Electronic
Research Archive

Research article

Ship detention prediction using anomaly detection in port state control: model and explanation

Ran Yan* and Shuaian Wang

Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

* Correspondence: Email: angel-ran.yan@connect.polyu.hk; Tel: +85227664614; Fax: +85223302704.

Abstract: Maritime transport plays an important role in global supply chain. To guarantee maritime safety, protect the marine environment, and enhance the living and working conditions of the seafarers, international codes and conventions are developed and implemented. Port state control (PSC) is a critical maritime policy to ensure that ships comply with the related regulations by selecting and inspecting foreign visiting ships visiting a national port. As the major inspection result, ship detention, which is an intervention action taken by the port state, is dependent on both deficiency/deficiencies (i.e., noncompliance) detected and the judgement of the inspector. This study aims to predict ship detention based on the number of deficiencies identified under each deficiency code and explore how each of them influences the detention decision. We innovatively view ship detention as a type of anomaly, which refers to data points that are few and different from the majority, and develop an isolation forest (iForest) model, which is an unsupervised anomaly detection model, for detention prediction. Then, techniques in explainable artificial intelligence are used to present the contribution of each deficiency code on detention. Numerical experiments using inspection records at the Hong Kong port are conducted to validate model performance and generate policy insights.

Keywords: Port state control (PSC); ship detention; anomaly detection; isolation forest (iForest)

Abbreviations: COVID-19: Coronavirus disease 2019; EEDI: Energy Efficiency Design Index; iForest: isolation forest; IMO: International Maritime Organization; ISM: International Safety Management; iTree: isolation tree; MLC: Maritime Labour Convention; MoU: memorandums of understanding; PSC: port state control; PSCO: port state control officer; ROC: receiver operating

characteristic; ROC AUC: area under the ROC curve; SEEMP: Ship Energy Efficiency Management Plan; SHAP: Shapley additive explanation; SOLAS: International Convention for the Safety of Life at Sea; TAN: tree augmented naive Bayes; UNCTAD: United Nations Conference on Trade and Development; USCG: United States Coast Guard; XAI: explainable artificial intelligence

1. Introduction

Maritime transport is the backbone of globalized trade and the manufacturing supply chain [1−6]. Even during the difficult time due to the COVID-19, international shipping still takes significant responsibilities of transporting products and goods around the world [7−9]. Being the most dangerous transportation mode due to the adverse sea and weather conditions, the lack of rescue force at sea, and the nature of the goods transported (e.g., crude oil and chemicals), various international regulations and conventions on shipping activities are implemented to guarantee maritime safety [10−12], with the International Convention for the Safety of Life at Sea (SOLAS) as the most important [13−15]. Seafarers engaged in ocean shipping play a key role to the movement and growth of the maritime industry, and their working and living conditions as well as the health and medical care are mainly protected by the Maritime Labour Convention (MLC) [16−17]. In recent years, more policies and regulations on environmental sustainability are imposed on the maritime industry, with the Energy Efficiency Design Index (EEDI) for new ships and the Ship Energy Efficiency Management Plan (SEEMP) for all ships as representatives [18−21].

Ships whose conditions are substantially below the standards required by the relevant convention or whose crew is not in conformity with the safe manning document are called substandard ships [22]. Ship flag state, which is the country under ship registration, is the first line of defense against substandard shipping. However, due to the internationality property of the shipping industry, it is believed that flag states cannot perform their duties well [22−24]. Under this condition, port state control, or PSC, is proposed as the second line of defence of substandard shipping, whose main goal is to inspect foreign visiting ships so as to eliminate substandard shipping. Currently, nine regional memorandums of understandings (MoUs) on PSC in addition to the United States Coast Guard (USCG) are established.

Table 1. Deficiency code list of Tokyo MoU [25].

Code	Item	Code	Item
D1	Certificates & documentation	D10	Safety of navigation
D2	Structural condition	D11	Life saving appliances
D3	Water/Weathertight condition	D12	Dangerous goods
D4	Emergency system	D13	Propulsion and auxiliary machinery
D5	Radio communication	D14	Pollution prevention
D6	Cargo operations including equipment	D15	International Safety Management (ISM)
D7	Fire safety	D18	Labour conditions
D8	Alarms	D99	Other
D9	Working and living conditions		

For foreign visiting ships to a port, the port authority first selects ships of higher risks for inspection, and ship inspection is carried out by an inspector called PSC officer (PSCO). During an inspection, a condition found not to be in compliance with the requirements of the relevant conventions

is recorded as a deficiency [22]. Deficiency codes in different MoUs are slightly different. For example, the list of deficiency codes of Tokyo MoU, which is responsible for the Asia-Pacific region, is shown in Table 1. If the condition of the ship or its crew is substantially below the standards, an intervention action, named detention, can be taken by the port state. The detention decision is highly subjective to the deficiency/deficiencies detected onboard, and a deficiency leading to detention is called a detainable deficiency. Although lists of typical but not exhaustive detainable deficiencies under relevant regulations and conventions are provided in the PSC manuals published by the IMO and individual MoUs, the descriptions are relatively abstract. Therefore, the detention decision is still highly subject to professional judgements of PSCOs, where it is observed that a detainable deficiency code may not lead to detention sometimes.

