Marine Policy 108 (2019) 103528

Contents lists available at ScienceDirect

Marine Policy

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

Marine piracy prediction and prevention: Policy implications

Mengjie Jin[a], Wenming Shi[b], K.-C. Lin[c], Kevin X. Li[d,*]

[a] Jiangsu Key Laboratory of Modern Logistics, Department of Logistics Management, School of Marketing and Logistics Management, Nanjing University of Finance and Economics, No.3 Wenyuan Road, Xianlin College Town, Nanjing, China

[b] Office F96, Swanson Building, Maritime and Logistics Management, Australian Maritime College, University of Tasmania, Launceston, TAS, 7250, Australia

[c] Department of Politics and International Studies, University of Cambridge, Cambridge, CB39DT, United Kingdom

[d] Ocean College, Zhejiang University, 1 Zheda Road, Zhoushan, Zhejiang, 316021, China

ARTICLE INFO

Keywords:
Piracy risk
Piracy prediction
Piracy prevention
Shipping safety
Policy

ABSTRACT

Marine piracy has traditionally been a serious threat to marine security, highlighting the great significance of piracy risk assessment and prevention to the marine industry and global economy. This study uses data on piracy attacks between 1994 and 2017 to estimate the probability of a vessel being attacked by pirates and the success rate of piracy attacks. The results of the employed binary logistic regression model show that small vessels and open registry vessels are more likely to be targeted by pirates. Further, the probability of successful boarding by pirates is higher for a vessel at berth or anchor, at night, in territorial waters and port areas, and in South America, the South China Sea, and the Strait of Malacca. Moreover, the rate of successful boarding will decreases from 99.42% to 86.71% if crews take passive measures and decreases further to 9.39% if crews take active measures. These results can not only be used by stakeholders to estimate the likelihood of piracy incidents, but can also provide information on how and the extent to which the identified factors influence piracy, helping marine stakeholders take appropriate anti-piracy measures. Moreover, these results can also be used for crew training and education to counter pirates as well as to help insurance companies set a premium rate for the different risk levels calculated by the model.

1. Introduction

Marine piracy is an organized, violent, and acquisitive crime that appeared with the emergence of the world's seaborne trade [1]. It is a complex phenomenon consisting of different criminal behaviors, modus operandi, and targets. The United Nations Convention of the Law of the Sea (UNCLOS) defines piracy as “any illegal acts of violence, detention or depredation, committed for private ends by the crew or passengers of a private ship or a private aircraft and directed: (i) on the high seas, against another ship (or aircraft), or persons or property on board that ship or aircraft; (ii) against a ship, aircraft, persons or property in a place outside the jurisdiction of any state” [2]. This definition restricts acts of piracy to the “high seas” and “outside the jurisdiction of any state”, which was used by International Maritime Organization (IMO). To overcome the limitation of the narrow definition of piracy, and to facilitate the collection of data, the IMO introduced a separate definition for ‘armed robbery against ships’ since July 2012, classified separately any reported incidents of piracy and armed robbery at sea (international or territorial waters) vis-à-vis acts of armed robbery allegedly committed in port areas, as well as attempted acts of armed robbery. This definition includes acts against vessels in port or at anchor, and regardless of whether they are inside or outside territorial waters when attacked [3]. As the motivation of maritime piracy is very different from other sea-based crimes, we adopt IMO's definition that emphasizes the economic motivation of the perpetrators, as to differentiate from maritime terrorism and other crimes at sea [4]. That is using the 1982 UNCLOS definition of piracy and the IMO definition of armed robbery, and the data provided by IMO will be used for purpose of this paper.

Modern-day pirates generally seek to profit from the confiscation of cargo and/or the holding of crew members for ransom [5]. The consequences of piracy include the loss of cargoes and vessels, ransom payments, and the death, injury, or kidnapping of seafarers, which can cause enormous losses to marine stakeholders. Taking marine oil transportation as an example, piracy has long been an inevitable risk factor. Oil tankers have become pirates' primary hijacking targets in recent years. According to the reports of IMB, the number of piracy attacks on crude oil tankers had increased from 9 in 2006 to 19 in 2017

* Corresponding author.
E-mail addresses: mengjiejin@nufe.edu.cn (M. Jin), Wenming.Shi@utas.edu.au (W. Shi), kcl35@cam.ac.uk (K.-C. Lin).
URL: http://kxli@zju.edu.cn (K.X. Li).

https://doi.org/10.1016/j.marpol.2019.103528
Received 21 September 2018; Received in revised form 7 May 2019; Accepted 9 May 2019
Available online 30 May 2019
0308-597X/

and with a peak in 2011 with 61. In 2017, the number of piracy attacks on product oil tankers and crude oil tankers were 29 and 23, ranking second and third among all vessels [6]. The Gulf of Oman and the entrance to the Persian Gulf have been identified as hot spots for Somali pirates, as they run along the navigation path of oil tankers from the Strait of Hormuz [7]. 12% of the world's oil supply passes through the Gulf of Aden, and much through Nigerian waters and the Straits of Malacca: these are the world's major piracy hubs [8,9]. The rampant piracy increases the possibility of being attacked for the tanker vessels sailing through the Strait of Malacca [10].

As the backbone of international trade and the global economy, the prediction and prevention of piracy attacks are significant for marine transport [11]. For this reason, this study performs a risk assessment by identifying the factors behind piracy attacks and assessing the likelihood of piracy attacks to find measures that can reduce and prevent risk. Three main questions are answered in this study. First, how do pirates select the vessels they plan to attack? Second, how do vessels' characteristics, environmental information, geographic factors, pirates' capacity, and crews' actions determine the success rate of piracy attacks? Third, how can the probability of piracy accidents for individual vessel are predicted?

By answering these questions, this study offers two advances over previous research. First, it helps shipping stakeholders (e.g., seafarers, shipping companies, insurance underwriters, port authorities) estimate the likelihood of piracy incidents. The second contribution of this study is to provide information on how and the extent to which the identified factors influence piracy attacks, helping marine stakeholders take preventive measures to combat piracy. In particular, the results can help insurance companies set a premium rate for marine piracy insurance.

The remainder of this paper is organized as follows. Section 2 reviews previous studies of piracy risk analysis. Section 3 presents the variables and methodology used in this study. Section 4 shows the results of the binary logistic regression models used to assess piracy risk. Section 5 discusses and concludes.

2. Literature review

The literature on piracy risk can be mainly classified into macro-level and micro-level analyses. Macro-level analysis examines the nature and extent of piracy risk from a variety of perspectives such as the cause of piracy, the impact of piracy, and changes in piracy patterns, which can advance our overall understanding of marine piracy.

Hallwood and Thomas [5] believed that the motivation of Somali pirates is for financial gain though some Somali pirates claimed that they are motivated for the 'honor' of protecting Somalia's interests; and they contended that pirates in other parts of the world where pirates significant were also motivated by financial gain. So factors such as poverty, income and socioeconomic conditions were settled as one main cause of piracy in some researches. For example, Modarress et al. [7] found that the absence of military forces, increasing marine trade, global poverty, human trafficking, and weapons smuggling all contribute to increased piracy attacks. Khondaker et al. [1] found that the root causes of marine piracy vary widely depending on geographic location, socioeconomic conditions, moral values, and political and economic stability. Daxecker and Prins [12] found that factors such as military capacity, population size, coastline length, trade volume, and state fragility are statistically related to piracy attacks. Marlow [13] paid particular attention to regulatory approaches and the impact of piracy on the shipping industry and argued that security measures must be managed effectively. By employing a static model where the pool of potential pirates is homogeneous and divides its time between fishing and pirate activities, Guha [14] found multiple equilibria between the factor of patrolling and piracy, that is, an “efficient” one with no patrolling and low piracy, a less efficient equilibrium with intermediate levels of both piracy and patrolling and a highly inefficient high-patrolling high-piracy equilibrium. Mejia et al. [15] investigated the changes in and modus operandi of acts of piracy between 1996 and 2008 and found that piracy attacks are organized and show shifting location patterns. In addition, poverty and political instability were found to be the root causes of piracy. Similarly, Modarress et al. [7] found that both the scale and the capacity of piracy had risen remarkably in recent years. Since varying pirate behaviors and piracy attack patterns make it more difficult to detect piracy attacks, it is thus essential to analyze individual attacks to predict and prevent future piracy attacks.

