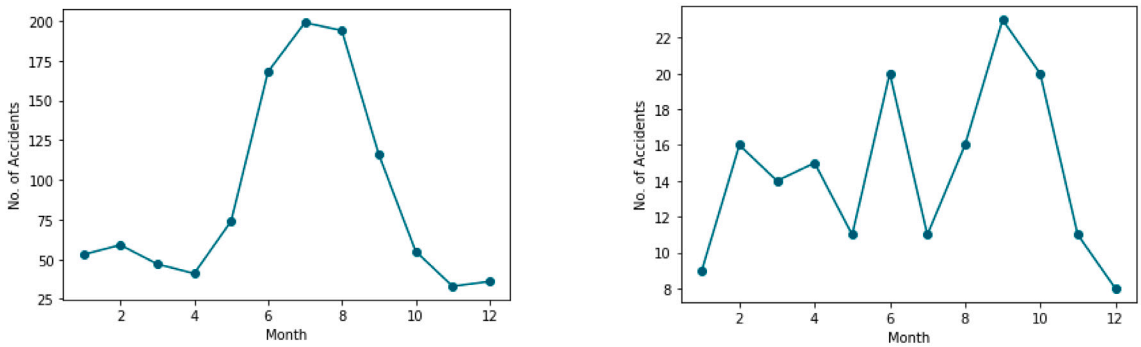


(a) Number of incidents by *Year* in NWP



(b) Number of incidents by *Year* in NSR

Fig. A.4. Comparison of incidents in NWP and NSR by year.



(a) Number of incidents by *Month* in NWP

(b) Number of incidents by *Month* in NSR

Fig. A.5. Comparison of incidents in NWP and NSR by month.

Table A.2
List of input features for the best performing model for each technique for the entire Arctic data.

Classification model	Features
Logistic Regression	Length, Age, Latitude, Longitude, Month, Temperature at 2 m, Wind, Vessel Type
Naïve Bayes	Length, Age, Latitude, Longitude, Temperature at 2 m, Wind, Vessel Type
Support Vector Machine	Length, Age, Latitude, Longitude, Temperature at 2 m, Vessel Type
Random Forest	Length, Age, Latitude, Longitude, Month, Temperature at 2 m, Wind, Vessel Type

Table A.3
List of input features for the best performing model for each technique for the NWP data.

Classification model	Features
Logistic Regression	Length, Age, Latitude, Longitude, Month, Temperature at 2 m, Wind
Naïve Bayes	Length, Age, Latitude
Support Vector Machine	Length, Age, Latitude, Longitude, Month, Temperature at 2 m, Vessel Type
Random Forest	Length, Age, Latitude, Longitude, Month, Temperature at 2 m, Wind, Vessel Type

Table A.4
List of input features for the best performing model for each technique for the NSR data.

Classification model	Features
Logistic Regression	Length, Latitude, Longitude, Month
Naïve Bayes	Length, Age, Latitude, Longitude, Vessel Type
Support Vector Machine	Length, Age, Latitude, Longitude, Temperature at 2 m, Vessel Type
Random Forest	Length, Age, Latitude, Longitude, Month, Temperature at 2 m, Wind, Vessel Type

Table A.5
Evaluation metric values obtained from test data of all classification methods for the entire dataset.

Classification model	Accuracy	Precision (macro)	Recall (macro)	F-1 score (macro)
Logistic Regression	0.496	0.475	0.464	0.468
Naïve Bayes	0.532	0.529	0.494	0.502
Support Vector Machine	0.548	0.509	0.502	0.494
Random Forest	0.580	0.560	0.536	0.543

Table A.6
Evaluation metric values obtained from test data of all classification methods for the NWP incidents.

Classification model	Accuracy	Precision (macro)	Recall (macro)	F-1 score (macro)
Logistic Regression	0.502	0.381	0.386	0.382
Naïve Bayes	0.530	0.346	0.384	0.358
Support Vector Machine	0.507	0.412	0.406	0.389
Random Forest	0.544	0.426	0.416	0.410

Table A.7
Evaluation metric values obtained from test data of all classification methods for the NSR incidents.