Meanwhile, as the economic loss of ship detention can be very high and ship schedule might be delayed due to detention, ship operators take various measures to reduce the detention risk of their ships, while PSCOs are also cautious to detain a ship. Therefore, it is observed that ship detention rate is very low at a port. Take the Hong Kong port which belongs to the Tokyo MoU as an example, the annual detention rate ranges from 2.96 to 3.67% between 2015 and 2019. In all the member states within the Tokyo MoU, the detention rate is below 5% in the recent decades [26]. Given the facts that only when a significant deficiency or several significant deficiencies are found will a ship be detained and the ship detention rate is very low, it is justifiable to regard ship detention as a type of 'anomaly' and ships without detention as normal behavior. This is because anomalies are pattern in data that do not conform to a well defined notion of normal behavior, where the normal behavior refers to the conditions (e.g., the types of deficiencies detected and the number of the deficiencies) of most of the ships which are not detained, and the nonconformity is the detection of the detainable deficiencies that warrant a detention. Given this setting, this study develops an anomaly detection approach based on isolation forest, or iForest, for ship detention prediction, with the number of deficiencies under each deficiency code as the input and ship detention as the prediction target. Managerial implications are generated from the anomaly detection model and the prediction results.

2. Literature review

As PSC plays a major role in international marine policy and management, and the inspection results, including deficiency and detention conditions, are important to both port authorities and ship operators, there are many studies on predicting ship deficiency and detention conditions. We review these two streams of literature in this section.

Most of the related studies aim to predict ship deficiency condition, either from the perspective of the total number of deficiencies of an inspection or from the perspective of the number of deficiencies under each deficiency code. Early examples of ship deficiency number prediction based on machine learning models include [27,28] which use support vector machine and [29] which uses a combination of k-nearest neighbor and support vector machine. In recent years, there is a boom in developing more advanced machine learning models for ship deficiency number prediction. For example, to predict the total number of deficiencies a ship has, tree augmented naive Bayes (TAN), which is a type of Bayesian networks, is developed in [23], and an XGBoost model considering shipping domain knowledge is proposed in [30]. Tailored random forest models under the framework of smart "predict, then optimize" are developed to predict the number of deficiencies under each deficiency category of a ship, which are informed by the structure and property of the following PSCO assignment model [31].

In contrast, there are much fewer studies on developing ship detention prediction models, as this is not a trivial task due to its highly imbalanced nature, i.e., un-detained ships highly outnumber detained ships. The related literature includes a TAN Bayesian network developed in [32], a balanced random forest model proposed in [33], and a support vector machine calibrated in [34]. However, to the best of our knowledge, the above studies only use standard supervised machine learning models for ship deficiency number and detention prediction. In addition, although ship detention is led by the detection of detainable deficiency or deficiencies, how each deficiency code contribute to the detention decision is unclear. To bridge this gap, this study innovatively view ship detention as a type of anomaly in data and adopt an unsupervised anomaly detection model based on iForest for ship detention prediction. The prediction results are then explained by an explainable artificial intelligence (XAI) method based on SHAP.

3. Research data

The whole dataset used in this study contains 2943 initial inspection records at the Hong Kong port from 2015 to 2019 with at least one deficiency detected. As we aim to explore the influence of each deficiency code on the detention decision, the input features are the number of deficiencies under each deficiency code, which is considered as continuous. The distribution of the input features is given in Table 2.

Table 2. Distribution of input features over the whole dataset.

Deficiency code	01	02	03	04	05	06	07	08
Mean	0.3123	0.0557	0.3782	0.3177	0.2508	0.0153	1.0527	0.0605
Min	0	0	0	0	0	0	0	0
Max	13	12	8	5	7	3	12	2
Standard deviation	0.8015	0.3691	0.7457	0.6508	0.5745	0.1359	1.3930	0.2496
Deficiency code	09	10	11	12	13	14	15	18
Mean	0.4757	0.7842	0.5814	0.0034	0.1549	0.2514	0.0849	0.0663
Min	0	0	0	0	0	0	0	0
Max	6	12	9	1	14	5	4	6
Standard deviation	0.7770	1.2401	0.9148	0.0582	0.6102	0.5771	0.3187	0.3410

