Center for Maritime Economics and Logistics
Erasmus University Rotterdam

Erasmus University Rotterdam

MSc in Maritime Economics and Logistics

2018/2019

Risk analysis of maritime accidents

By

Fengoudakis Matthaios

Acknowledgements

With the finish of this master, I cannot but mention the contributions of some individuals who have helped me not only to complete this study but also supported me throughout the entire duration of my study at MEL. I would first like to thank my supervisor professor Ted Welten whose guidance was very important for this study. I would like to express also my appreciation to every professor or lecturer at the MEL program, who I came in contact with over the last year.

I would like also to thank my friends from Greece for their continuous moral support throughout this year especially Maria for her faith in me. Moreover, I want to express my gratitude for my fellow students Antonis, Nondas and Dimitris whose contribution made the last year an incredible experience.

The pursuit of my career dreams and this study would not be possible without the support of my parents via many ways. Their encouragement and faith to me and my skills helped me not only to complete this work but also to overcome any obstacle I encountered.

Abstract

Maritime accidents have been studied for a few decades now. However, despite the advancement of technology used on the vessels, maritime accidents still occur, therefore an analysis of their causes is essential in the preventing future accidents. In this study we assess the risk that vessel carry and their probability of participating in a maritime accident based on quantitative and qualitative characteristics of the vessel. To do so, we gathered public data regarding the size, age, type, operational strategy and history of vessels which participated in maritime accident. In order to study the relationship between these characteristics and the risk we developed a Bayesian Neural Network with the help of algorithms. In the first part of our assessment we examined the established relationships and the strength of these relationships. In the second part we calculated the conditional probabilities for each one of these characteristics and we identified conditions under which the probability of accident is higher. In the last part of our study we study how the operational strategy selected by ship-management company is affecting the risk that a vessel is bearing.

We find out that for each type of vessel there is a combination of conditions and circumstance which are related with the higher probability of accident. Therefore, the prediction of when an accident will happen is not possible. But the identification of these circumstances is viable through our methodology, which can act as a decision-making tool. The human element seems to be the main factor in maritime accidents. On the other hand the flag state under which is registered a vessel as well as the performance of it during inspections seem to be two trustworthy indicators of risk.

Table of Contents

List of Tables

List of figures

List of abbreviations

BBNs	Bayesian Belief Networks
DAG	Directed Acrylic Graph
EMCIP	European Marine Casualty Information Platform
EMSA	European Maritime Safety Agency
FOC	Flag of Convenience
FSA	Formal Safety Assessment
IMO	International Maritime Organization
ITF	International Transport Workers' Federation
MARPOL	International Convention for the Prevention of Pollution from Ships
MoU	Memorandum of Understanding
Rel. Freq.	Relative Frequency
SOLAS	Safety of Life at Sea

1. Introduction

1.1. *Background*

One imperative industry that is of significant importance to the global economy is the maritime industry. According to International Maritime Organization (IMO) website and the latest statistical data the 90% of the word trade is transported by Sea. Consequently, its contribution is estimated to be in billions of euro per year and millions of people are employed in operations related to this industry.(“ICS | Shipping and World Trade,” n.d.).The main mean of transportation that is used to facilitate the seaborn trade is the vessel. The total number of vessels registered under a flag for the age 2018 was 50732. Moreover, the fleet size the last years has increased as it is described in Table 1

Table 1 Fleet size, data from UNCTAD, created by author.

Year	2014	2015	2016	2017	2018
Total number of ships registered under a flag.	47 797	48 240	49 445	50 146	50 732

Those vessels are operated worldwide, therefore their activity is spread in different geographical areas. The degree of this segmentation is illustrated in Table 2. (EMSA, 2017)This table presents the numbers of ships which are sightseen as they were reported by AXSMarine, Vessel traffic and Marine traffic.

Table 2 Total distinct number of ships sighted by area year 2017, Database: Equasis

Geographical area	**Total distinct number of ships sighted by area**	**Percentage**
Australia and New Zeeland	6788	2,78%
East Africa	9465	3,87%
East Asia	30194	12,35%
Mediterranean Sea	30706	12,56%
Middle America and Gulf of Mexico	12175	4,98%
New guinea Pacific	4407	1,80%
North America East Coast	14316	5,85%
North America West Coast	7674	3,14%
North Asia	178	0,07%
North Europe	9800	4,01%
South America East Coast	10314	4,22%
South America West Coast	3445	1,41%
South Asia	27837	11,38%
South East Asia	32300	13,21%
South Africa	9518	3,89%
West Africa	17025	6,96%
West Europe	18427	7,53%
Total	244569	100,00%

From the above tables derives that the activity of ships worldwide is increased and segmented. There are sea areas such as South/East Asia and Mediterranean Sea which present more traffic than others. In addition, shipping industry is known as a competitive industry where cautious planning of many assets is crucial for the involved companies in order to survive in this market. Since the planning is dynamic and there are a lot of involving parts, mistakes happen. As George Williams Curtis mentioned “It is not the ship so much as the skillful sailing that assures the prosperous voyage”.

Though the safety standards of the maritime industry are high, accidents or incidents are existing. Many times, the aforementioned mistakes or the severe sea weather conditions contribute to the creation of a marine incident. The definition of marine incident is well-defined by maritime laws. According to the Australian Law marine incident (AIMS, 2015) is considered an event which includes the following:

- Death or injury to, any person on the vessel or associated with the navigation of it.
- The loss or of a vessel or collision of it with another vessel or object.
- The grounding, sinking, flooding or capsizing of a vessel
- Fire on board or structural which affects the safety of the vessel.
- A close quarters situation

1.2. *Problem identification*

In recent era, the development of ship designing, and advanced technological systems is significant. Despite the sophisticated technology of vessels, the number of marine incidents is not decreased. According to the European Maritime Safety Agency (EMSA) and their report of 2018 (Maritime & Agency, 2018) the number of marine incidents for 2011 were around 2000 and after 2014 are between the range of 3200 to 3500 reported incidents per year. In addition to the existing type of losses such as grounding and collision, new types are occurred like cyber incidents or technological failures. On the other hand, the number of losses in maritime sector the last has been decreased by 38% (Allianz Global, 2018) which is a result of the not only of the technological development of the ships but also the quality of the crew, who is more educated and well trained. However as explained earlier the safety in maritime industry remains still an issue for some occasions. The main reasons of the stable number of incident are considered to be the traffic density of some sea routes and the sea environment. (Yang, Zhao, & Ma, 2018)

The marine incidents are bearing risks not only for the shipowner and the economic system but also for the environment and the society. In order to manage those risk, the involving parts must first understand the root causes and then assess and model them. IMO has developed a tool named Formal Safety Assessment (FSA). This tool describes a methodology which can be used to illustrate av event, its causes and the results. The purpose of this tool is to be used in the process of decision making by the involving parts and to enhance the level of safety in maritime sector.(IMO, 2002). Under this procedure each incident is categorized and combined with a scenario which describes the sequence of events that are included to the incident. Furthermore, the accident has an outcome and is related with a frequency. However a risk assessment model should not treat the accident as a statistical number but it should also provide information about the root causes and aid the process of decision making by providing feasible solutions and define base on logical criteria which is the optimal solution(Mazaheri, Montewka, & Kujala, 2014a).

Risk should be described as a combination of uncertainty and damage since it includes a level of uncertainty and the probability of some kind of loss.(Kaplan, S; Garrick, 1981). According to Kaplan et all (1981) risk is a set of three elements. The basic procedure of analyzing those elements is to try to predict first which event may happen? (scenario), second what is the probability of this event to happen? (likelihood) and last which will be the outcome if this happens? (consequence). The product of these elements is the first definition of risk.

$$R = \{S_i \; L_i \; C_i\} \; (1)$$

R=Risk

S= is a scenario

L= the likelihood of the scenario to occur

C=the consequences/damage/outcome of the scenario.