Although macro-level analyses are generally suitable for describing piracy crime, they rarely identify the impact factors, preventive implications, or policy responses of individual piracy attacks [16]. This information is critical for shipping stakeholders such as shipowners, ship operators, and insurers. Hence, micro-level factors such as vessel type, size, freeboard (the height of a ship's sides above the water), and voyage need to be analyzed to obtain the true picture of the nature of piracy and formulate counter-piracy suggestions [3].

Some research has found relationships between individual factors and piracy attacks. Popular factors are vessels' characteristics such as vessel type, vessel size, vessel flag, freeboard, and speed. Pirates first choose the vessels they plan to attack and many researchers found the piracy acts are not random occurrence, and it is affected by some vessel characteristics. For example, Mejia et al. [17] found that vessels under Malaysian, Singaporean, and Indian registries are more likely to face acts of piracy than other flags. They also showed that general cargo vessels, tankers, and chemical and product carriers face fewer attacks than liquid gas tankers, while containers are significantly more exposed to attack. The results of Shane and Magnuson [16] showed that the risk of an attack for larger vessels (e.g., container ships, tankers, bulk cargo ships) is 1.6 times higher than that for all other types of vessels. Kiourktsoglou and Coutroubis [18] found that Somali pirates seem to prefer to attack open-registered vessels whose navy is absent (these vessels suffer 8.1 times more attacks than those whose navy is present). Bateman and Mathai [19] focused on vessels using the Straits of Malacca and Singapore and described the relative vulnerability of commercial vessels to piracy attacks. They found that vessels with a lower freeboard and smaller crews such as tugs, gas carriers, product tankers, and fishing vessels are most at risk of attack. Further, older and smaller vessels and substandard ships tend to be hijacked by Somali pirates [3].

The success rate of piracy attack is also affected by series of ship characteristics. Pristrom et al. [20] found that bulk carriers followed by tankers and general cargo ships were attacked the most between 1994 and 2009, with ships having a low freeboard and a speed below 15 knots particularly vulnerable to piracy. Wong and Yip [21] used logit and probit models to estimate the success rate taking account of vessel type, size, flag, and status. They found that larger vessels and ro-ro vessels are safer given their higher freeboards, while vessels anchored and berthed are more vulnerable. However, vessel flag is not a significant factor in piracy attacks as pirates with commercial objectives do not target specific flags. By using a probit model, Psarros et al. [22] found that the success rate of an attack is the highest for oil tankers followed by general cargo ships, bulk carriers, container ships, and chemical tankers. Shane and Magnuson [16] suggested that larger vessels have a lower probability of experiencing a successful attack.

Environmental factors are also used to predict the probability of piracy attacks in marine transportation, using micro-level data. Pristrom et al. [23] used a Bayesian network model to analyze marine piracy incidents and found that the most significant node is the environmental condition which may lead to a likelihood for an attack as low as approximate 7.5% or up to a maximum of 21.2%. Wong and Yip [21] found that vessels sailing in winter and in the South East, Far East, Indian Subcontinent, and Americas provide a weaker defense against pirates. Huang et al. [24] calculated the probability of encountering pirates by combining the average probability of piracy and an adjustment coefficient derived from Geographic Information System (GIS). Okeahalam and Otwombe [25] found that successful attacks are most likely in poor regions that have a low military capacity, in ports or territorial waters in close proximity to a port, and when vessels are anchored. Shane and Magnuson [16] suggested that environmental factors such as night-time hours, a ship's berth status, and sailing around the African continent all significantly increase the chance of a successful attack. Bateman [3] found that the greatest concentration of piracy incidents between 2003 and 2009 was off the Horn of Africa and in the Red Sea. Somalia is a particular hotspot for piracy activities [3,13].

Some researchers use pirates' capacity and crews' actions to predict the probability of piracy attacks. Based on piracy in the Indian Ocean as a case study, Liwång et al. [26] identified piracy threat scenarios by using pirates' capability such as skiff speed, threat intent, and a ship's vulnerability, and then employed event tree analysis to describe the probability of a vessel being detected, successfully approached, and successfully boarded by pirates. They found that a risk-based approach could increase ship security by assisting in selecting risk control options. Shane and Magnuson [16] found that the presence of weapons significantly increases the chance of a successful attack, whereas the number of pirates during the attack does not. Wong and Yip [21] found that as a coordinated result of location, vessel type, season effect and vessel status, the success rate is lower if a greater number of pirates are involved in the attack. Ships' security measures taken by crews are often the first and only measure preventing criminal acts at sea, and commercial vessels must assume that they are on their own if attacked [27]. The most common measure taken by people who detect pirates is raising the alarm to alert all crews to take counter-piracy measures. By employing situational crime prevention theory, Shane and Magnuson [16] used micro factors to analyze the success rate of piracy and suggested that the proactive self-protective measures taken by crews can reduce the probability of a successful attack. Kiourktsoglou and Coutroubis [18] found a correlation between crew nationality and piracy, with Filipino crews disproportionally vulnerable to Somali pirates. Wong and Yip [21] found that the probability of a successful attack decreases as the actions by crew increase.

Binary models are the most popular method of identifying driving factors and the corresponding probability of piracy attacks, as they can estimate the probability of categorically dependent variables based on several independent variables. Consequently, this study uses a binary logistic regression model to identify and predict the probability of piracy attacks that, in this study, are divided into two processes: detected by pirates and successfully boarded by pirates. This research firstly predicts the probability of a vessel being attacked by pirates by employing factors including vessel size, vessel type, vessel flag, and geographic location. Second, the probability of a vessel being successfully attacked by pirates is quantified by factors including ships' characteristics, environmental indicators, pirates' capacity level, and crews' action level. The product of the probabilities is then used to represent the probability of a piracy accident.

3. Methodology

3.1. Probability model specifications

As suggested by previous studies, a binary logistic regression model is suitable for analyzing the relationship between a discrete target-dependent variable and a number of independent variables [21,22,25]. In model I, for each vessel $i$in our database, the dependent variable y_i takes two values: $y_i = 1$ if vessel i has been attacked by pirates and 0 otherwise. $X = (x_1, x_2 ... x_k)$ indicates a vector of the independent variables, explaining the binary outcome y_i, and $p_i = P(y_i = 1)$ is the probability of an incident occurring for tanker i. The initial binary logistic regression model can be expressed as

$$Logit\,(p_i) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k \tag{1}$$

where $Logit\,(p_i) = \ln\left(\frac{p_i}{1-p_i}\right)$ is called the logit transformation and $\frac{p_i}{1-p_i}$ is defined as the odds. That is, the log odds are predicted by a linear combination of the independent variables. As a result, the binary logistic regression model can be rewritten as

$$p_i = \frac{e^{\beta_0+\beta_1 x_1+...+\beta_k x_k}}{1 + e^{\beta_0+\beta_1 x_1+...+\beta_k x_k}} \tag{2}$$