The prediction target is ship detention: -1 if detained and 1, otherwise. In our dataset, there are a total of 144 records with detention among all the 2943 initial inspection records, and the detention rate is 4.89%, showing that the dataset is highly imbalanced. There are several reasons for such a low detention rate. The main reason is that detention of a ship is a serious matter involving many issues and interested parties [22], given the complex structure of the international shipping industry. Even if the main purpose of PSC is to prevent a substandard ship proceeding to sea, hard efforts should be made to avoid a ship being unduly detained or delayed. If this happens, the port authority should be obliged to compensation for any loss or damage suffered [22]. Therefore, PSCOs are very cautious about the decision of detention. In addition, conscientious ship operators and managers also spare no effort to make sure that their ships are of satisfied conditions and are of a low probability to be detained in PSC inspection.

After examining the inspection records, we find that detention is mainly caused by the identification of at least one detainable deficiency. However, we also observe that a specific deficiency code can be regarded as a detainable deficiency in some records while not in others. Actually, to the best of the authors' knowledge, there is no general rule of which deficiency is always detainable, and it is clearly stated in the procedure for PSC published by the IMO that "the PSCO should use professional judgement to determine whether to detain the ship" [22]. Therefore, it can be seen that detention highly depends on the experience, expertise, and the judgement of the PSCOs, in addition to the actual ship overall conditions. Hence, detained ships are highly likely to exhibit unique properties in their deficiencies detected, in terms of the total number of deficiencies, the specific types and combination of deficiency codes, and the number of deficiencies under each deficiency code, compared to ships that are not detained. Considering also that there are few detained ships, we regard ship detention as anomalies, and thus we adopt anomaly detection method for ship detention prediction in this study.

4. Anomaly detection methods for ship detention prediction

In this study, the prediction of ship detention is achieved by an anomaly detection method based on iForest model, which was proposed by [35,36]. iForest is unsupervised and is based on an ensemble of tree-based models. The assumptions of iForest are that the anomaly class is "few and different": it only contains the minority of samples that are much fewer than normal samples, and the feature values of the anomaly samples are very different from the normal samples. The concept of anomaly in iForest is analogous to detention in PSC inspection, where the ships with detention only constitute a very small number of all the inspected ships (no more than 5%), while they are detained because several serious deficiencies/one or more detainable deficiencies/a large number of deficiencies are detected, which make them different from the majority of the inspected ships.

The basic idea of iForest is to explicitly isolate anomalies, and thus one tree in iForest is also called an isolation tree (iTree). 'Isolation' means to separate a sample from the rest of the samples. In tree structure, it means splitting each sample to separate leave nodes. Given the "few and different" characteristics of anomalies, they are thus more likely to be split into a leaf node near the root, i.e., being isolated, thanks to their very different feature values, while the normal samples are more like to be split to deeper leaf nodes as they are similar to each other. In other words, a sample with a shorter path length from the root node to the leaf node it belongs to is more likely to be an anomaly in an iTree. Similarly, an example with a shorter average path length in the iForest model is more likely to be an anomaly. The proportion of the anomaly samples in the whole data set is denoted by ε and is preset. The iForest model is constructed based on the concept of isolation in a binary tree structure, which is fundamentally different from other distance and density based measures for anomaly detection.

Denote a training set containing n samples and m features by $D = (\mathbf{x}_i, y_i)$, $\mathbf{x}_i \in R^m$, $i = 1, ..., n$. An iTree is constructed using a random subset of D whose ratio to the whole dataset is denoted by α, $\alpha \in (0,1)$, and a random subset of the m features whose proportion to the total number of features is denoted by β, $\beta \in (0,1)$. The height of an iTree is limited to $l = \lceil \log_2(n \times \alpha) \rceil$.

Using iForest for anomaly detection contains two stages, namely training stage and evaluating stage. The training stage constructs a certain number of iTrees, and the evaluating stage calculates the anomaly score of each sample. Specifically, in the training stage, a node in an iTree is split using a random feature and a random value between the minimum and maximum values of this feature, and the construction process terminates when the tree reaches the height limit l, or all samples are

separated. As an iTree is of high uncertainty, iForest model is developing by formulating an ensemble of a total of T iTrees. Therefore, we consider a total of three hyperparameters in the iForest model in our study: α, β, and T. The procedure of constructing an iTree is given in Procedure 1, and the procedure of developing an iForest is given in Procedure 2.