Classification model	Accuracy	Precision (macro)	Recall (macro)	F-1 score (macro)
Logistic Regression	0.457	0.352	0.321	0.333
Naïve Bayes	0.486	0.433	0.488	0.429
Support Vector Machine	0.514	0.458	0.507	0.458
Random Forest	0.600	0.401	0.424	0.412

References

[1] Overland J, Dunlea E, Box JE, Corell R, Forsius M, Kattsov V, et al. The urgency of arctic change. Polar Sci 2019;21:6–13.
[2] Arctic-Council. Navigating the future of Arctic shipping. 2021, URL https://arctic-council.org/news/navigating-the-future-of-arctic-shipping. [Accessed 18 November 2021].
[3] Buixadé Farré A, Stephenson SR, Chen L, Czub M, Dai Y, Demchev D, et al. Commercial Arctic shipping through the Northeast Passage: routes, resources, governance, technology, and infrastructure. Polar Geogr 2014;37(4):298–324.
[4] Marchenko N, Andreassen N, Borch OJ, Kuznetsova S, Ingimundarson V, Jakobsen U. Arctic shipping and risks: emergency categories and response capacities. TransNav: Int J Mar Navig Saf Sea Transp 2018;12(1).
[5] Willmer G. How to protect the Arctic as melting ice opens new shipping routes. 2018, URL https://ec.europa.eu/research-and-innovation/en/horizon-magazine/how-protect-arctic-melting-ice-opens-new-shipping-routes. [Accessed 01 January 2023].
[6] Xu S, Kim E, Haugen S, Zhang M. A Bayesian network risk model for predicting ship besetting in ice during convoy operations along the Northern Sea Route. Reliab Eng Syst Saf 2022;223:108475.
[7] Li Z, Yao C, Zhu X, Gao G, Hu S. A decision support model for ship navigation in Arctic waters based on dynamic risk assessment. Ocean Eng 2022;244:110427.
[8] Fu S, Yu Y, Chen J, Xi Y, Zhang M. A framework for quantitative analysis of the causation of grounding accidents in arctic shipping. Reliab Eng Syst Saf 2022;226:108706.
[9] Li H, Ren X, Yang Z. Data-driven Bayesian network for risk analysis of global maritime accidents. Reliab Eng Syst Saf 2023;230:108938.
[10] Zhang Y, Meng Q, Ng SH. Shipping efficiency comparison between Northern Sea Route and the conventional Asia-Europe shipping route via Suez Canal. J Transp Geogr 2016;57:241–9.
[11] Eguíluz VM, Fernández-Gracia J, Irigoien X, Duarte CM. A quantitative assessment of Arctic shipping in 2010–2014. Sci Rep 2016;6(1):1–6.
[12] Fu S, Zhang D, Montewka J, Yan X, Zio E. Towards a probabilistic model for predicting ship besetting in ice in Arctic waters. Reliab Eng Syst Saf 2016;155:124–36.
[13] Afenyo M, Khan F, Veitch B, Yang M. Arctic shipping accident scenario analysis using Bayesian Network approach. Ocean Eng 2017;133:224–30.
[14] Lasserre F. Analysis of maritime transit trends in the arctic passages. In: International law and politics of the arctic ocean: essays in honor of donat Pharand. Brill Academic Publishing; 2015.
[15] Faury O, Cariou P. The Northern Sea Route competitiveness for oil tankers. Transp Res A Policy Pract 2016;94:461–9.
[16] Gunnarsson B. Recent ship traffic and developing shipping trends on the Northern Sea Route—Policy implications for future arctic shipping. Mar Policy 2021;124:104369.
[17] Li X, Otsuka N, Brigham LW. Spatial and temporal variations of recent shipping along the Northern Sea Route. Polar Sci 2021;27:100569.
[18] Pizzolato L, Howell SE, Dawson J, Laliberté F, Copland L. The influence of declining sea ice on shipping activity in the Canadian Arctic. Geophys Res Lett 2016;43(23):12–46.
[19] Baksh A-A, Abbassi R, Garaniya V, Khan F. Marine transportation risk assessment using Bayesian Network: Application to Arctic waters. Ocean Eng 2018;159:422–36.
[20] Aziz A, Ahmed S, Khan F, Stack C, Lind A. Operational risk assessment model for marine vessels. Reliab Eng Syst Saf 2019;185:348–61.
[21] Kum S, Sahin B. A root cause analysis for Arctic Marine accidents from 1993 to 2011. Saf Sci 2015;74:206–20.
[22] Zhang Y, Hu H, Dai L. Real-time assessment and prediction on maritime risk state on the Arctic Route. Marit Policy Manag 2020;47(3):352–70.
[23] Zhang G, Thai VV. Expert elicitation and Bayesian Network modeling for shipping accidents: A literature review. Saf Sci 2016;87:53–62.
[24] Bye RJ, Aalberg AL. Maritime navigation accidents and risk indicators: An exploratory statistical analysis using AIS data and accident reports. Reliab Eng Syst Saf 2018;176:174–86.
[25] Rawson A, Brito M, Sabeur Z, Tran-Thanh L. A machine learning approach for monitoring ship safety in extreme weather events. Saf Sci 2021;141:105336.