In model II, for each vessel j that has been attacked by pirates in our database, the dependent variable y_j is 1 if vessel j has been boarded by pirates and 0 otherwise. $Z = (z_1, z_2 ... z_l)$ indicates a vector of the independent variables, explaining the binary outcome y_j, and $p_j = P(y_j = 1)$is the probability of a vessel attacked by pirates being successfully boarded by pirates. The initial binary logistic regression model can be expressed as

$$Logit\,(p_j) = \alpha_0 + \alpha_1 z_1 + ... + \alpha_l z_l \tag{3}$$

The binary logistic regression model can be rewritten as

$$p_j = \frac{e^{\alpha_0+\alpha_1 z_1+...+\alpha_l z_l}}{1 + e^{\alpha_0+\alpha_1 z_1+...+\alpha_l z_l}} \tag{4}$$

The Law of Total Probability is used to show the estimation procedure of successful boarding probability. Suppose, event A is a ship is attacked by pirates, the complementary event of A is A', that is a ship is not attacked by pirates; event B is a ship is successfully boarded by pirates. Then the probability of a ship being successfully boarded by pirates can be expressed as

$$\begin{aligned} P(B) &= P(A)P(B/A) + P(A')P(B/A') = P(A)P(B/A) + P(A') \\ &\quad * 0 = P(A)P(B/A) = p_i * p_j \end{aligned} \tag{5}$$

3.2. Data

This study built two datasets. The first comprises 52,014 samples, with 3569 vessels being attacked by pirates before October2, 017.[1] The data on vessels being attacked are based on piracy and armed robbery, as provided by the GISIS [37]. Data from GISIS (Annual) provides 7159 piracy accident records between 1994 and October 2017, and 5787 vessels are selected after dropping non-commercial vessels such as dredgers and tugs. However, some vessel information variables in the models are not provided in the GISIS. Therefore, by matching with the World Fleet Register Database of Clarkson and online information, 3569 vessels are identified for the first stage analysis of piracy attacks. The vessels without piracy attacks are provided by the World Fleet Register Database of Clarkson. The second dataset is extracted from the first dataset. After removing incomplete records, we finally identified 1736 vessels attacked by pirates, including 1074 successful boarding records and 662 unsuccessful boarding records, which were used to analyze the attack success rate.

3.3. Variables

3.3.1. Probability of piracy attacks

Mejia et al. [17] showed that piracy attacks are not randomly selected. Choosing the vessels to attack and gaining access to those vessels is thus the first step of a piracy incident. In this situation, pirates can exploit vessels' outwardly displayed signs such as vessel type and flag, which influence pirates' decisions to attack [16]. To examine how

[1] The GISIS provides 7159 piracy accident records between 1994 and October 2017, and 5787 vessels are selected after dropping non-commercial vessels such as dredgers and tugs. However, some vessel information variables in the models are not provided in the GISIS. Therefore, by matching with the World Fleet Register Database of Clarkson and online information, 3569 vessels are identified for the first-stage analysis of piracy attacks.

Table 1
List of the variables used in model I.

Variable	Measurement	Variable Type	Reference
Y	1 if incident, 0 otherwise	Dummy	
Vessel size	Gross tonnage	Continuous	
Vessel flag	1 if vessel is an open register, 0 otherwise	Dummy	Mejia et al. [17]; Kiourktsoglou and Coutroubis [18]
Vessel type			Mejia et al. [17]; Shane and Magnuson [16]
Bulk carrier	1 if vessel is a bulk carrier, 0 otherwise	Dummy	
Chemical tanker	1 if vessel is a chemical tanker, 0 otherwise	Dummy	
Container	1 if vessel is a container, 0 otherwise	Dummy	
Gas carrier	1 if vessel is a gas carrier, 0 otherwise	Dummy	
General cargo ship	1 if vessel is a general cargo ship, 0 otherwise	Dummy	
Oil tanker	1 if vessel is an oil tanker, 0 otherwise	Dummy	
PCC	1 if vessel is a PCC, 0 otherwise	Dummy	
Reefer	1 if vessel is a reefer, 0 otherwise	Dummy	
Ro-ro	1 if vessel is a roro, 0 otherwise	Dummy	
Other	1 if vessel is another type, 0 otherwise	Dummy	

Table 2
Descriptive statistics of the vessel types and flags in model I.

Variable	Number of attacks	Number of non-attacks	Total	Percentage of all piracy attacks	Percentage of this type of vessel
Vessel type					
Bulk carrier	875	10193	11068	24.52%	7.91%
Chemical tanker	549	3084	3633	15.38%	15.11%
Container	561	4611	5172	15.72%	10.85%
Gas carrier	158	1770	1928	4.43%	8.20%
General cargo ship	439	12890	13329	12.30%	3.29%
Oil tanker	830	8952	9782	23.26%	8.48%
Other type	30	3726	3756	0.84%	0.80%
PCC	30	749	779	0.84%	3.85%
Reefer	58	1298	1356	1.63%	4.28%
Ro-ro	39	1172	1211	1.09%	3.22%
All types	3569	48445	52014	100.00%	6.86%
Vessel flag					
Open registry	1567	18790	20357	43.91%	7.70%
Traditional registry	2002	29655	31657	56.09%	6.32%
All flags	3569	48445	52014	100%	6.86%

Data Source: The data on vessels being attacked are based on piracy and armed robbery, as provided by the GISIS. The vessels without piracy attacks are obtained by subtracting the number of vessels attacked by their IMO number from the total number of that class of vessel in the World Fleet Register Database of Clarkson.

pirates' vessel selection may be affected by vessels' outward signs, we collect data on vessel size, vessel flag, and vessel type, which are shown in Table 1 and summarized in Table 2.

3.3.1.1. Vessel type and vessel size. Vessel type is one of the most important characteristics of vessels. As shown in Table 2, the percentages of attacked vessels vary with vessel type according to the records in our database. As suggested by Shane and Magnuson [16]; they employed data about piracy attacks provided by the IMB Piracy Reporting Center for 2000–2012, and found larger vessels are more likely to be targeted if not considering vessel status. We use gross tonnage to measure vessel size in this research with the unit of 10,000 GT and propose the following two hypotheses:

H1. An increase in vessel size raises the probability of a piracy attack.

H2. The probability of a piracy attack varies with vessel type.

3.3.1.2. Vessel flag. The flag state of a commercial vessel is the state under whose laws the vessel is registered or licensed. There are two types of vessel registries: the traditional registry and open registry. The traditional registry is open only to ships of its own nation, while the open registry is open to foreign-owned ships. Traditional-registered vessels usually have a naval presence off the Somali Basin and in the broader area of the Gulf of Aden, while open-registered vessels have no naval presence off East Africa [18]. Table 2 shows that 7.52% of open registry vessels are attacked by pirates compared with 5.84% of traditional registry vessels. Hence, we propose the following hypothesis:

H3. Open registry vessels have a higher probability of piracy attacks.

3.3.1.3. Geographic location. Because of data availability, another important factor neglected in previous studies is the geographic location of vessels. Pirates are usually concentrated in several areas in the world because of economic, political, and geographic factors. Vessels in some areas such as Somalia and the Strait of Malacca are thus more likely to be attacked by pirates [3,13]. As a result, we propose the following hypothesis:

H4. Vessels navigating off the east coast of Africa have a higher probability of piracy attacks.

The location information of vessels that have not been attacked by pirates is difficult to obtain as their locations vary over time. Hence, we use descriptive statistics to analyze the influence of geographic location and employ an adjustment coefficient extracted from GIS data to adjust the probability of piracy attacks.

3.3.2. Probability of successful piracy attacks

After selecting the target, the pirates will approach and board the vessel. However, a successful boarding depends on many factors such as vessel type, speed, and freeboard. We use four group factors, along with the circumstances of individual attacks, to analyze the success rate of piracy attacks to propose appropriate suggestions for stakeholders to combat piracy. The four factors are vessels' characteristics, environmental factors, pirates' capacity, and crews' action, as shown in Table 3.