Procedure 1. Construction of an iTree

Input	Training set D with n samples and m features, α (the proportion of the total number of samples used to construct an iTree), β (the proportion of the total number of features used to split a node), l (the height limit of an iTree).
Output	An iTree $f(t)$.
Step 1.	Formulate the input data set to the current iTree, which is constituted by αn random samples from D. Initialize the current tree height $l' = 0$.
Step 2.	While $l' < l$, do
Step 2.1.	Randomly select a feature from the βm candidate features and denote the selected feature by m'. Randomly select a value of feature m' and denote the selected value by $v(m')$.
Step 2.2.	Divide the current node into two child nodes, where the left child node contains samples with values for feature m' less than or equal to $v(m')$, and the right child node contains samples with values for feature m' larger than $v(m')$. Update $l' = l'+1$.
Step 2.3.	Repeat Step 2.1 and Step 2.2 in a depth-first manner on the child nodes.
	End while
Return	iTree $f(t)$.

Procedure 2. Construction of an iForest

Input	Training set D with n samples and m features, T (the number of iTrees contained in the iForest model), α, β, l.
Output	An iForest $f(x)$.
Step 1.	Initialize the iForest model $f(x) = 0$.
Step 2.	For $t = 1,...,T$: Construct iTree $f_t(x)$ using *Procedure 1*. Update $f(x) \leftarrow \dfrac{t-1}{t} f(x) + \dfrac{1}{t} f_t(x)$.
Return	iForest $f(x)$.

In the evaluating stage, anomaly scores are calculated for the samples considering their positions on each iTree in the iForest constructed. The anomaly score of sample i is denoted by $s(i)$, $i = 1,...,n$, and is calculated by

$$s(i) = 2^{-\frac{\sum_{t=1}^{T} h_t(i)}{T \times c(n)}}, \tag{1}$$

where $h_t(i)$ is the path length of sample i in iTree t, which is measured by the number of edges sample i traverses in iTree t from the root node to the leaf node it falls in, and $c(n)$ is the normalizer calculated by

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n}, \tag{2}$$

where $H(x)$ is the harmonic number that can be estimated by $\ln(x) + 0.5772156649$ (Euler's constant) [36]. $c(n)$ is an estimation of the average path length of samples in the iForest given the total number of samples in the data set as n, and it is used to normalize the anomaly score calculated for each sample. Obviously, we have $0 \le s(i) \le 1$ and $0 < h_t(i) \le n-1$. Especially, the average path length of sample i over all iTrees in the iForest model is $E(h_i) = \frac{\sum_{t=1}^{T} h_t(i)}{T}$. There are three special cases for $s(i)$ given $E(h_i)$ as follows:

1) When $E(h_i) \to 0$, that is, the leaf nodes where sample i lies in over the iForest model are very close to the root node, $s(i) \to 1$, which indicates that sample i is almost definitely an anomaly;

2) When $E(h_i) \to \alpha n - 1$, that is, the leaf nodes where sample i lies in over the iForest model are nearly at the end of a tree and are very far from the root node, $s(i) \to 0$, which indicates that sample i is almost impossible to be an anomaly; and

3) When $E(h_i) \to c(n)$, that is, the average path length of sample i over the iForest model is close to the average path length of samples in the iForest, meaning that the current sample is quite normal among all the samples, $s(i) \to 0.5$. In particular, if all the samples in the data set have $s(i) \approx 0.5$, then the entire sample may not really have obviously distinct anomalies.

Finally, in our problem, given the proportion of anomalies ε, samples with the $\varepsilon \times n$ highest $s(i)$ are predicted to be anomalies.

5. Numerical experiment

The whole dataset containing 2943 records is randomly divided into training set (80%, 2354 records, 114 detentions) and test set (20%, 589 records, 30 detentions). As the overall detention rate is about 5%, while the cost of ignoring a ship that should actually be detained can be very high, we set the proportion of anomaly samples, i.e., ε, to 0.1. One can find that the value of ε is much larger than the ratio of detained ships in the whole data set which is 4.89%. The main reason of doing so is that ship detention is a very serious outcome to both the vessel and the port authority, as it is highly related to vessels' navigation safety and management plannings. Therefore, the port authority may wish to identify as many ships that should be detained as possible at the cost of inspecting more ships and thus pay higher inspecting costs. Therefore, we set the value of ε much higher than the overall detention rate over the whole dataset, with the aim at catching as many ships that should be detained as possible. Although iForest is unsupervised, we have a labeled dataset (i.e., ship detention condition is known in the dataset), so we use 10-fold cross validation with grid search on the training set for hyperparameter tuning. The metric is area under the ROC (receiver operating characteristic) curve (ROC AUC). The search space and the selected values of the hyperparameters are shown in Table 3.

Table 3. Search spaces and selected values of hyperparameters in the iForest.

Parameter	Number of iTrees T	Sample subset ratio α	Feature subset ratio β
Search space	50 to 500 with 50 as the interval	0.2 to 1 with 0.2 as the interval	0.2 to 1 with 0.2 as the interval
Selected value	450	0.2	0.2

Then, the iForest model is constructed using the selected hyperparameter values on the training set, and its performance is validated on the test set. The confusion matrix on the test set is shown in Table 4.