Scenario can be defined also as a sequence of events triggered by an initiating event and or conditions. Each scenario of an accident can be divided into two parts. The first part could be before the accident (pre-accident). The second part could be placed after the accident. (pro-accident). Therefore, in order to alleviate the risk two strategies may be used. One strategy will focus on how to prevent the incident (proactive) and the second strategy will focus on how to minimize the losses after the incident (reactive).

1.3. *Research question*

Based on the definition of risk explained earlier the main research question of this research will be:

Does the vessel's characteristics and operational profile relate with the probability of participating in an accident and is there a modelling feasibility of predicting this probability?

The main goal of this research is to first establish the relationship between the input data and the scenario of an accident to be occurred. Input data is defined as various variables which are related with the vessels engaged in incidents. Specifically, those variables are operational and vessel characteristics such as flag of registration, age of the vessel, location, time, type of the vessel. In doing so, a model will be developed which will illustrate quantitatively this relationship and allow us to measure this probability and assess the risk. The results will indicate which vessel or operation strategy bear the less risk. Under the premise that maritime industry behaves rational and will select the strategy that offers the less risk.

We will investigate how the decisions related to operation strategy of a vessel affect the risk of an accident. To support our research objective, we answer some sub-questions, which have equivalent importance. The sub-questions are:

1. **What is risk assessment framework and how it is build?**
2. **Which are the most important root causes of a marine incident?**
3. **Which operational strategies bear less risk for a vessel?**

1.4. *Research methodology*

The research that is going to be conducted is basically quantitative. Various sources will be used to acquire statistical data related to vessels and maritime incidents. We will study only merchant vessels.

Sub question (1) will be used to build the model. Its results will indicate the fundamental variables that should be included in the research. The correlation of the above variables and the probability of an incident to occur will be established first through the literature review. Moreover, a theoretical background of the existing models will be provided and explain their application and relevance. The results of the theoretical research will be compared also with the empirical data and the fitting of the model will be measured through statistical indicators.

The sub-question (2) will help us describe the initiating events of our scenario, in other words the roots causes of a marine incident. The empirical data from IMO reports will be analyzed and help us categorize the scenarios

The sub question (3) is related with the consequence of a marine incident. This question will help us identify which approach a ship-owner should adapt in order bear less risk.

1.5. *Thesis structure*

The chapter 1 of this thesis will be the introduction. In this chapter it will be presented the background of this research and the identified problem that derives from it. The main research question will be described as well as the sub-questions and the methodology that will be used. In chapter 2 we will explain the theoretical framework of this study as well as some theoretical concepts related in this study. We will also form an extensive literature review related to risk assessment of marine accidents. In this chapter we will also explain our assumptions. In chapter 3 we will elaborate in detail the methodology that we will use to build up the model and the data that are required. Moreover, we will discuss how we are going to process the data and the workflow of our study. In chapter 4 we will present the result and the main points from the model. We will describe how the result are interpreted and contribute to optimizing the decision making regarding the risk a vessel is bearing. A sensitivity analysis will be carried out in order to assess the effect of different vessel operational strategies on the results. Chapter 5 will describe the conclusions of the research. The major results will be summarized and their relationship with our research questions will be illustrated. Finally, the limitation of the study will be highlighted, and we shall suggest field of further future investigation.

2. Theoretical background

2.1. *Introduction*

This chapter seeks to review the literature that forms the theoretical background of this research. We fist look in the main theoretical concepts that will consist our framework, such as what is maritime accident, the main categories of it. Next we will explain theoretically parts of our framework such as what is risk and hot it is assessed. In section 2.6 we will explain some theoretical concepts which are related with our selected variables and we believe require further explanation. in section we will illustrate our theoretical framework and the selected variables. In the next section we will discuss in detail the conducted literature review on the subject and we will summarize the main findings. In chapter 2.9 we will update our initial theoretical framework based on the literature review and the key findings.

2.2. *Maritime accident / incident*

2.2.1. Introduction

In this section we will define what is maritime accident. As mentioned in the previous chapter this thesis aims to study and analyze maritime accidents.

2.2.2. Definition

Accidents are the result of many complex consequences. In this chapter we will present the theoretical framework of our research, starting from the definitions of maritime incidents and accidents and their difference.

The Term of “Maritime accident” is used when:

- A damage to a ship or to a facility related to the ship’s operation is occurred.
- An injury or a death of people is occurred. These people are engaged in the construction equipment or operation of a ship(Board & Accident, 2017)

On the other hand, “Maritime incident” is defined as a situation where there is risk that an accident will occur.

2.2.3. Categories of Maritime accidents

As mention above the maritime accidents can be divided in two main categories. The first one is about damages related to the operation of the ship and the second one is about casualties whose root cause is related to the equipment, structure or operation of the ship. Each category can be divided in more subcategories based on two criteria, the severity and the type of the accident.

According to our research there are many ways to classify the accidents. Each Organization or database which is reporting them may use a different system of categorization. In Table 3 below we are representing some of these systems.

Table 3 Accident classification comparison

Japan Transport Safety Board		DAMA database[1]		EMCIP database[2]
Damages	Casualties	Damages	Casualties	Damages
Collision	Fatality	Ship-ship collision	Injury	Capsizing/listing

[1] Accident Database DAMA includes reports of maritime casualties as they are given to Finish authorities.(Ladan & Hänninen, 2012)

[2] European Marine Casualty Information Platform – EMCIP is a database operated by EMSA(European Maritime Safety Agency)

Grounding	Fatality and injury	Collision with an offshore platform	Death	collision
Sinking	Missing person,	Collision with a bridge of quay	Poisoning	contact
Flooding,	Injury	Grounding or standing		Damage to ship or environment
Capsizing		Severe tilting		Grounding/standing
Fire		Leakage		Fire/explosion
Explosion,		Environmental damage		Flooding/ fundering
Missing		Storm Damage		Hull failure
Damage to facilities		Machinery Damage		Loss of control
		Fire / explosion in the machinery area		missing
		Fire/ explosion in cargo area		Non-accidental type
		Fire / explosion in other areas		
		Helicopter accident		
		Near accident		
		Unknown		

Toffoli et al. at their research they used only six categories which were: Grounding, Fire/explosion, Collision, Stranded/Wrecked, Contact, Hull/machinery. Their study was based on the database of Lloyds world casualty.(Chalikias, Ntanos, & Milioris, 2015)

The data generally is divided in two parts, the factual data and the casualty analysis data. In other words, the factual data describes an initial event known as *accidental event* which occurred and trigged another event named casualty event which resulted to a casualty. This methodology is known as CASMET method.(Caridis, 1999) This methodology is used broadly by the European Agency of Safety to develop the EMCIP approach. According to EMCIP approach there two types of factors than can cause a casualty event. First there are factors that trigger the accidental event and there are factors which are named contributing factors which are enlarging the consequences of the initial event. (Ladan & Hänninen, 2012) Figure 1 illustrates the structure of this approach. It is a three-tier system which classify an incident based on the subcategories of each tier.