3.3.2.1. Vessels' characteristics. We use three variables to describe vessels' characteristics, namely freeboard, type, and status. Pirates usually board vessels by climbing the mooring ropes or anchor chain, and thus the freeboard is critical for pirates to board the vessel. Large vessels usually have a high freeboard, with a decreasing attack success rate as the size increases [21,22]. The success rate also varies with the type of vessel [16,17]. In this research, vessel type was coded by nine binary variables to categorize vessels into bulk carrier, chemical tanker, container, gas carrier, general cargo ship, oil tanker, pure car carrier

Table 3
List of the variables used in model II.

Variable	Measurement	Variable Type	Reference
Y	1 if successfully boarded, 0 otherwise	Dummy	Wong and Yip [21]; Psarros et al. [22]
Vessels' Characteristics			
Vessel size	Gross tonnage	Continuous	Wong and Yip [21]; Psarros et al. [22]
Vessel type			
Bulk carrier	1 if vessel is a bulk carrier, 0 otherwise	Dummy	Wong and Yip [21];
Chemical tanker	1 if vessel is a chemical tanker, 0 otherwise	Dummy	Psarros et al. [22];
Container	1 if vessel is a container, 0 otherwise	Dummy	Shane and Magnuson [16];
Gas carrier	1 if vessel is a gas carrier, 0 otherwise	Dummy	Okeahalam and Otwombe [25]
General cargo ship	1 if vessel is a general cargo ship, 0 otherwise	Dummy	
Oil tanker	1 if vessel is an oil tanker, 0 otherwise	Dummy	
PCC	1 if vessel is a PCC, 0 otherwise	Dummy	
Reefer	1 if vessel is a reefer, 0 otherwise	Dummy	
Ro-ro	1 if vessel is a roro, 0 otherwise	Dummy	
Other			
Freeboard		Continuous	Liwång et al. [26]
Vessel status	1 if vessel is steaming, 0 otherwise	Dummy	Wong and Yip [21]; Shane and Magnuson [16]; Okeahalam and Otwombe [25]
Environmental Factors			
Light conditions	1 if the vessel is attacked by pirates in daytime (6am–6pm)	Dummy	Liwång et al. [26]; Shane and Magnuson [16]
Geographic area of attack			
East Africa	1 if the attack was launched from East Africa, 0 otherwise	Dummy	Psarros et al. [22];
Indian Ocean	1 if the attack was launched from the Indian Ocean, 0 otherwise	Dummy	Wong and Yip [21];
Strait of Malacca	1 if the attack was launched from the Strait of Malacca, 0 otherwise	Dummy	Shane and Magnuson [16];
South America	1 if the attack was launched from South America, 0 otherwise	Dummy	Okeahalam and Otwombe [25]
South China Sea	1 if the attack was launched from the South China Sea, 0 otherwise	Dummy	
West Africa	1 if the attack was launched from West Africa, 0 otherwise	Dummy	
Area of attack			
Port areas	1 if the attack was located in port areas, 0 otherwise	Dummy	Okeahalam and Otwombe [25]
Territorial waters	1 if the attack was located in territorial waters, 0 otherwise	Dummy	
Pirates' Capability	The capacity can be measured by the principal component of these ordinal variables	Discrete	
Number of pirates	Level 1: 1–4 people; Level 2: 5–10 people Level 3: more than 10 people	Discrete	Psarros et al. [22]; Wong and Yip [21]; Shane and Magnuson [16]
Types of weapons used	Level 1: other weapons; Level 2: knives; Level 3: guns; Level 4: rocket-propelled grenades	Discrete	Wong and Yip [21]; Shane and Magnuson [16]
Number of pirates' boats	Level 1: 1 boat; Level 2: 2–5 boats; Level 3: More than 5 boats	Discrete	
Crews' Action	Level 1: No action Level 2: Passive measures (alarm raised, ship security alert system engaged, messages sent to other people; crew assembled and vigilant) Level 3: Anti-piracy measures (use hoses, lights, flares, evasive maneuvers, and increasing speed)	Discrete	Psarros et al. [22]; Shane and Magnuson [16]

(PCC), reefer, ro-ro, and other types. Another important factor is vessel speed. Some shipowners assume that a ship can be made almost impossible to board with a high speed and high freeboard [26]. We use the status of the vessel to represent vessel speed, which is coded as a binary variable (1 if the vessel is steaming, 0 if the vessel is berthed or anchored). Based on these analyses, we propose the following hypotheses:

H5. The probability of successful boarding by pirates decreases as vessel freeboard increases.

H6. The probability of successful boarding by pirates varies by vessel type.

H7. The probability of successful boarding by pirates is smaller when a vessel is steaming.

3.3.2.2. Environmental factors. We use three variables to describe environmental factors, namely light conditions, geographic area of attack, and location of attack. Shane and Magnuson [16] found a significant relationship between a successful attack and night-time hours because it is difficult to detect pirates at night. We use a binary variable to represent light conditions, coded as 1 if the attack is between 6am and 6pm. The success rate also varies with the geographic area of attack because of different geographic environments and pirates' experience. To examine the impact of the geographic area of attack on success rates, we identify the following seven areas based on attack frequencies: East Africa, the Indian Ocean, the Strait of Malacca, South America, the South China Sea, West Africa, and other areas.[2] The locations of piracy attacks are divided into three categories: international waters, territorial waters, and port areas in the GISIS. Therefore, we take international waters as a benchmark and use two binary variables to represent the location of attacks. Based on these

[2] The geographic areas of attacks are divided into 17 areas in the GISIS. Therefore, we recoded them into seven areas, as the numbers of piracy attacks in these six areas are far greater than those in other areas.

analyses, we propose the following hypotheses:

H8. The probability of successful boarding in daytime is lower than at night.

H9. The probability of successful boarding by pirates varies with the geographic area of attack.

H10. The probability of successful boarding in international waters is lower than that in port areas and territorial waters.

3.3.2.3. Pirates' capability. Pirates' capability, an important factor for assessing the probability of successful boarding, can be described as the number of pirates, the number of pirates' boats, and types of weapons and then grouped by three discrete variables [16,21,22]. As a result, pirates' capability is expressed as an index, which is the principal component of the three variables. We propose the following hypothesis:

H11. The probability of successful boarding by pirates increases as pirates' capability increases.

3.3.2.4. Crews' action. Once pirates are detected, the master and crews will take actions to defend themselves from attack. These actions can be divided into three levels: Level 1: No action; Level 2: passive anti-piracy measures (alarm raised, ship security alert system engaged, messages sent to other people, crew assembled and vigilant); and Level 3: Active anti-piracy measures (use hoses, lights, flares, evasive maneuvers and increasing speed). We propose the following hypothesis:

H12. The probability of successful boarding by pirates decreases as the level of crews' action increases.

4. Empirical results

4.1. Model I—Piracy choice model

The first model aims to evaluate the factors behind piracy attacks and investigate pirates' likely behavior when selecting target vessels. By using the maximum likelihood estimation method in STATA 14.0 software, Equation (2) can be estimated. Table 4 presents the results. As previously stated, the coefficients provide information on whether changes in the given independent variables increase or decrease the probability of being selected by pirates, while the marginal probability provides further information on the extent of the changes in probability.

We can see that vessel size, vessel flag and vessel type all significantly affect the probability of piracy incidents. In particular, the impact of vessel size on the probability of being attacked by pirates is negative; that is to say, the larger size, the lower is the likelihood of being targeted by pirates. Hence, H1 is rejected. Large vessels usually have a higher freeboard and speed as well as more crew members, which make it more difficult for pirates to board. For this reason, pirates are less likely to choose larger vessels as targets, especially pirates with poor capability. The marginal probability result suggests that the probability decreases by 0.001% as vessel size increases by 1%. As indicated by the results for vessel type, the probability of piracy attacks for vessels in order from high to low is chemical tankers, containers, oil tankers, gas carriers, bulk carriers, reefers, PCCs, general cargo ships, and ro-ro, and H2 is accepted. Our results also show that vessels under open registries are more likely to encounter pirates than those under traditional registries. The reason for this difference might be that open registries are more appealing to pirates as they have no naval presence off risky areas such as East Africa [18]. Hence, H3 is supported. In addition, open-registered vessels usually enjoy a more relaxed regulatory structure and fewer precautions, making it easier for pirates to board. Therefore, open registry vessels provide an opportunity for piracy attacks due to relaxed regulations on the part of flag states, allowing serious violation and crimes in unprotected oceans [7].