Table 4. Confusion matrix of iForest on the test set.

	Ships detained	Ships not detained
Predicted ships detained	21	33
Predicted ships not detained	9	526

The overall ROC AUC score is 0.82, indicating that the probability that a randomly chosen ship with detention is predicted to have a higher anomaly score than a randomly chosen ship without detention is 0.82, which is much better than random guessing. Furthermore, precision of the iForest model is 0.69, showing that 69% of the ships that are predicted to be detained are indeed detained. The ratio is much higher than the detention rate 4.89% on the whole data set, showing that the proposed iForest model is more efficient in catching ships that should be detained. Recall of iForest model is 0.82, indicating that 82% of the ships that are detained by the port authority are accurately predicted by the iForest model. Then, F1 score of the iForest model, which reaches a trade-off between precision and recall by calculating their harmonic mean, is 0.73.

To explore the contribution of each deficiency code to the detention decision, we further use SHAP (Shapley additive explanations) method [37] on the training set to explain the importance of each deficiency code in the training set. The relative importance of the deficiency codes after normalizing is shown in Figure 1.
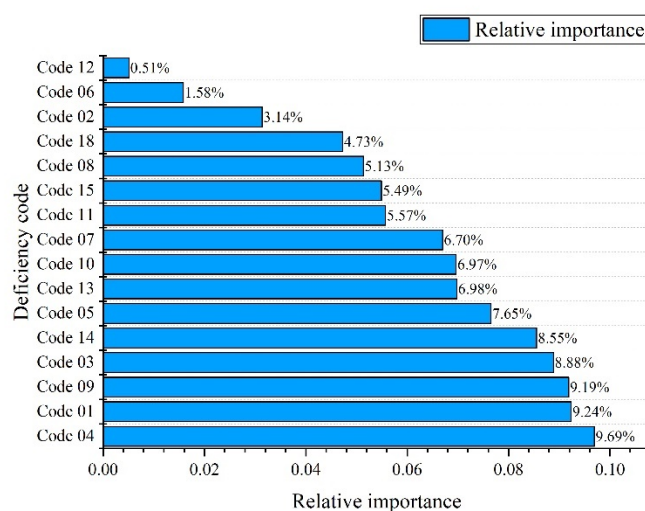

Figure 1. Relative importance of deficiency codes based on SHAP.

Figure 1 shows that code 04—emergency system contributes the most to ship detention, followed by code 01—certification & documentation, and code 09—working and living conditions. These three types of deficiency codes are very commonly detected in detained ships: among all the 714 ships with deficiency (deficiencies) under code 04, 98 are detained (detention rate at 13.73%); among all the 609 ships with deficiency (deficiencies) under code 01, 94 are detained (detention rate at 15.44%); among all the 1009 ships with deficiency (deficiencies) under code 09, 87 are detained (detention rate at 8.62%).

Meanwhile, code 06—cargo operations including equipment and code 12—dangerous goods has the least contribution to ship detention.

Although the above analysis is effective in identifying the main contributors to ship detention, it is not comprehensive due to the distribution of deficiency codes: deficiencies under some deficiency codes inherently receive more attention from the port states, and thus they are common in records no matter whether detention is observed. Consequently, their estimated contribution to ship detention can be weakened. The typical example is deficiency code 07—fire safety, which is a common detainable deficiency and is detected in 132 records among all the 144 records with detention. Meanwhile, it is also detected in 1519 records without detention, indicating that it is a popular deficiency code that are commonly detected on the inspected ships. Consequently, its universality among the inspection records heavily reduces its measured contribution to detention, and thus it only ranks at No. 9 among all deficiency codes. To alleviate this issue, we further consider the frequency of each deficiency code to be regarded as detainable deficiency in the training set when evaluating their importance. We calculate the product of such frequency and the relative importance score given by SHAP as the integrated importance for each deficiency code, which is shown in Figure 2.
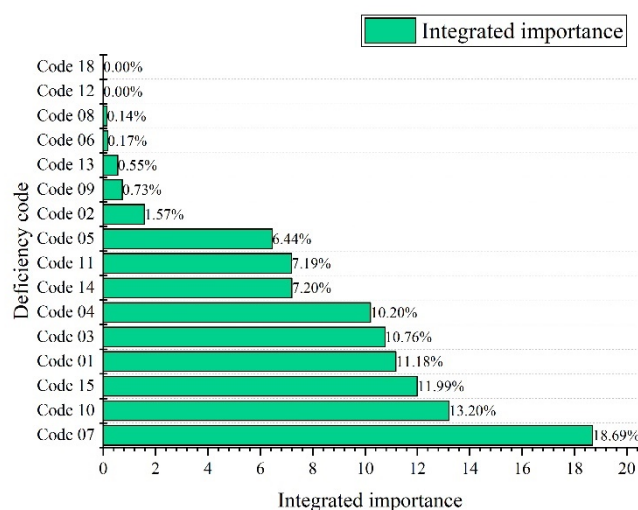

Figure 2. Integrated importance of deficiency codes based on SHAP and detainable deficiency frequency.