Tier 1 using a common taxonomy to classify the outcome of the accident. The subcategories used in this tier are described in Table 3 above. Tier 2 is referred to the accidental event which caused the event in tier 1. The main causes of the accidental event that EMSA is using at its reporting are:

- Environmental effect
- Equipment Failure
- Hazardous Material

- Human Erroneous Action
- Other Agent or Vessel
- Unknow

The last Tier is Tier 3 which is related with Contributing factors. These factors are usually related either to shipboard operations or to decisions and policies made by the shore management team. A summary of the contributing factors is presented by category below in Table 4

Table 4 Tier 3 subcategories, Source: EMCIP website

Shipboard operations	Shore management
Social environment	Business ethics
Supervision	Organization & general management
Crewing	Operations management
Work conditions	Safety & environment management
Physical stress	Crew management
Insufficient gear & equipment	System acquisitions
Maintenance	Vessel Design
Sea state/ weather conditions	Maintenance policy
Emergency alertness	Emergency alertness policy

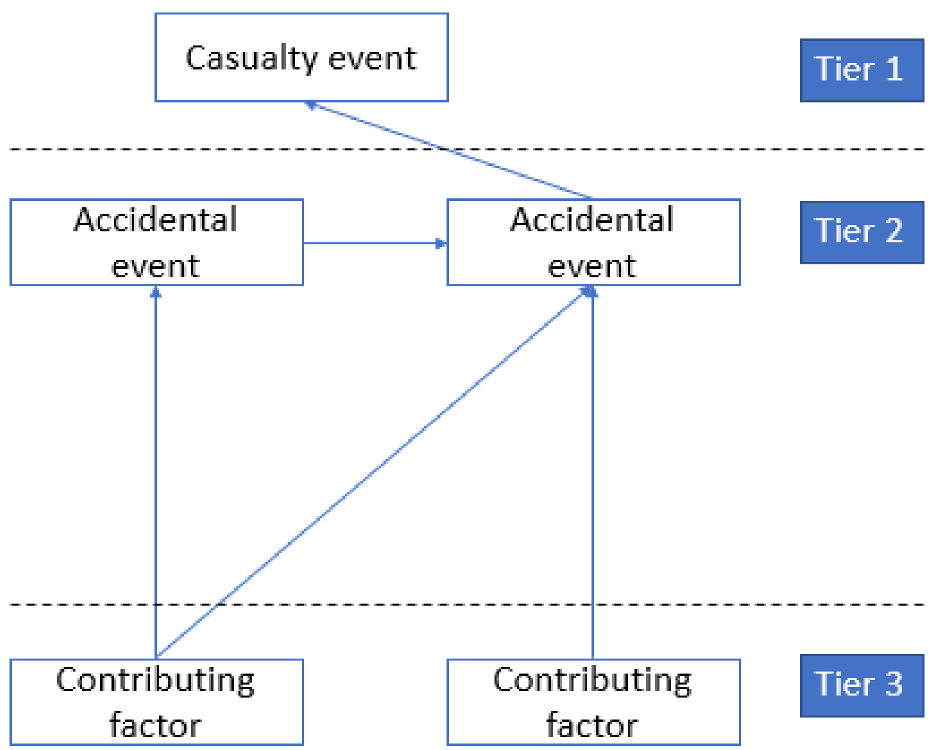

Figure 1 EMCIP approach

Another way to classify a marine accident is based on the severity of the accident and the casualties. According to the IMO which has issued the SOLAS regulation I/21 and MARPOL, articles 8 and 12, there are four types of maritime accidents(IMO, 2008):

- Very serious casualties
- Serious casualties
- Less serious casualties
- Marine incidents

More specifically according to the regulations mentioned above the *very serious casualties* are cases where occurred total loss of the ship, loss of life or severe pollution[3] to the environment. *Serious casualties* are named casualties to ships which do not meet the requirements as “very serious casualties” and which are related with collision, grounding, explosion, fire, contact, ice damage, heavy weather damage, hull cracking or hull defect. These events may have one or more of the following results:

- halt of main engines, sever accommodation damage or structural damage
- pollution (the quantity of the pollution is irrelevant)
- a breakdown which leads towage or shore assistance.

The term “less serious casualties” is referred to casualties which are not qualified as the aforementioned categories. Finally, the last category will be marine incidents which bear risk or are “near misses” events.

It is worth mentioning that administrations are obliged to provide data in their reports. The reports are composed by 10 annex tables. The type of information each annex is enclosing is presented in Figure 24 of Annex

2.3. ***Root Cause of Accident***

2.3.1. Introduction

In this section we will describe theoretically the root causes of a maritime accident which will be examined in detail through our empirical analysis.

2.3.2. Categories of root causes

In the literature generally there are two main root causes of a marine accident and lead to a ship loss. Faulkner in his research proved that 60% of ship loss is caused due to operational reasons (collision, fire, machinery damage) while the remaining 40% is related with designing or maritime causes (capsizing, water ingress, hull breaking). (Faulkner, 2013) However a big part of the above events was caused by human error. Moreover, the sea/weather condition is also a significant factor since it may prevent the crew from keeping the ship under control. (S, K, & M, 1997). The fact that human error is a significant factor is proved also by other studies. Marine safety reports from the area of New Zealand are in line with the above fact: 49% of the marines incidents reported human error as the root cause, while 16% reported environmental factors and 35% technical reasons. (Zealand, 1996) Also the EMSA stated that 58% of the investigated accidental events are caused by human error or action (Maritime & Agency, 2018) while 19,8% of those events reported supervision as a contributing factor, in other words human actions related to supervision contributed to the accidental event or worsen the consequences.

[3] Severe pollution according to MEPC 37/22, paragraph 5.8 is a case where serious catastrophic effects were created upon the environment or would have been produced without an action to prevent them.

Hetherington et al cited that the main reasons for human errors were related to false judgement and inappropriate lookout. Another cause of accident is the insufficient implementation of the safety regulations. (O'Neil, 2003, pp. 95-97) Data from another research by U.S. Coast Guard prove that almost 80% of the investigated incidents are caused by some form of human error. (Esbensen, 1985). Heea et all at their paper stated that at least 80% of the marine accidents were cited as result of organizational or human error. Same range of attribution to human error (70-90%) is indicated also by other researchers.[4] Specifically the trigging events were mostly attributed to actions by the operating staff, thus more than the 80% of the contributing factors were attributed to organizational decisions which influenced the personnel. (D.D., B.D, R.G., K.H., & R.B., 1998). However, the impact of human factor is different per type of accident. According to Pazara et all 89% - 96% of the collisions,75% of fire and explosions and 79% of grounding are attributed to human error. (Pazara HR, 2008) Except of many researchers also many authorities have acknowledged the human factor as the most important in maritime accidents. Such authorities are the "Australian Transport Safety Bureau", "The Marine Accident Investigation Branch of United Kingdom" or "The National Transportation Safety Board of U.S.A." (ÖZDEMİR & GÜNEROĞLU, 2015)

As Carridis cited in his report in 1999 it is not easy to decide the causes of a marine incident or accident and especially for the accidental events. This procedure requires objective criteria which are not always available, also this procedure focus about what happened and not the reason (Tier 2 of EMCIP approach).(Caridis, 1999).

2.4. ***Risk***

2.4.1. Introduction

As explained in the first chapter the goal of this thesis is to assess the risk related to maritime accident. in the previous two section we defined the maritime accident and its root causes. In this section we will explain what risk is and how it is defined.

2.4.2. Definition of risk

The etymology of risk is originated from the Latin word "*resecum*" (the one that cuts) and it was related with the dangers that sailors had to face such us cliffs. From the legal perspective risk is classified in three groups.