[26] Wang H, Liu Z, Wang X, Graham T, Wang J. An analysis of factors affecting the severity of marine accidents. Reliab Eng Syst Saf 2021;210:107513.
[27] Zhang M, Kujala P, Hirdaris S. A machine learning method for the evaluation of ship grounding risk in real operational conditions. Reliab Eng Syst Saf 2022;226:108697.
[28] Merrick JR, Dorsey CA, Wang B, Grabowski M, Harrald JR. Measuring prediction accuracy in a maritime accident warning system. Prod Oper Manage 2022;31(2):819–27.
[29] PAME. Compendium of Arctic Ship Accidents. URL https://pame.is/projects-new/arctic-shipping/pame-shipping-highlights/457-compendium-of-arctic-ship-accidents.
[30] Cox DR. The regression analysis of binary sequences. J R Stat Soc Ser B Stat Methodol 1958;20(2):215–32.
[31] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. Vol. 112. Springer; 2013.
[32] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20(3):273–97.
[33] Liaw A, Wiener M. Classification and regression by randomforest. R News 2002;2(3):18–22, URL https://CRAN.R-project.org/doc/Rnews/.
[34] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321–57.
[35] Kandel R, Baroud H, Bennartz R. Arctic Maritime Incident Records with Weather Observations (2005–2017). Arctic Data Center; 2023, http://dx.doi.org/10.18739/A2DJ58J37.
[36] Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al. The ERA5 global reanalysis. Q J R Meteorol Soc 2020;146(730):1999–2049.
[37] Riviera. Indian rig fire blamed on supply ship collision. 2009, URL https://www.rivieramm.com/news-content-hub/news-content-hub/indian-rig-fire-blamed-on-supply-ship-collision-50454. [Accessed 06 November 2022].
[38] JOC. Maersk container ship catches fire in collision off China coast. 2016, URL https://www.joc.com/maritime-news/container-lines/maersk-line/maersk-container-ship-catches-fire-collision-china_20160508.html. [Accessed 06 November 2022].
[39] DW. China searches for missing crew after tanker collision. 2018, URL https://www.dw.com/en/dozens-missing-after-tanker-collision-in-east-china-sea/a-42053626. [Accessed 06 November 2022].
[40] TurkishMaritime. Major tanker fire after collision with barge near Shanghai. 2020, URL http://www.turkishmaritime.com.tr/major-tanker-fire-after-collision-with-barge-near-shanghai-38465h.htm. [Accessed 06 November 2022].
[41] CIAIM. Loss of control, grounding and complete loss of cargo ship LUNO on the breakwater of the outer harbour of Bayonne (France). 2014, URL https://www.mitma.gob.es/recursos_mfom/comodin/recursos/s34-2014-ingles.pdf. [Accessed 06 November 2022].
[42] Safety4Sea. Lessons learned: Loss of steering leads to grounding. 2018, URL https://safety4sea.com/lessons-learned-loss-of-steering-leads-to-grounding/. [Accessed 06 November 2022].
[43] MarineInsight. Real life incident: Grounding leads to major listing of ship, declared total constructive loss. 2021, URL https://www.marineinsight.com/case-studies/real-life-incident-grounding-leads-to-major-listing-of-ship-declared-total-constructive-loss/. [Accessed 06 November 2022].
[44] Blenkey N. China searches for missing crew after tanker collision. 2021, URL https://www.marinelog.com/legal-safety/shipping/deviation-from-planned-route-seen-in-grounding-of-giant-ore-carrier/. [Accessed 06 November 2022].
[45] Dixon G. Portuguese tanker stable after hull breached in grounding off Azores. 2022, URL https://www.tradewindsnews.com/casualties/portuguese-tanker-stable-after-hull-breached-in-grounding-off-azores/2-1-1165418. [Accessed 06 November 2022].
[46] Martin T, Tsamados M, Schroeder D, Feltham DL. The impact of variable sea ice roughness on changes in A rctic O cean surface stress: A model study. J Geophys Res: Oceans 2016;121(3):1931–52.
[47] USNA. Introduction to maneuvering. 2020, URL https://www.usna.edu/NAOE/_files/documents/Courses/EN455/AY20_Notes/EN455CourseNotesAY20_Chapter6.pdf. [Accessed 21 May 2023].
[48] TSB. Marine transportation safety investigation report M18C0225. 2018, URL https://www.tsb.gc.ca/eng/rapports-reports/marine/2018/m18c0225/m18c0225.html. [Accessed 21 May 2023].
[49] Fan S, Blanco-Davis E, Yang Z, Zhang J, Yan X. Incorporation of human factors into maritime accident analysis using a data-driven Bayesian network. Reliab Eng Syst Saf 2020;203:107070.