Figure 2 shows that when combining SHAP with detainable deficiency, code 07—fire safety becomes the determinant deficiency code of ship detention, followed by code 10—safety of navigation, code 15—ISM, and code 01—certificate & documentation, which are more in compliance with the annual reports published by the Tokyo MoU [26,38,39]. Meanwhile, as deficiency codes 12 and 18 are not served as detainable deficiency in the training set, their integrated importance is zero. While taking the deficiency importance generated by SHAP and the combination of SHAP and detainable deficiency frequency into account, port authorities can conduct more targeted ship selection and inspection to improve the overall effectiveness of PSC before an inspection if combined with studies on predicting ship deficiency condition under different deficiency codes (e.g., in [31]) and during an inspection when the deficiencies of a ship are identified. In addition, the importance scores of deficiency codes can also shed light on the design of high-risk ship selection scheme by the port states and on the inspection focus during an inspection for PSCOs. Meanwhile, ship operators can maintain and self-inspect their ships more efficiently to reduce ship detention risk in PSC.

6. Conclusions

PSC inspection is the safeguard of maritime safety, the marine environment, and seafarers' rights. The outcome of an inspection includes deficiency and detention, while the detention decision highly depends on the deficiencies detected, regarding both the number of deficiencies and their types, and the judgement of PSCOs. This study aims to predict ship detention based on the deficiencies detected and explore how each deficiency code contributes to detention by innovatively viewing detention as a type of anomaly in data. An iForest model is constructed using 2943 PSC inspection records at the Hong Kong port from 2015 to 2019 with the number of deficiencies under each deficiency code as the input and ship detention as the prediction target. The overall ROC AUC score of the iForest model is 0.82, indicating it is much more accurate than random guessing. The prediction results are then explained by SHAP, and we found that code 04—emergency system, code 01—certification & documentation, and code 09—working and living are the main contributors to ship detention. If the probability of being a detainable deficiency of a deficiency code is considered, code 07—fire safety is the determinant deficiency code of ship detention, followed by code 10—safety of navigation, code 15—ISM, and code 01—certificate & documentation.

The proposed iForest model is initially designed to be used by port authorities, and the specific practical applications including designing ship selection criteria, targeting high-risk ships, guiding onboard inspection process, and rationalizing detention decisions. If they are used by ship owners, it may not be necessarily a bad thing: as deficiency codes associated with a higher importance score are more likely to lead to detention and can thus be regarded to be more serious, it can promote ship owners and operators to pay more attention to these critical aspects on their ships to prevent them from being a substandard ship and thus being detained. Therefore, it can be expected that ship owners and operators would put more efforts to maintain their ships in better condition, especially in those critical aspects. This this way, the deterrence of PSC inspection has been exerted, as the ships' conditions are improved, and thus the navigation safety is enhanced, and the marine environment is better protected.

In addition to the prediction of ship detention in PSC, iForest based anomaly detection models can be applied to address other similar problems in maritime transport, such as ship anomaly sailing status detection and risk alarming. It can also be applied to address practical problems in other transportation modes, such as aircraft safety assessment and maintenance in air transportation and traffic accident warning in road transportation. Besides, iForest model is only sensitive to global anomaly points while is not sensitive to local anomaly points. This means that ships with overall distinct features are more easily identified, but if ships only have a few distinct features, they might not be accurately captured. It should also be noted that as iForest randomly selects features and values for node splitting, some features or many of the values of certain features might not be selected. This may reduce the reliability of the iForest model, indicating that iForest might not be suitable to address high-dimensional data with a large number of features and values. In addition, iForest is an unsupervised anomaly detection method which is initially designed to be applied to data sets without labels. Therefore, for data sets with labels available, supervised anomaly detection methods can also be used.

This study is the pioneer that develops anomaly detection model for ship detention prediction, with SHAP model used for model explanation. The prediction and explanation results are beneficial to both port states and ship operators and can contribute to enhancing the role PSC plays to guarantee maritime safety, protect the marine environment, and guarantee good living and working conditions of seafarers.