1. Certain/ unacceptable risks where there is a cause-effect relationship between the event and the casualty. This relationship is supported by scientific findings.
2. Residual/Acceptable risks which are the result of human activities. This link is not proved by science, but it is based on speculations.
3. Uncertain Risks are not supported by science; thus their existence cannot be excluded.(Liuzzo, Bentley, Giacometti, Bonfante, & Serraino, 2014)

Maritime accidents as explained before are unwanted events, risk is named the probability of occurrence of these events which are followed by multiple consequences. Many resources are defining risk as a function of the probability of a danger event happening and its consequences. This is known as the traditional definition of risk.[5] In the engineering science the risk is defined as the expected loss, however this approach does not describe situation where large consequences are combined with small

[4] (Lützhöft M, 2011, pp. 280-286) (Faturachman D, 2012)
[5] ("ISO 31000:2018 - Risk management -- Guidelines," n.d.), (Norway, 2016)

probability and vice versa. (Aven, 2010) According to the FSA guidelines risk is defined as the outcome of a consolidation between a probability (P) and the consequences (C) of a specific event.[6]

$$R = P\ x\ C\ (2)$$

According to Vanem et al risk is the possibility of consequences of unwanted events in different scenarios. (Vanem, Anta, Østvik, Del, & Comas, 2008)Moreover, the risk of accidents can be defined also as the combination of the likelihood and the severity of probable human injury, or damage to a infrastructure or the environment. (Aven, 2010, pp. 623-631)

In order to quantify the above approach FSA proposed a risk index which is equal with the product of P and C. Specifically the risk index is equivalent with sum of the logarithm of the two factors as it is described in equation

$$Risk\ index = \log(probability) + \log(consequence)\ (3)$$

The main drawback of this approach is that it may result in two similar situations A and B with the same risk index, but the impact of these situations may differ. For example, situation A may represent a situation with frequent events but minor consequences such as minor injuries or equipment damage. On the other hand, situation B may be a case where rare events are followed by large consequences such as multiple fatalities or total loss of the vessel. An improvement of this approach was introduced by Kaplan and Garrik in 1981 which is suitable for maritime traffic.

Particularly if the marine traffic is considered as a system then risk can be defined as a combination of three factors. The first factor is the Scenario (S) where it is described a situation that something is going wrong in the system. The second factor is the likelihood (L) of this scenario and the last factor is the consequences (C) of the hypothetical scenario.[7].In order to create a complete set of scenarios, all possible scenarios must be included. However the complete knowledge on the system is not possible to be achieved.(Haimes, 2009) Therefore Aven proposed that instead of consequences and probability it should be used consequences and uncertainty. In 2015 the Society for Risk Analysis proposed a modified definition of risk where the triplet of (C,Q,K) would be used. C is used to describe the consequences, Q is the measured uncertainty expressed in the form of boundary possibility and K denote the background knowledge on which C and Q are based.(Xia, Xiong, Dong, & Lu, 2017) The same approach was used also by Mazaheri et all who defined risk as the system illustrated in equation 4 below.

$$R \sim \{S, L, C|BK\}\ (4)$$

S describes a group of variables which act explanatorily for a specific scenario. These variables can be modified and change their relationship with each other based on the Background Knowledge (BK) about the analysis. L stands for a set of likelihood which is related with the corresponding consequences for a specific scenario and a specific collection of input and assumptions.(Mazaheri et al., 2014a)

Background knowledge (BK)

Moreover, Montewka et al proposed that background is a combination of understanding (N) and knowledge (K) which can be used both to describe the uncertainty of a system. Knowledge is focusing on establishing a proposition based on facts. In order

---

[6] .(IMO, 2002)(Montewka, Goerlandt, & Kujala, 2014)

[7] (Haimes, 2009)(Aven, 2011)

to succeed this establishment, the events must be tangible and the source of information reliable. On the other hand, understanding enables the comprehension between various facts and fictions. In that way the analyst can create new scenarios and explore new situations and solutions. At his work Baumberger presented four reasons that distinct understanding from knowledge, the main points are presented below in Table 5.(Baumberger, 2014)

Table 5 Comparison Knowledge and Understanding, Table made by Author, source: Baumberger

Knowledge	Understanding
Is collection of beliefs	Is a combination of belief and non-belief states
Can be divided in parts, partially holistic	Completely Holistic, cannot be divided in discrete parts
Not gradual. Either you know a subject or not.	Can be scaled, there are four scales of understanding. (Understanding in breath, depth, significance and accuracy
Based on facts	Based on evidence but not factive

Understanding and knowledge are significant for a risk analysis, based on them the analyst will decide if the results are trustworthy and informative enough to be used in the decision-making procedure.(Montewka, Ehlers, et al., 2014) Background knowledge can be expressed in a risk analysis with the use of Bayesian Belief Networks (BBNs). BBNs are graphical probabilistic models which describe the probability of corresponding scenarios and quantify their consequences. Moreover, BBNs models are suitable in cases where understanding on a subject is incomplete and the knowledge is limited.[8] BBNs also follow the definition of risk as combination of three factors. (Kaplan, 1997)

Scenario (S)

An initial stage of a risk analysis is the is to define the scenarios. At this stage the knowledge related to the subject is described, therefore it is crucial to select the associated variables and define correctly the initial assumptions. This stage is also affecting the following stages and the risk analysis framework. The main goal when describing risk is to understand in depth the scenarios. The identification of scenarios is related also with the identification of the cause of an undesired event, analyze it and mitigate the hazardous factors. The structure of a scenarios is usually divided in two parts, the qualitative and the quantitative. The qualitative part is describing the structure of the scenario while the quantitative is describing the content of the scenario. The description of the content is succeed through the description of events and their relationship with each other.(Montewka, Ehlers, et al., 2014)

Likelihood (L)

The Likelihood of a event is usually quantified in one of the following forms: the relative frequency, subjective probability and a combination of these two named the probability of frequency. (Montewka, Ehlers, et al., 2014). Frequency is used when there is a repetitive situation and illustrated how often an event is occurring. Probability is used when a situation happens only once, and we want to describe its success rate. Probability

[8] (Hänninen & Kujala, 2012) ,(Uusitalo, 2007)

of frequency is expressing the knowledge of a repetitive event whose frequency is uncertain and this knowledge is described by an probability curve.(Kaplan, 1997).

2.5. ***Risk analysis framework***

2.5.1. Introduction

A theoretical framework related to risk is a useful decision-making tool however it has to be validated first. In this section we will describe what methods can be used to validate a risk analysis framework.

2.5.2. Validation of a framework

In order to build up a risk analysis framework, there is a procedure of five parts that the analyst has to follow. First is to define the objective of the model, second the variables that he will use, third establish the qualitative part of the framework as explained above, by explaining the content of the scenarios. Fourth step is to build up the quantitative part and the last step will be to validate this framework.

More specifically in the first step the accidental scenarios will be selected and be described in detail. The analyst at this stage must distinct the initial or accidental events from the contribution factors and model the casualties. Next, the analyst will select the related variables and establish a relationship between the variables and the objective of the model. In the third part the structure of the model will be developed graphically. In the quantitative part all the variables will be presented and the acquired data. In the last part as mentioned earlier the framework will be validated. The methods that can be used to validate such a framework are:

- Sensitivity analysis

Through this analysis the variables with the higher impact on the product of the model are defined

- The Value of information analysis

This analysis has as goal to highlight the most informative variables. These variables make the probability mass of the outcome to be scattered.

- The influence analysis

This analysis is used in order to quantify the impact of changing the assumptions used in the framework or other corresponding variables. Usually this analysis is used in cases of limited knowledge or understanding on a phenomenon.

- The last method of analysis is the comparison with real data.

Once the results are finalized they are compared with the available data in terms of risk and severity.(Montewka, Ehlers, et al., 2014)(DNV GL, 2016)

2.5.3. Risk Evaluation – Risk management

In the previous sections of this chapter we described what is risk, how to identify it and what is the procedure to analyze it. These steps are part of a process called Risk assessment which concludes with the part of Risk Evaluation.