Another important factor affecting the probability of piracy attacks is the geographic location of vessels. Following Huang et al. [24]; we use an adjustment coefficient extracted from GIS data, an effective tool for spatial analysis and the visualization of geographic data, to adjust the probability calculated by Model I. We drop the area of land when calculating the number of grids because vessels do not sail on land. The world map is divided into 12 × 6 (72) grids in ArcGIS 2015 (Fig. 1), with 70.8% (361 million square kilometers) of the Earth's surface covered by oceans [28]. That is, oceans account for 50.976 of the grids (72 × 70.8%). By comparing the ArcGIS 2015 map with the Automatic Identification System map on the Ship Finder website, we find that vessels mainly navigate among ocean areas between 60S and the Arctic circle. Hence, we should further drop the ocean area within the Arctic Circle and South Ocean to obtain the most accurate adjustment coefficients. The number of grids in the Arctic Ocean area is 6.4[3] and the number of grids in the South Ocean is 7.059.[4] As a result, the total number of navigation grids N is 37.517 (50.976–6.4-7.059).

In the considered dataset, 5779 of the piracy attack records include the geographic location. The average number of piracy attacks in each grid $\overline{N}$ is 5779/37.5 = 154.107. Accordingly, as shown in Table 5, the average number of piracy accidents in each area $\overline{n}$ can be calculated by dividing the number of piracy attacks by the number of grids. Then, the adjustment coefficient of area i is defined as $B_i = \overline{n}/\overline{N}$, and the adjusted probability is defined as $P_2 = P_1 * B_i$, where P_1 is the probability calculated using Model I.

Table 4
Results from Model I—Piracy choice model.

Variable	Coefficient (Standard Error)	Marginal Probability
Constant	-4.85^{***} (0.18)	
Vessel size	-0.01^{*} (0.01)	−0.001
Vessel flag	0.08^{**} (0.04)	0.004
Vessel type		
Bulk carrier	2.38^{***} (0.19)	0.25
Chemical tanker	3.10^{***} (0.19)	0.47
Container	2.74^{***} (0.19)	0.37
Gas carrier	2.43^{***} (0.20)	0.33
General cargo ship	1.45^{***} (0.19)	0.11
Oil tanker	2.47^{***} (0.19)	0.27
PCC	1.63^{***} (0.26)	0.17
Reefer	1.72^{***} (0.23)	0.19
Ro-ro	1.44^{***} (0.25)	0.14

Note:
* indicates statistical significance at the 10% level;
** indicates statistical significance at the 5% level;
*** indicates statistical significance at the 1% level.

4.2. Model II—Successful boarding model

As shown in Equation (4), the dependent variable is the probability of a vessel being successfully boarded by pirates, as estimated and reported in Table 6.

Regarding the vessels' characteristics variables, first, an increase (1%) in vessel freeboard is associated with a significant decrease (0.03%) in the probability of being successfully boarded by pirates. Hence, H5 is supported. Second, as suggested by the lack of significance

[3] The ocean area in the Arctic Ocean is approximately 16 million square kilometers, accounting for 16/30 of the Arctic Ocean (http://arctic.ru/geographics/). Hence, the number of grids in the Arctic Ocean area is 6.4 (12 × (16/30)).

[4] The area of the South Ocean is about 20 million square kilometers and that of Antarctica is about 14 million square kilometers. Thus, the number of grids in the South Ocean is 7.059 (12 × (20/34)).

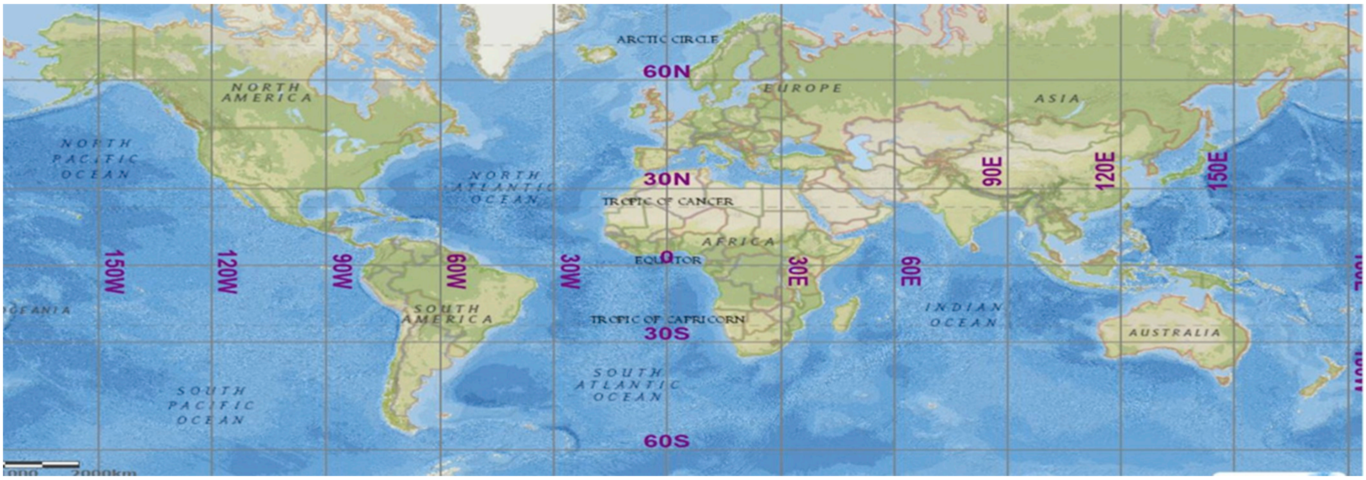

Fig. 1. ArcGIS 2015 map.

Table 5
Adjustment coefficients extracted from the GIS data.

	East Africa	Indian Ocean	Strait of Malacca	South America	South China Sea	West Africa	Other Areas	Total
Number of piracy attacks	973	957	578	550	1776	859	86	5779
Number of grids	3	5	0.5	7	0.5	5	16.5	37.5
	324.33	191.40	1156	78.57	3552	171.80	5.21	154.11
B_i	2.10	1.24	7.50	0.51	23.05	1.11	0.03	

Table 6
Results from model II—Successful boarding model.