This study is the pioneer that develops anomaly detection model for ship detention prediction, with SHAP model used for model explanation. The prediction and explanation results are beneficial to both port states and ship operators and can contribute to enhancing the role PSC plays to guarantee maritime safety, protect the marine environment, and guarantee good living and working conditions of seafarers.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. M. Eiden, G. Abramowicz, W. Filipiak, D. Małyszko, J. Małyszko, K. Węcel, A framework for the quality-based selection and retrieval of open data-a use case from the maritime domain, *Electron. Markets*, **28** (2018), 219–233. https://doi.org/10.1007/s12525-017-0277-y

2. S. Wang, X. Chen, X. Qu, Model on empirically calibrating stochastic traffic flow fundamental diagram, *Commun. Transp. Res.*, **1** (2021), 100015. https://doi.org/10.1016/j.commtr.2021.100015

3. R. Yan, S. Wang, L. Zhen, G. Laporte, Emerging approaches applied to maritime transport research: Past and future, *Commun. Transp. Res.*, **1** (2021), 100011. https://doi.org/10.1016/j.commtr.2021.100011

4. C. L. Tsai, D. T. Su, C. P. Wong, An empirical study of the performance of weather routing service in the North Pacific Ocean, *Marit. Bus. Rev.*, **6** (2021), 280–292. https://doi.org/10.1108/MABR-11-2020-0066

5. V. Zisi, H. N. Psaraftis, T. Zis, The impact of the 2020 global sulfur cap on maritime CO2 emissions, *Marit. Bus. Rev.*, **6** (2021), 339–357. https://doi.org/10.1108/MABR-12-2020-0069

6. L. Wu, Y. Adulyasak, J. F. Cordeau, S. Wang, Vessel service planning in seaports, *Oper. Res.*, **(2022)** 2022. https://doi.org/10.1287/opre.2021.2228

7. *UNCTAD*, Review of maritime transport 2021, Accessed 20 December 2021. Available from: https://unctad.org/webflyer/review-maritime-transport-2021.

8. S. Wang, R. Yan, A global method from predictive to prescriptive analytics considering prediction error for "Predict, then optimize" with an example of low-carbon logistics, *Cleaner Logist. Supply Chain*, **4** (2022), 100062. https://doi.org/10.1016/j.clscn.2022.100062

9. R. Yan, S. Wang, Integrating prediction with optimization: Models and applications in transportation management, *Multimodal Transp.*, **1** (2022), 100018. https://doi.org/10.1016/j.multra.2022.100018

10. W. Yi, L. Zhen, Y. Jin, Stackelberg game analysis of government subsidy on sustainable off-site construction and low-carbon logistics, *Cleaner Logist. Supply Chain*, **2** (2021), 100013. https://doi.org/10.1016/j.clscn.2021.100013

11. *IMO*, Maritime safety, Accessed 25 February 2022. Available from: https://www.imo.org/en/OurWork/Safety/Pages/default.aspx.

12. D. Huang, S. Wang, A two-stage stochastic programming model of coordinated electric bus charging scheduling for a hybrid charging scheme, *Multimodal Transp.*, **1** (2022), 100006. https://doi.org/10.1016/j.multra.2022.100006

13. A. P. C. Chan, W. Yi, F. K. Wong, Evaluating the effectiveness and practicality of a cooling vest across four industries in Hong Kong, *Facilities*, **34** (2016), 511–534. https://doi.org/10.1108/F-12-2014-0104

14. W. Yi, Y. Zhao, A. P. C. Chan, Evaluating the effectiveness of cooling vest in a hot and humid environment, *Ann. Work Exposures Health*, **61** (2017), 481–494. https://doi.org/10.1093/annweh/wxx007

15. W. Yi, S. Wu, L. Zhen, G. Chawynski, Bi-level programming subsidy design for promoting sustainable prefabricated product logistics, *Cleaner Logist. Supply Chain*, **1** (2021), 100005. https://doi.org/10.1016/j.clscn.2021.100005

16. X. Chen, D. Z. Long, J. Qi, Preservation of supermodularity in parametric optimization: Necessary and sufficient conditions on constraint structures, *Oper. Res.*, **69** (2021), 1–12. https://doi.org/10.1287/opre.2020.1992

17. J. Zhang, D. Z. Long, R. Wang, C. Xie, Impact of penalty cost on customers' booking decisions, *Prod. Oper. Manage.*, **30** (2021), 1603–1614. https://doi.org/10.1111/poms.13297

18. R. Yan, H. Mo, S. Wang, D. Yang, Analysis and prediction of ship energy efficiency based on the MRV system, *Marit. Policy Manage.*, **2021** (2021), 1–23. https://doi.org/10.1080/03088839.2021.1968059

19. R. Yan, H. Mo, X. Guo, Y. Yang, S. Wang, Is port state control influenced by the COVID-19? Evidence from inspection data, *Transp. Policy*, **123** (2022), 82–103. https://doi.org/10.1016/j.tranpol.2022.04.002

20. Z. Sun, R. Zhang, Y. Gao, Z. Tian, Y. Zuo, Hub ports in economic shocks of the melting Arctic, *Marit. Policy Manage.*, **48** (2021), 917–940. https://doi.org/10.1080/03088839.2020.1752948

21. W. Ma, T. Lu, D. Ma, D. Wang, F. Qu, Ship route and speed multi-objective optimization considering weather conditions and emission control area regulations, *Marit. Policy Manage.*, **48** (2021), 1053–1068. https://doi.org/10.1080/03088839.2020.1825853

22. *IMO*, Procedure for port state control, 2021, Accessed 3 March 2022. Available from: https://www.register-iri.com/wp-content/uploads/A.115532.pdf.