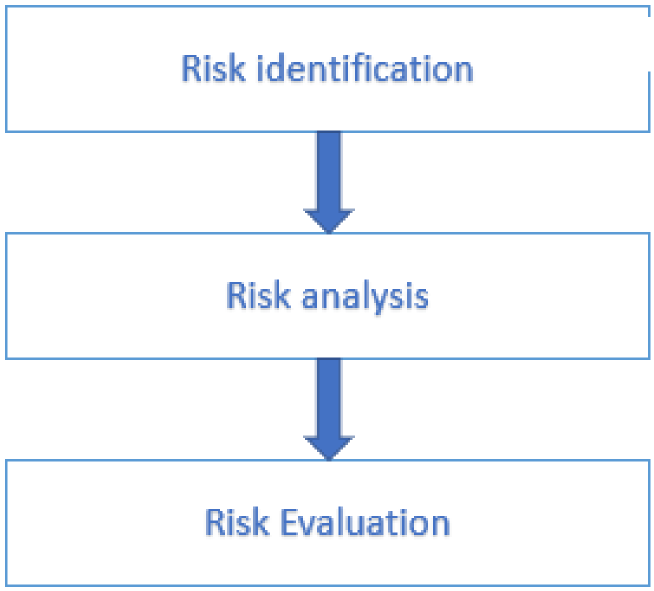

Figure 2 Risk assessment process

Risk evaluation is usually done by an organization. The aim of this evaluation is to decide which actions can be taken in order to alleviate the risk and it is a subject discussed often. However, it is considered a highly complex procedure since it includes a lot of factors even legal regulations. There are two main types of approach to alleviate a risk, the first on is react after the event in order to minimize the damage and the second one is to react proactively in order to avoid the undesired event. Regarding the second type of approach a mechanism was proposed which utilizes risk policy as a business strategy to alleviate risk by focusing in preparation and not in reaction. The mechanism is identifying the risks and manage them in advance instead of reacting in them.(Gurning, 2011) However not all companies are understanding the importance of risk management. As Mitroff and Alpasin (2003) cited the the percentage of companies of Fortune 500 that are prepared to handle crises or disruption is calculated between five and 25 per cent.(Mitroff & Alpaslan, 2003)

2.5.4. Risk Criteria

Part of Risk evaluation is also to evaluate the results of the risk assessment based on some criteria called risk criteria. The outcome of Risk evaluation is illustrated in a matrix called Risk matrix where there are three main types of risk. First is the acceptable risk, the second is acceptable risk with caution and the last one is not acceptable risk. However, it is very difficult to define precisely the term acceptable risk by an organization or by the society. These criteria are based on personal beliefs, preferences and expectations. They also differ from time to time.(Det Norske Veritas, 2001)

2.6. ***Theoretical Concepts related to the variables***

2.6.1. Introduction

The previous section described some theoretical concepts related to the framework of this study. In this section we will explain some theoretical concepts related to the variables we will use which we think they require further analysis.

2.6.2. Flag state and port state control.

Every vessel is registered to a country and carries its flag. The country is giving to the vessel a nationality and the holder of the flag has some duties and rights. The role of the Flag state is to monitor and control effectively the vessels under its jurisdiction in terms of administration, technical and society matters (The United Nations Convention on the Law of the Sea, 1982).In other words the flag state determines the applicable laws governing all the activities on the ship. (Committee & Senate, 2009). According to Brook and Pallis at their report in 2008 the term “effectively” can be interpreted as the transportation of good and or people by sea without occurring any hazard related to safety, environment and health. In order to measure the term effectively the concept of Flag Performance was created which is defined as “how effectively a flag state is utilizing its resources (financial or human capital) to ensure the safety and labor conditions of the crew on board and the protection of the environment” (Graziano, 2018) As a result there are flag states who are considered better in terms of performance than others.

Nowadays there are two main initiatives which are used to measure Flag State performance. The first one is a list published by Paris MoU and Tokyo MoU. They use a scale of three levels: White, Grey, Black. The second one is a Performance table issued by the International Chamber of Shipping.(Graziano, 2018). Paris MoU and Tokyo MoU are two organizations which are collaborating with maritime authorities and Port State Controls in order to eliminate situations where the vessels are not meeting the required minimum standards.

2.6.3. Flag of convenience

As mentioned earlier every vessel is registered to a Flag state. The choice of the state is based on many criteria, economics and operational. The management team of the vessel is able to register her in any state even if they are not related to it. There are some states whose associated obligations are considered less strict, these Flag states are called “Flag of Convenience” (FOC). The registration in a FOC is a tool that companies are using often to reduce the operational costs or other financial obligations such as tax, labor legislations and investment controls. FOC are creating loopholes, which are exploited by some shipowners to sail sub-standard vessels (Hamad, 2016). Many maritime safety and security issues are related to these types of flags since they facilitate poor labor conditions, cover up of ownership and insufficient regulations. (Shaughnessy, 2007)Totally there are 35 flags of convenience according to the International Transport Workers’ Federation (ITF) which are presented below inTable 6.(International Transport Workers’ Federation, 2019). However, this list includes some of the most reputable open registries such as Panama, Liberia and The Marshal Islands. According to UNCTAD report of 2018 under these three flags is sailing the 41.44 % of the world dead weight tonnage(Asariotis et al., 2018)

Table 6 Flags Of Convenience (FOC) Source: ITF website, Table made by Author.

1. Antigua and Barbuda	2. Bahamas	3. Barbados
4. Belize	5. Bermuda (UK)	6. Bolivia
7. Cambodia	8. Cayman Islands	9. Comoros
10. Cyprus	11. Equatorial Guinea	12. Faroe Islands (FAS)
13. French International Ship Register (FIS)	14. German International Ship Register (GIS)	15. Georgia
16. Gibraltar (UK)	17. Honduras	18. Lebanon
19. Liberia	20. Malta	21. Madeira
22. Marshall Islands (USA)	23. Mauritius	24. Moldova
25. Mongolia	26. Myanmar	27. Netherlands Antilles
28. North Korea	29. Panama	30. Sao Tome and Príncipe
31. St Vincent	32. Sri Lanka	33. Tonga
34. Vanuatu	35. Jamaica	

2.7. ***Theoretical Framework before the literature review***

2.7.1. Introduction

In this section we will explain the theoretical framework of this study which will include the theoretical concepts explained before. This framework will be prior to the literature review. We will explain the variables under consideration and their expected relationships.

2.7.2. Selected variables and framework

As we explained earlier the total risk of a vessel participating in an accident is a combination of the probability of the accident scenario and the corresponding consequence. In this research we will examine accident-scenarios whose consequences will be scaled based on the IMOs' reports.

Each maritime accident has one or more root causes which from now one will be called "initial events". These events were the initial events which may had ignited a chain of events that lead to an accident.

This risk assessment will study three main categories of factors which influence the risk of a maritime accident. First will be the static factor which is related with the characteristics of the vessel and the operational profile in the moment of the accident. This factor is mainly influenced by the decisions of the shipping company. Second will be a geographical factor which is related with the location of the accident and the time. Third will be a factor related with the Accidental events which are considered as causes of the accident. These events are usually influence by the performance of the crew, the maintenance policy of the shipping company and unpredicted events.

The goal of this research is to study the above factors for each accidental scenario and assess the associated probability and consequences. The total risk of a vessel will be the sum of the individual risks for each type of accident.

The variables that we will investigate are described shortly in the Table 7 below.

N	Variable	Short description
1	Accidental Event	The root causes of the accident will be described
2	Detained previously	It will indicate if it was detained before by an authority
3	Event Location	Describes the location of the ship during the event
4	Fatalities	It will indicate if there were or not fatalities / missing personnel
5	Flag state	The flag state will be classified based on the Classification of Paris MoU
6	Injuries	It will indicate if there were injuries
7	Sea area	The sea region where the accident occurred
8	Season	The season of the year will be indicated
9	Severity of Casualty	Based on the reports of IMO a classification will occur regarding the severity of the damages
10	Ship Age	Based on the year it was built
11	Ship operational condition	Describes the operational condition of the ship during the accident
12	Ship size	The gross tonnage
13	Ship type category	Based on the cargo, vessels will be classified
14	Time of the day	The time of the accident
15	Type of accident /scenario /outcome	The scenario of accident will be described
16	Vessel Length	It will indicate the length of the vessel
17	Flag of convenience	If the flag is a flag of Convenience
18	Crew	Number of crew
19	Chang flag	If the vessel has changed flag before the accident
20	Change Name	If the vessel has changes, her name before the accident
21	Operation capability after	If the Vessel was able to continue her journey after the accident

Table 7 Variables selected prior to Literature review

Our goal is to build a BBNs which will adapt the framework illustrated below. Based on that framework will study the relationships between the variables.