Variable	Coefficient (Standard Error)	Marginal Probability	Variable	Coefficient (Standard Error)	Marginal Probability
Constant	8.72*** (0.84)		**Environmental Factors**		
Vessels' Characteristics			Light conditions	−0.68*** (0.21)	−0.16***
Vessel freeboard	−0.12** (0.05)	−0.03**	Geographic area of attack		
Vessel status	−0.90*** (0.28)	−0.20***	Indian Ocean	0.44 (0.33)	0.09***
Vessel type			Strait of Malacca	1.17*** (0.38)	0.21***
Bulk carrier	−0.41 (0.49)	−0.09	Other area	2.93** (1.33)	0.31***
Chemical Tanker	−0.25 (0.50)	−0.05	South America	1.25*** (0.41)	0.22***
Container	−0.50 (0.50)	−0.12	South China Sea	1.19*** (0.31)	0.23***
Gas carrier	−0.50 (0.61)	−0.12	West Africa	0.85*** (0.32)	0.17***
General cargo ship	−0.11 (0.51)	−0.02	Area of attack		
Oil tanker	0.08 (0.49)	0.02	Port areas	0.43 (0.34)	0.09
PCC	−0.62 (1.55)	−0.15	Territorial waters	0.71*** (0.26)	0.15***
Reefer	−0.36 (0.83)	−0.08	**Pirates' Capability**	−0.24*** (0.09)	−0.05***
Ro-ro	0.34 (1.61)	0.07	**Crews' Action**	−3.03*** (0.19)	−0.66***

Note:
** indicates statistical significance at the 5% level;
*** indicates statistical significance at the 1% level.

of the vessel type variables, we reject H6, indicating that vessel type is not important for determining the piracy attack success rate. This is maybe because when vessels are attacked by pirates, they are vulnerable to an attack irrespective of their type. The evidence in Table 7 also supports this result. Specifically, bulk carriers (25.63%), oil tankers (22.58%), containers (15.67%), chemical tankers (14.40%), and general cargo ships (11.18%) are most likely to be attacked, but the success rates of piracy attacks are very close (about 60%). Third, the probability of successful boarding is lower when a vessel is steaming compared with at anchor or berthed, suggesting that H7 is supported. Generally speaking, when vessels are underway, they are less vulnerable to piracy because of their speed. Indeed, it is particularly dangerous for a small craft to attempt to get alongside a large vessel traveling at its normal operational speed [19]. Hence, more attention needs to be paid to anchored or berthed vessels, as they are more vulnerable to piracy attacks. The GISIS website provides a case study of pirates attacking a berthed tanker. On August 5, 2016, four robbers armed with long knives boarded the forecastle of a berthed tanker in West Africa. The duty pumpman on routine rounds was taken hostage and threatened. The robbers submerged two hoses into the forward tank dome and commenced stealing the cargo. Once the cargo was filled into their boat, the robbers released the pumpman and fled.

In terms of environmental factors, first, the probability of being successfully boarded by pirates is lower when an attack occurs in the daytime, indicating that H8 is supported. This can probably be attributed to the fact that visibility at night is worse, making it difficult to detect pirates and take actions against them. As reported in Table 7, most piracy attacks occur at night (71.14%) with a higher success rate of 75.63%. Second, the significantly positive coefficients of other areas, South America, the South China Sea, the Strait of Malacca, and West Africa show that the probabilities of successful boarding in these areas are higher than those in East Africa, suggesting that H9 is supported. Table 7 shows that piracy attacks are most likely to be successful in other areas (92.86%), followed by South America (86.08%), the Strait

Table 7
Statistics on piracy data.

Variable	Boarded	Not Boarded	Success rate	Percentage of piracy attacks
Y	1074	662	61.87%	
Vessel type				
Bulk carrier	285	160	64.04%	25.63%
Chemical tanker	152	98	60.80%	14.40%
Container	174	98	63.97%	15.67%
Gas carrier	44	23	65.67%	3.86%
General cargo ship	121	73	62.37%	11.18%
Oil tanker	224	168	57.14%	22.58%
PCC	2	9	18.18%	0.63%
Reefer	22	6	78.57%	1.61%
Ro-ro	6	5	54.55%	0.63%
Other	44	22	66.67%	3.80%
Vessel status				
Steaming	289	555	34.24%	48.62%
At anchor	785	107	88.00%	51.38%
Light conditions				
Daytime	140	361	27.94%	28.86%
Night-time	934	301	75.63%	71.14%
Area of attack				
Port areas	565	77	88.01%	36.98%
Territorial waters	350	97	78.30%	25.75%
International waters	159	488	24.57%	37.27%
Geographic area of attack				
East Africa	43	320	11.85%	20.91%
Indian Ocean	150	125	54.55%	15.84%
Strait of Malacca	137	27	83.54%	9.45%
South America	136	22	86.08%	9.10%
South China Sea	395	78	83.51%	27.25%
West Africa	200	89	69.20%	16.65%
Other Area	13	1	92.86%	0.81%
Pirates' Capability				
Number of pirates				
Level 1: 1–4 people;	634	193	76.66%	47.64%
Level 2: 5–10 people;	386	358	51.88%	42.86%
Level 3: more than 10	54	111	32.73%	9.50%
Types of weapons used				
Level 1: other weapons;	107	63	62.94%	9.79%
Level 2: knives;	613	37	94.31%	37.44%
Level 3: guns;	347	432	44.54%	44.87%
Level 4: rocket-propelled grenades	7	130	5.11%	7.89%
Number of pirates' boats				
Level 1: 1	993	423	70.13%	81.57%
Level 2: 2–5	79	221	26.33%	17.28%
Level 3: more than 5	2	18	10.00%	1.15%
Crews' Action				
Level 1	171	1	99.42%	9.91%
Level 2	848	130	86.71%	56.34%
Level 3	55	531	9.39%	33.76%

of Malacca (83.54%), the South China Sea (83.51%), West Africa (69.20%), the Indian Ocean (54.55%), and East Africa (11.85%), while the majority of piracy attacks occur in the South China Sea (27.25%), East Africa (20.91%), West Africa (16.65%), and the Indian Ocean (15.84%). In this regard, we should pay more attention to the South China Sea and West Africa because of the large number of piracy attacks and high success rates in these areas. The probability of successful boarding in East Africa is much smaller, although the number of attacks in this area is high (20.91%). This could be because masters and crews usually take great precautions when vessels pass through piracy-prone areas in East Africa such as Somalia, which makes them less vulnerable to piracy attacks. Moreover, many vessels are protected by naval escorts from their respective countries when navigating in Western Indian/East African areas, such as NATO and EU NAVFOR. Further, a number of countries operate only under their own country's mandate such as Russia, Iran, and China [23]. Third, vessels are more likely to be successfully boarded in territorial waters compared with international waters, indicating that H10 is supported. Table 7 shows that 36.98% of piracy attacks occur in port areas and 25.75% in territorial waters, below the share of international waters (37.27%). However, the success rates in port areas (88.01%) and territorial waters (78.30%) are much larger than that in international waters (24.58%). One possible explanation for this finding is that vessels usually sail at a high speed in international waters, which makes it difficult to board them, while it is easier to board vessels when they stop at ports or slow in territorial waters. Another possible explanation is the coordinated presence of navies and coast guards in international waters.

Turning to pirates' capability, it is surprising that the probability of successful boarding decreases with an increase of pirates' capability, suggesting that H11 is rejected. After considering other factors, it is modus operandi of pirates rather than pirates' capability result in the less likely successful attack. For example, a hijacking of a vessel underway may require more pirates but be less successful, whereas a sneak attack on a vessel at anchor may involve fewer numbers but a higher success rate. By analyzing pirate numbers and their weapons and equipment, we find that most pirates work in small groups (1–4 people, 47.64%) or medium-sized groups (5–10 people, 42.86%). Moreover, 47.24% of pirates are armed with simple weapons such as knives and axes and 44.87% are armed with guns. In most cases, the intention of pirates is not to hijack a ship; rather, they look for opportunities to board vessels secretly to steal cash, the crew's valuables, the ship's stores, and other goods. Accordingly, it is difficult to detect these pirates because they act surreptitiously on such a small scale. On the contrary, if pirates act in large scale, crews can easily detect them even from distance, and crews can take actions in advance to counter piracy, which make it more difficult for pirates to board the vessels. These reasons all lead to a decrease in the success rate when pirates' capability increases.