23. S. Wang, R. Yan, X. Qu, Development of a non-parametric classifier: Effective identification, algorithm, and applications in port state control for maritime transportation, *Transp. Res. Part B*, **128** (2019), 129–157. https://doi.org/10.1016/j.trb.2019.07.017

24. L. Zhang, L. Guan, D. Z. Long, H. Shen, H. Tang, Who is better off by selling extended warranties in the supply chain: the manufacturer, the retailer, or both?, *Ann. Oper. Res.*, **2020** (2020), 1–27. https://doi.org/10.1007/s10479-020-03728-z

25. *Tokyo MoU*, List of Tokyo MoU deficiency codes, Accessed 28 October 2018. Available from: http://www.tokyo-mou.org/publications/tokyo_mou_deficiency_codes.php.

26. *Tokyo MoU*, Annual report on port state control in the Asia-Pacific region 2019, Accessed 17 July 2020. Available from: http://www.tokyo-mou.org/doc/ANN19-f.pdf.

27. R. Xu, Q. Lu, W. Li, K. X. Li, H. Zheng, A risk assessment system for improving port state control inspection, in *Proceedings of 2007 International Conference on Machine Learning and Cybernetics*, (2007), 818–823. https://doi.org/10.1109/ICMLC.2007.4370255

28. R. Xu, Q. Lu, K. X. Li, W. Li, Web mining for improving risk assessment in port state control inspection, in *Proceedings of 2007 International Conference on Natural Language Processing and Knowledge Engineering*, (2007), 427–434. https://doi.org/10.1109/NLPKE.2007.4368066

29. Z. Gao, G. Lu, M. Liu, M. Cui, A novel risk assessment system for port state control inspection, in *Proceedings of 2008 IEEE International Conference on Intelligence and Security Informatics*, (2008), 242–244.

30. R. Yan, S. Wang, J. Cao, D. Sun, Shipping domain knowledge informed prediction and optimization in port state control, *Transp. Res. Part B Methodol.*, **149** (2021), 52–78. https://doi.org/10.1016/j.trb.2021.05.003

31. R. Yan, S. Wang, K. Fagerholt, A semi-"smart predict then optimize" (semi-SPO) method for efficient ship inspection, *Transp. Res. Part B Methodol.*, **142** (2020), 100–125. https://doi.org/10.1016/j.trb.2020.09.014

32. Z. Yang, Z. Yang, J. Yin, Realising advanced risk-based port state control inspection using data-driven Bayesian networks, *Transp. Res. Part A*, **110** (2018), 38–56. https://doi.org/10.1016/j.tra.2018.01.033

33. R. Yan, S. Wang, C. Peng, An artificial intelligence model considering data imbalance for ship selection in port state control based on detention probabilities, *J. Comput. Sci.*, **48** (2021), 101257. https://doi.org/10.1016/j.jocs.2020.101257

34. S. Wu, X. Chen, C. Shi, J. Fu, Y. Yan, S. Wang, Ship detention prediction via feature selection scheme and support vector machine (SVM), *Marit. Policy Manage.*, **2021** (2021), 1–16. https://doi.org/10.1080/03088839.2021.1875141

35. F. Liu, K. Ting, Z. Zhou, Isolation forest, in *Proceedings of 2008 Eighth IEEE International Conference on Data Mining*, (2008), 413–422. https://doi.org/10.1109/ICDM.2008.17

36. F. Liu, K. Ting, Z. Zhou, Isolation-based anomaly detection, *ACM Trans. Knowl. Discovery Data*, **6** (2012), 1–39. https://doi.org/10.1145/2133360.2133363

37. S. Lundberg, S. Lee, A unified approach to interpreting model predictions, *Adv. Neural Inf. Process. Syst.*, **30** (2017).

38. *Tokyo MoU*, Annual report on port state control in the Asia-Pacific region 2017, Accessed 27 October 2018. Available from: http://www.tokyo-mou.org/doc/ANN17-f.pdf.

39. *Tokyo MoU*, Annual report on port state control in the Asia-Pacific region 2018, Accessed 23 August 2019. Available from: http://www.tokyo-mou.org/doc/ANN18-f.pdf.