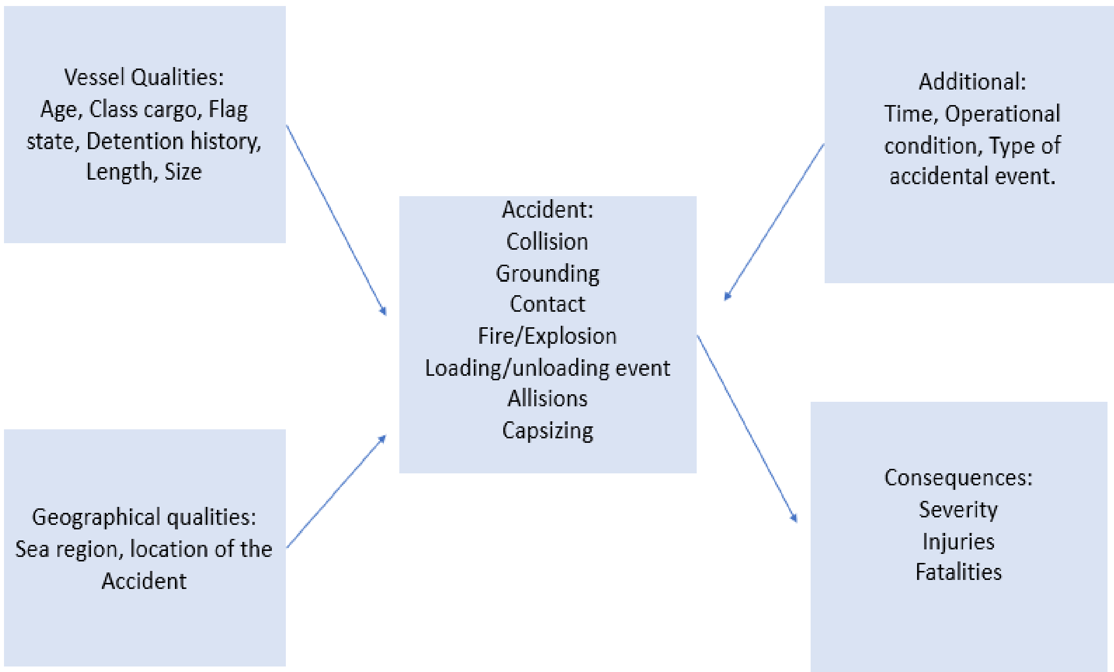

Figure 3 Theoretical Framework prior to literature review

2.8. ***Literature research***

2.8.1. Introduction

In this section of the second chapter we will represent based on our systematic literature review some previous works regarding maritime accidents, the variables that was used, the methodology and the main findings regarding these variables.

2.8.2. Literature review

Chalikias et all studies maritime accidents in Greece for vessel bigger than 1000 GT. They investigated all types of accidents between ships for the years from 1974 to 2010. The databased they used it was provided by the Greek Ministry of the transportation. Their independent variables were the size, the age and the type of vessel as well as the initial event of the accident. Their dependent variable was the accident area, the type of accident and the outcome. Their goal was to examine the relationship between the two types of variables with the help of chi square test to examine this relationship. Their findings indicated that ships under 5000 grt are suffering more accidents. Regarding the ship age they proved that older vessels are more prone to accidents. The most often accident was grounding and the location with the higher frequency was coastal waters.(Chalikias et al., 2015)

Willem et all also studied maritime accidents which were caused by human error. They analyzed the causes of the human error and they proved that lack of attention and proper training is one of the main reasons. They used casual networks with the form of event tree and the reliability index to examine their sample.(Wagenaar, Willem, & Groeneweg, 1987)

Montewka et all have studied the case of ship collision, specifically for ROPAX vessels. Their independent variables were first the ship particulars such as the length of the vessel, the type and the width. Second, they used information related to the time of the accident such as the velocity of the vessel and the route, the sea condition. For their research they used Bayesian neural networks to express their scenario and the corresponding probabilities. Their dependent variable were the outcome of the accident and the severity of the damages. They proved what sea conditions and big wave heights can affect negatively the stability and operation capability of the vessel. Through this study also it is proved that BBNs can be used as a methodology of a holistic study of maritime accidents. (Mazaheri, Montewka, & Kujala, 2014b)(Bitner-gregersen & Monbaliu, 2006)(Montewka, Ehlers, et al., 2014)

Ozdemir and Guneroglu studied the importance of human factor in maritime casualties. The study in detail the contribution factors in the maritime accidents. They used a decision-making trial and evaluation laboratory method in order to simulate their study. As independent variables they used characteristics of the vessel, the crew and the ship management company. Each criterion was based on team of experts which assisted the research. Their findings show that the most important factor of accident is with the human element and specifically is lack of “ability, knowledge and skill”. The second most important factor is the physical conditions on the ship. And the least most important the cargo.(ÖZDEMİR & GÜNEROĞLU, 2015)

Faghih-Roohi et all used a Markov modelling in combination with Markov Chain Monte Carlo (MCMC) simulation in order to predict maritime accidents based on the severity of their causalties the upcoming years. They acquired data from the accidents reports of the Australian safety Bureau from 2005 to 2010. Their results showed that the accident rate will drop the upcoming years and so is the severity of the accidents.(Faghih-Roohi, Shahrzad, Xie, & Ng, 2014)

Another relevant work also proved that human factor is crucial. Hanninen and Kujala investigated the collision occurred in the Gulf of Finland. They examined data regarding the time of the accident, the characteristics of the vessel, the crew and the navigational equipment. For their research they utilized BBNs. The proved that during night and conditions where the visibility is low and the fatigue of the crew is increased, the probability of collision is also increased.(Hänninen & Kujala, 2009)

Vanem et all studied the risk of accident for LNG carriers. They study the general categories of accident with the help of event trees. Their data included information regarding the condition of the vessel during the accident and they calculated the corresponding probabilities of the consequence for each type of accident. As total risk they defined the sum of the individual probabilities for each category of accident. Their findings demonstrated that accident with the higher risk is the collision. Moreover serious damages on the vessel are also observed at grounding or Allision.(Vanem et al., 2008)

Aziz et all studied the maritime accident which were caused by machinery or hull failure. The methodology that they used was the “bow-tie” model where they input several data related to the root causes and the condition of the vessel at the moment of the accident. They created 4 scenarios which represent a machinery or hull failure and they calculated the corresponding probability for each one. The scenario with the higher frequency was the Propulsion failure.(Aziz, Ahmed, Khan, Stack, & Lind, 2019)

Other relevant work also proved that the risk of maritime accident is higher during winter that in summer due to the metocean condition. More specifically Fernades et al investigated the maritime accidents at the coasts of Portugal and their impact in the environment l. They implemented the fundamental equation of risk we described in section 2.2. with the help of risk indexes. They also proved that high traffic near the coast line is not accompanied with higher risk of sea pollution. (Fernandes, Braunschweig, Lourenço, & Neves, 2016)

Another relevant study that calculated causation probability with the help of algorithm is the study of Silveira et al. They investigated the probability of collision of in the Coast of Portugal based on the traffic pattern and the AIS data of the nearby vessels. Their findings indicate that tanker present the higher probability of collusion due to their lack of maneuverability. Also they demonstrate that coastal waters present higher probability of collision than port areas.(Silveira, Teixeira, & Soares, 2013)