Finally, as shown in Table 6, the crew's action has a significantly negative impact on the success rate of piracy attacks, indicating that active measures can combat piracy effectively. This result is also supported by the statistical evidence in Table 7. The rate of successful boarding decreases from 99.42% to 86.71% if crews take passive measures and further decreases to 9.39% if crews take active measures. For example, assembling the crew is a common measure to resist pirates, as the crew may be visible from the water, which gives the attacking pirates the impression that the crews are well-organized and vigilant and thus able to resist the pirates' efforts [16]. The GISIS website provides a case study of successfully repelling pirates as a result of active measures. On June 22, 2009, four skiffs with pirates armed with guns and rocket-propelled grenades chased a general cargo ship named BOLAN in East Africa and opened fire on it. The master increased the sailing speed, made feasible evasive maneuvers, contacted coalition warships, and released dunnage bundles along with sharply cut empty drums. Crews threw sawdust at the pirates to reduce their visibility. Later, coalition warships stopped the skiffs and confiscated the pirates' weapons.

4.3. Results of the model tests

Table 8 reports diagnostic information to assess model performance. First, the goodness of fit of the aforementioned models is evaluated by using the likelihood ratio test [29]. As shown in Table 8, the likelihood ratio statistics of Model I and Model II are both significant at the 1% level, which indicate that the models developed in this study fit the data well. Second, this research calculates the number of observations required taking statistical power into account. The guideline proposed by Ref. [30] is used to calculate the minimum number of cases to provide

Table 8
Test results of the models.

Test	Model I	Model II
Likelihood Ratio Test	1199.89***	1361.76***
Pseudo R^2	0.05	0.59
Correctly Classified	93.13%	90.03%
AUC	0.65	0.94
Number of Observations Required	1603	372
Number of Observations	51985	1736

Note:
*** indicates statistical significance at the 1% level.

effective power, as shown in Table 8. The number of observations in Model I and Model II are well above the minimum number of observations required, illustrating the effect of sample size on power. Third, as suggested by Hanley and McNeil [31] and Daxecker and Prins [12]; the correctly classified rate and Receiver Operating Characteristic (ROC) plots can be used to determine the overall predictive power of models, which are also considered in this study. More specifically, the correctly classified rate is the percentage of cases correctly classified by a model using threshold of 0.5, while the ROC plots can be developed by examining the area under the ROC curve (AUC). Here, the value ranges between 0.5 and 1 with larger AUC values indicating better predictive power [32]. Table 8 shows that both models perform well with correctly classified rates of 93.13% and 90.03%, respectively. The estimated AUC values are 0.65 for Model I and 0.94 for Model II. There is no absolute value for the AUC value to evaluate the quality of a classifier, as it depends on the context, and no consensus about an adequate level of predictive validity exists in the field of marine risk assessment. Hence, we use the conservative interpretation of AUC values (i.e., below 0.60: low accuracy; 0.60–0.70: marginal accuracy; 0.70–0.80: modest accuracy; 0.80–0.90: moderate accuracy; and 0.90 and above: high accuracy) proposed by Sjöstedt and Grann [33] to classify the accuracy of a diagnostic test. The results show that Model II has high accuracy, which means its explaining variable can identify the probability of successful boarding during piracy attacks. Model I has marginal accuracy, which may be due to the limited explaining variables; indeed, just three types of variables are included in Model I because of limited information.

A good classifier not only has high prediction accuracy in training tests, but also performs well when applied to other samples. Hence, by referring to Kohavi [34] and Bengio and Grandvalet [35]; the K-fold cross-validation test can be applied to estimate how accurately Model II would perform in practice. Under K-fold cross-validation, the dataset is randomly split into K mutually exclusive subsets with a proportion 1/K of the samples removed for testing purposes [34,35]. The advantages of K-fold cross-validation compared with conventional validation is that it can overcome the problem of lacking sufficient data to partition them into separate training and test sets without losing significant modeling or testing capability [36]. As shown in Table 9, the correctly classified rates of the five testing sets are 87.32%, 91.48%, 87.92%, 90.99%, and 89.46%, which indicates that Model II performs well in predicting successful piracy attacks.

Table 9
Predictive performance of Model II.

Statistic	S1	S2	S3	S4	S5
Number of training sets	1397	1372	1380	1381	1404
%Correct	91.05%	89.58%	90.43%	89.79%	90.10%
Number of testing sets	339	364	356	355	332
%Correct	87.32%	91.48%	87.92%	90.99%	89.46%
Pseudo R^2	0.61	0.58	0.60	0.58	0.58

5. Discussion and conclusion

This study explores piracy risk prediction and prevention by estimating the probability of a vessel being attacked by pirates and the success rate of such an attack. The results can not only be used by stakeholders to estimate the likelihood of piracy attacks, but can also help them take appropriate anti-piracy measures. In particular, the results can help insurance companies set a premium rate for marine piracy insurance. First, this study uses data on piracy attacks between 1994 and 2017 and the world merchant fleet to predict the probability of a vessel being attacked by pirates as well as that of a successful attack given various factors associated with attack. The results of Model I show that piracy attacks are not random. Small vessels are more likely to be targeted by pirates as they have a low freeboard and slow speed, thus making them relatively easy to board. Tankers, containers, and bulk carriers are more likely to be attacked by pirates because of their perceived value and potential ransom value [16]. In particular, chemical tankers and oil tankers have a low freeboard and are therefore easier to board [17]. In addition, vessels under open registries are more likely to face acts of piracy than those under traditional registries as they lack any naval presence off risky areas. However, because naval missions off Yemen and Somalia are expensive, shipowners need to consider the costs, risks, and benefits to make the best decision. By contrast, vessels in the South China Sea, the Strait of Malacca, and East Africa are more likely to counter piracy attacks. These results can help marine organizations, governments, military forces, shipowners, ship operators, and shipping companies better understand piracy attacks and provide information for them to avoid piracy risk.

As a piracy attack may not necessarily lead to boarding, it is useful to analyze the factors affecting the probability of a successful piracy attack. The factors in the proposed model include vessels' characteristics such as freeboard, status, and vessel type; environmental factors such as light conditions, geographic area of attack, and area of attack; pirates' capability; and the measures taken by the crew. The results show that the probability of successful boarding by pirates is higher for vessels berthed or anchored, at night, and in territorial waters. Geographically, piracy attacks are most likely to be successful in other areas (92.86%), followed by South America (86.08%), the Strait of Malacca (83.54%), and the South China Sea (83.51%). In this regard, the South China Sea and West Africa need to be paid great attention in view of the large number of piracy attacks and high success rate in these areas. The influence of crews' actions is as expected: the probability of successful boarding decreases as the level of crews' action increases. The rate of successful boarding falls from 99.42% to 86.71% if crews take passive measures and further decreases to 9.39% if crews take active measures. In addition, the success rate of low-capacity pirates' attack is higher because of the difficulties detecting small surreptitious groups of pirates. These results not only can be used to estimate the likelihood of a successful piracy attack, but can also provide information for marine stakeholders to take appropriate anti-piracy measures as well as for crew training and education to counter piracy.

An example is that more attention needs to be paid to anchored or berthed vessels, vessels at night, and vessels in territorial waters and port areas, as these are more vulnerable to piracy attacks. Crews should thus maintain vigilance when vessels navigate at night, in territorial waters and port areas, and when anchored or berthed, especially in risky areas that have a large number of piracy attacks and high success rates such as the Strait of Malacca and the South China Sea. Further, given that the success rate of a piracy attack depends heavily on the actions of crew members according to our results, the shipping industry could do more to help counter piracy by ensuring that crews are adequate in terms of size, well trained, and follow best practice guidelines [3].

Finally, although our study provides an in-depth understanding of piracy risk, there are still some limitations. First, owing to data unavailability, factors such as vessel speed, weather and patrolling were not considered in the probability models, although these factors are also important to piracy, for example, coordinated patrolling of Straits of Malacca by Singapore, Malaysia and Indonesia from 2004 reduced the probability of pirate attacks to 0.00019 in 2005 [8]. On one hand, the lack of data is because traffic and weather information is missing in many records in the GISIS database and it is difficult to obtain weather records when vessels are attacked. On the other hand, the lack of data from vessels not attacked by pirates is because the research sample in this study consists of vessels sailing around the world, and not in a specific area, making it difficult to choose the research period and obtain real-time data from all vessels, as traffic and weather information are time-varying. In future research, we will thus employ more advanced methods for predicting the probability to compensate for this lack of data. Second, the geographical areas used in this research are broad, i.e. the geographic areas of attacks are divided into several areas in the GISIS of IMO, which cannot get deep and specific results about locations of piracy attacks. For example, in the South China Sea, different causes and effects apply to a vessel at anchor in port in Indonesia to those to a vessel loitering in the eastern approaches to Singapore Strait, but both areas are classified as the South China Sea. Another example is that the IMO includes attacks in Singapore Strait within the Malacca Strait, resulting that we cannot differentiate between attacks in Singapore Strait and Malacca Strait. A potential way to overcome this limitation is to use each incident's Latitude and Longitude information, which points out the future research direction. Third, while our findings can be used for piracy prediction and prevention, neither the root causes and motivation of piracy attacks nor the behavioral patterns of pirates are discussed in this research. As pirates' behavior and piracy patterns vary over time and geographic areas, it is necessary to conduct a more comprehensive analysis to better protect against future piracy attacks.

Declarations of interest

None.

Acknowledgements

This study is supported in part by the Korea Foundation; Zhejiang University Education Foundation Global Partnership Fund; the Ministry of Education in China of Humanities and Social Science Youth Foundation Project under Grant No. 19YJC630242.

References

[1] A.N. Khondaker, S.M. Rahman, R.A. Khan, Dynamics of piracy in maritime transportation, J. Transportat. Secur. 6 (3) (2013) 193–207.
[2] UNCLOS, Article 101 of the united nations convention of the law of the Sea, [online] http://www.un.org/depts/los/convention_agreements/texts/unclos/unclos_e.pdf (1982) , Accessed date: 31 July 2018.
[3] S. Bateman, Maritime piracy in the Indo-Pacific region–ship vulnerability issues, Marit. Policy Manag. 37 (7) (2010) 737–751.
[4] D.R. Dillon, Maritime piracy: defining the problem, SAIS Rev. Int. Affairs 25 (1) (2005) 155–165.
[5] C. Paul Hallwood, Thomas J. Miceli, Maritime Piracy and its Control: an Economic Analysis, Palgrave Macmillan, New York, 2015.
[6] ICC-IMB, ICC International Maritime Bureau. Piracy and Armed Robbery against Ships Report for the Period 1 January-31 December 2017, (2018).
[7] B. Modarress, A. Ansari, E. Thies, The effect of transnational threats on the security of Persian Gulf maritime petroleum transportation, J. Transportat. Secur. 5 (3) (2012) 169–186.
[8] B. Guha, A.S. Guha, Pirates and traders: some economics of pirate-infested seas, Econ. Lett. 111 (2) (2011) 147–150.
[9] S. Zhu, X. Fu, A.K. Ng, M. Luo, Y.E. Ge, The environmental costs and economic implications of container shipping on the Northern Sea Route, Marit. Policy Manag. 45 (4) (2018) 456–477.
[10] S. Wang, D. Yang, J. Lu, A connectivity reliability-cost approach for path selection in the maritime transportation of China's crude oil imports, Marit. Policy Manag. 45 (5) (2018) 567–584.
[11] UNCTAD, Review of maritime transport, United Nations Conference on Trade and Development 2015, 2015.
[12] U.E. Daxecker, B.C. Prins, The new barbary wars: forecasting maritime piracy, Foreign Pol. Anal. 11 (1) (2015) 23–44.
[13] P.B. Marlow, Maritime security: an update of key issues, Marit. Policy Manag. 37 (7) (2010) 667–676.
[14] B. Guha, Pirates and fishermen: is less patrolling always bad? J. Econ. Behav. Organ. 81 (1) (2012) 29–38.
[15] M.Q. Mejia Jr., P. Cariou, F.C. Wolff, Piracy in Shipping. The Blackwell Companion to Maritime Economics, (2012), pp. 346–351.
[16] J.M. Shane, S. Magnuson, Successful and unsuccessful pirate attacks worldwide: a situational analysis, Justice Q. JQ 33 (4) (2016) 682–707.
[17] M.Q. Mejia Jr., P. Cariou, F.C. Wolff, Is maritime piracy random? Appl. Econ. Lett. 16 (9) (2009) 891–895.
[18] G. Kiourktsoglou, A.D. Coutroubis, Is Somali piracy a random phenomenon? WMU J. Maritime Affairs 11 (1) (2012) 51–70.
[19] S. Bateman, J. Ho, M. Mathai, Shipping patterns in the Malacca and Singapore straits: an assessment of the risks to different types of vessel, Contemp. S. Asia (2007) 309–332.
[20] S. Pristrom, K.X. Li, Z. Yang, J. Wang, A study of maritime security and piracy, Marit. Policy Manag. 40 (7) (2013) 675–693.
[21] M.C. Wong, T.L. Yip, Maritime piracy: an analysis of attacks and violence, Int. J. Shipp. Transp. Logist. (IJSTL) 4 (4) (2012) 306–322 4.
[22] G.A. Psarros, A.F. Christiansen, R. Skjong, G. Gravir, On the success rates of maritime piracy attacks, J. Transportat. Secur. 4 (4) (2011) 309.
[23] S. Pristrom, Z. Yang, J. Wang, X. Yan, A novel flexible model for piracy and robbery assessment of merchant ship operations, Reliab. Eng. Syst. Saf. 155 (2016) 196–211.
[24] D.Z. Huang, Y. Li, H. Hu, Application of Geographic information system to calculate the probability of piracy occurrence, Transportation Information and Safety (ICTIS), 2015 International Conference on IEEE, 2015, pp. 754–759.
[25] C. Okeahalam, K. Otwombe, Socioeconomic development and the risk of maritime piracy, J. Transportat. Secur. 9 (3–4) (2016) 125–160.
[26] H. Liwång, J.W. Ringsberg, M. Norsell, Quantitative risk analysis–Ship security analysis for effective risk control options, Saf. Sci. 58 (2013) 98–112.
[27] J. Kraska, B. Wilson, Fighting pirates: the pen and the sword, World Pol. J. 25 (4) (2008) 41–52.
[28] Physical Geography, [online] http://www.physicalgeography.net/fundamentals/8o.html (2006) , Accessed date: 31 July 2018.
[29] B.W. Turnbull, L. Weiss, A likelihood ratio statistic for testing goodness of fit with randomly censored data, Biometrics (1978) 367–375.
[30] P. Peduzzi, J. Concato, E. Kemper, T.R. Holford, A.R. Feinstein, A simulation study of the number of events per variable in logistic regression analysis, J. Clin. Epidemiol. 49 (1996) 1373–1379.
[31] J.A. Hanley, B.J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology 143 (1) (1982) 29–36.
[32] T. Fawcett, An introduction to ROC analysis, Pattern Recogn. Lett. 27 (8) (2006) 861–874.
[33] G. Sjöstedt, M. Grann, Risk assessment: what is being predicted by actuarial prediction instruments? Int. J. Forensic Ment. Health 1 (2) (2002) 179–183.
[34] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, t, Ijcai, vol. 14, 1995, Augus, pp. 1137–1145 2.
[35] Y. Bengio, Y. Grandvalet, No unbiased estimator of the variance of k-fold cross-validation, J. Mach. Learn. Res. 5 (Sep) (2004) 1089–1105.
[36] G. Seni, J.F. Elder, Ensemble methods in data mining: improving accuracy through combining predictions, Synth. Lectur. Data Min. Knowledge Discover. 2 (1) (2010) 1–126.
[37] GISIS, Global Integrated Shipping Information System (Annual) [online], https://gisis.imo.org/Public/Default.aspx , Accessed date: 4 January 2019.