According to Yang et al who studied maritime accidents with the help of K-medoid algorithms, most of the accidents is happening near the coastal line. Generally, dry cargo vessels seem to have the higher likelihood of a collision and in terms of ship tonnage, vessels less than 3000 GT present higher probability. Regarding the time frame of the accident, the analyst stated that the time interval with the higher frequency is between 8 to 12 pm and the season with the most collisions is the Spring period. As main root cause of the accident was indicated the human error. (Yang et al., 2018)

Pagiaziti at her thesis she studied the maritime accidents of RO-RO and passenger ships from 1990 to 2014 with the help of descriptive statistics. As independent variables she used the type of the vessel, the location of the accident. Dependent variables were the type of accident and the fatalities. Her findings proved that harbor area and coastal waters present the higher probability of accident however the less severity in terms of fatalities. Regarding the type of accident with higher number of fatalities , grounding seem to present the higher frequency.(Pagiaziti, 2014)

A statistical analysis on Tankers showed that most of the accidents are happening when the vessel is en route or in the open sea. It also proved that even the frequency of maritime accidents is decreased , the severity of the consequences towards the environment is increased.(Veritas, 2006)

Ventikos et all examined the maritime accidents in Aegean Sea with the help of probabilistic model and the kinetic energy. They gathered 10-year data regarding the speed of the vessel and the location of accident as well as the weather condition in the moment of the accident. they proved that areas such as strait where the average condition is rough demonstrate higher probability of accident with severe consequences. Also ship which are implementing the slow steam strategy are expecting to present smaller frequency of accidents.(Ventikos, Stavrou, & Andritsopoulos, 2017)

An extensive research of risk accidents was done by Stornes. He investigated maritime accidents in Norwegian Sea. He used many independent variables such as the type of the vessel, the ship particulars, the age, the operational condition, the visibility and the sea condition. He used the method of multinomic regression to prove that groundings are occurred mainly in coastal waters, collision in deep sea waters and allisions in port areas. Moreover, he proved that bigger vessels occurred higher severity of damages. The rest investigated relationships were not statistically significant according to his result.

The speed, hull form and displacement of the ship is also proved that are important factors, especially for the case of grounding.(Ancuţa, Stanca, Andrei, & Acomi, 2017).

2.8.3. Conclusion of literature review

This section is summarizing the main findings of prior literature review which were described in the previous section.

Author	Year	Independent Variables	Dependent variables	Methodology	Main findings
Chalikias et al	2015	Ship type Ship age	Type of accident Area of accident	Chi-square test	• Older vessels or bigger than 5000 grt are more prone to accidents.
Wagenaar et al	1987	Root causes of accidents	Final consequence of the root causes	Event tree	• Lack of Attention and proper training is the main causes of human error
Montewka et al	2012	Ship particulars and geographical characteristics	Outcome of accident	Bayesian neural network	• Established the BBN as method of analysis of maritime accidents.
Ozdemir and Guneroglu	2015	Contribution factors of a maritime accident	Maritime casualties	multiple criteria decision-making	• Lack of knowledge, ability and skill is the major contribution factor related to human error. • Second most important is the Physical condition of the ship.
Faghih-Roohi et al	2014	Rate of accident and severity of accidents	Probability of accident the upcoming years	Monte Carlo simulation	• The rate of accidents and their severity will drop
Hanninen and Kujala	2009	Ship characteristics, Time of accident, Visibility and weather condition Condition of crew	Probability of collision	BBN	• The probability of collision is higher during conditions with low visibility or when the crew is tired.
Vanem et al	2009	Condition of the vessel during the accident	Severity of consequences	Event trees	• Collision bear the higher risk

Aziz et al	2019	Condition of the vessel and root causes of machinery/hull failure	Probability of machinery/hull scenario based on four scenarios	Bow-tie model	• Propulsion failure is the machinery failure with the higher probability
Fernaded et al	2016	Traffic data and metocean condition	Probability of accident and oil spoil	Quantification of risk through risk indexes	• Metocean condition affects directly the probability of accident. Winter presents the higher probability.
Silveira et al	2013	Ship particulars, AIS data, Traffic patterns	Probability of collision	BBN, Algorithm	• Tankers and coastal waters present the higher probability.
Yang et all	2018	Ship characteristic, time, season	Probability of collision	K-medoid algorithms	• Dry bulk are prone to accidents. The time frame with the most frequency of accidents is 20-24. • Spring present the higher probability of collision. • Vessels smaller of 3000 grt are more prone to collisions.
Pagiaziti	2014	Type of ship, Location of accident	Frequency for each type of accident and the fatalities	Descriptive statistics	• Harbor area and Coastal waters present the higher probability of accident, • Grounding demonstrate the higher risk for human life.
Ventikos et al	2017	Location of the accident, speed of vessel, weather conditions	Probability of consequences	Probabilistic model in combination with kinetic energy	• Straits are presenting higher probability of accident. The same is expected of sea regions with severe weather conditions. • Ships with slower speed are expecting to present smaller frequency.
Stornes	2015	Ship particular, ship age, type, location of accident,	Type of accident, Casualties	multinomic regression	• Groundings are occurred mainly in coastal waters, collision in deep sea waters and allisions in port areas. • Dry bulk present high frequency of collisions.

		visibility, weather			• Bigger vessels occurred higher severity of damages. The rest investigated relationships were not statistically significant according to his result.

2.9. ***Framework after literature review***

After we concluded our literature review, we revised the theoretical framework of this research. We selected 21 variables which we believe have causal relationship with the initial event of an accident and the type of accident as it is illustrated in the Figure 4 below.

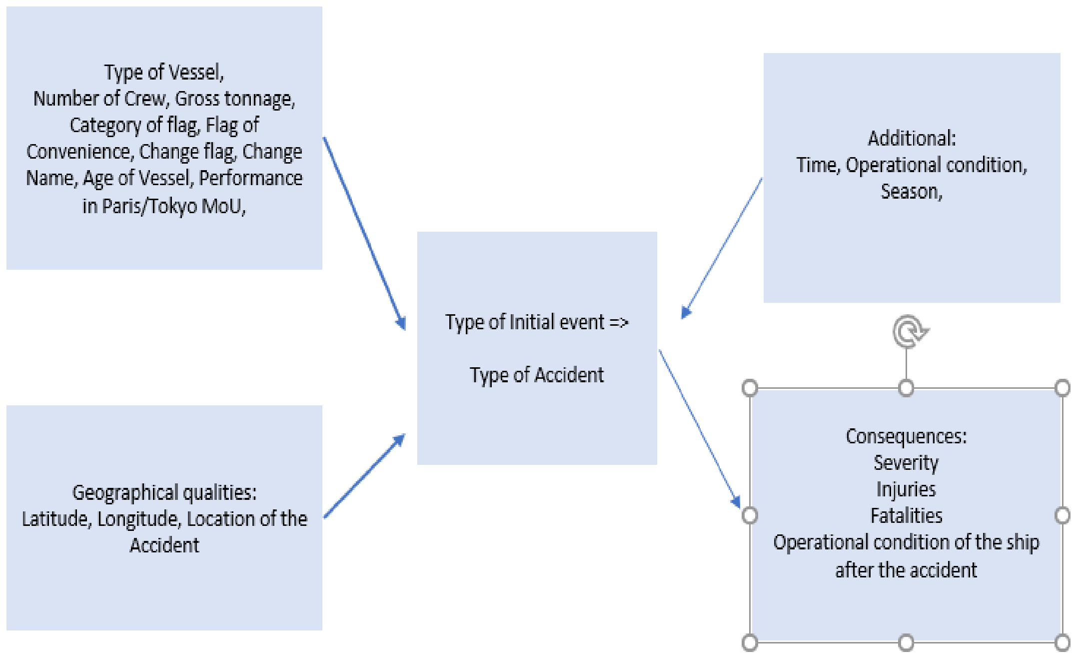

Figure 4 Revised Theoretical framework

The selected 21 variables will be:

Type of vessel	Number of Crew	Gross tonnage of the vessel
Category of the flag	Flag of convenience	Change of flag
Change of Name	Age of the vessel	Season
Time	Location of the accident	Operation condition
Performance in Paris/Tokyo MoU	Initial event	Type of Accident
Severity of Casualties	Operational Condition after	Injuries
Fatalities	Latitude	Longitude

Table 8 Selected variables after the literature review

The selected variables are a combination of quantitative and qualitative variables which we believe that related to maritime accidents. As dependent variable we choose the type of accident, the initial event and the consequences. However, we will study also the relationship between them. Moreover, we decided to include occupation accident together with the general type of accidents because we believe they have great impact for the crew on board. Finally, we will try to define which operational strategy selected by a ship management company bears the less risk.

2.10. *Conclusion*

Based on the literature review we can conclude that the risk a vessel is bearing is product of three factors, probability of each type of accident and the associated consequences. Based on the literature review we selected some variables and we will test their relationship with the maritime accident. We believe that these variables may act as indicators of conditions where the risk of a ship is increased. Therefore, our study will highlight these conditions. The selected are quantitative and qualitative variables which are related with the vessel, the location of the accident and the operational strategy of the ship managing company. In the next chapter we will explain the methodology we will use to test our framework.

3. Methodology

3.1. *Introduction*

This chapter is aiming on providing the reader all the theoretical background regarding the sources of data and the methodology to conduct a statistical analysis and modeling in order to answer the main research question of this thesis. The chapter will follow up this structure therefore in Ch. 3.2. the main sources of data will be presented as well as the sequence in which they will be used. In Ch. 3.3 we will explain the mechanism of the main statistical tools and algorithms that this research will utilize in order to reveal the characteristics of the data, establish causal relationship among the variables and construct a Bayesian neural network. Final, in Ch 3.4 the methodological approach of modeling the risk can be found divided in steps which follow chronological order.

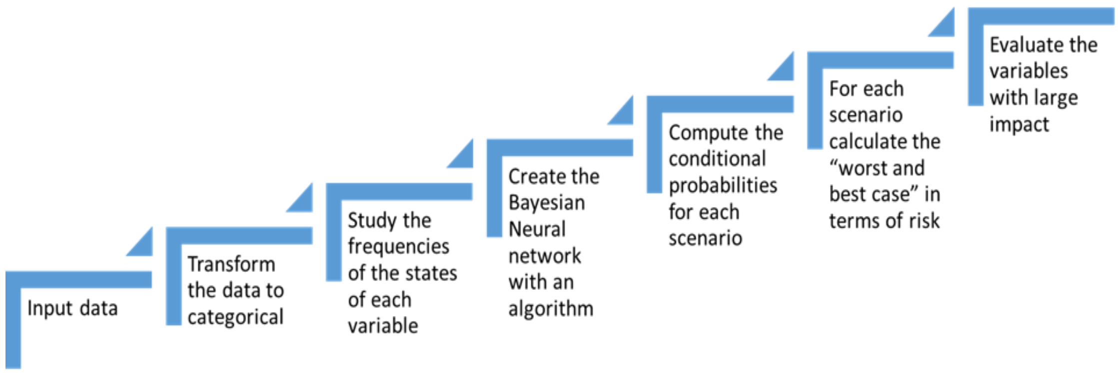

Figure 5 Chart flow of the methodology

3.2. ***Data collection***

In this section we will explain data sources and how we gathered the data. The main source of data will be the IMO's GISIS - Marine Casualties and Incidents Module. This database is considered has a collection of extensive reports regarding the maritime accidents as they reported to Maritime authorities. It is a project initiated by IMO who aims to increase traceability and transparency in the shipping industry. The database is accessible to the public, the only requirement is a free registration. The interface of this database is web-based and via and SQL type Queries and is allowing the researcher to make advanced research. GISIS can be considered as a trustworthy database since IMO is an organization respected worldwide and responsible for the safety and security of shipping industry.(Hassel & Lars Petter, 2009)

The second source of data will be the Paris MoU and Tokyo MoU. As mentioned previously those two organizations are authorized to collaborate with maritime authorities and publish data regarding ship detentions which are considered trustworthy worldwide. All the data will be discretized and transformed into numerical number in order to elevate their process via the selected algorithms.

3.3. ***Statistical analysis***

3.3.1. Introduction

In this section we will describe in detail the statistical tools which we will use to interpret our model through the statistical analysis. Statistical analysis is a methodology of collecting, analyzing and interpreting data.(Chatzaras, 2018). As mentioned earlier we will collect a sample[9] of reports which describe a target population[10] of vessels. All these selected vessels have a common characteristic, they are associated with a maritime accident or incident, but each individual has some attributes which are called variables and distinct them among other individuals of the population.(Isotalo, 2006)

The following subsections are based on a data set of nominal variables V_x (x = 1,2,3….k) where x is the number of variables and N the number of observations. The data set will be organized in a matrix composed of vectors with the following form

$$V_X = \begin{pmatrix} V_{11} \\ V_{21} \\ \cdot \\ \cdot \\ V_{N1} \end{pmatrix} (5)$$

A sample form of the matrix data is the following:

$$\begin{matrix} V_{11} & V_{12} & \cdot & \cdot & V_{1x} \\ V_{21} & V_{22} & \cdot & \cdot & V_{2x} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ V_{N1} & V_{N2} & \cdot & \cdot & V_{Nx} \end{matrix}$$

where V_{ij} is a value of the j : the variable collected from i : the observation, i = 1, 2, . . . , N and j = 1, 2, . . . , x.

3.3.2. Frequency

Since the data is nominal, in other words qualitative or categorical the calculations we will perform will be based on frequencies or the percentages of occurrence (Keller., 2014) Frequency is defined as the number of observations (v) which are included in a particular category of class of the qualitative variable.(Isotalo, 2006) Another indicator related to frequency is the relative frequency of the class which is equal to the rate of frequency to the total number of observations. The relevant frequency is a statistical tool which can provide information regarding a pattern that the data may follow. The class with the higher relative frequency is called mode class.

$$Relative\ frecuency\ of\ class\ z = \frac{Frequency\ of\ class\ Z}{Total\ number\ of\ observations\ (N)}\ (6)$$

Where z (z= 1,2,3…w) is the number of class that a variable V_x is categorized.

3.3.3. Probability

Another term that we will use in this research is the probability of an event which is defined as the number of times it will occur during a long run of repeated observations.

$$S = Values\ of\ Variable\ X = \{O_1, O_2, O_3, \dots . O_K\}$$

S is a set sample which represent a group of outcomes, the probability of each outcome/event is notated with $P(O_K)$ and satisfies always two requirements.

- $0 \leq P(O_K) \leq 1$
- $P(O_1) + P(O_2) + P(O_3) \dots P(O_K) = 1$

[9] Sample is considered a set of data selected from the examined population (Keller., 2014)

[10] The population under investigation which we want to study and make conclusions. Population is a group of all elements examined by a statistician. (Keller., 2014)

Probability is a useful tool because it links the sample with the population. Our research is based on the subjective approach where probability is described as the degree of belief, we hold in the occurrence of an event. (Keller., 2014)Moreover, the probability of an event is the sum of probabilities of individual events which compose the event. e.g. $P(Event\ \Omega) = P(A) + P(B) + P(C) + P(D) + P(E)$

3.3.4. Joint, marginal and conditional probability

According to Keller Intersection of the events A and E is the event that occurs when both event A and B occurs. Joint probability is defined as the probability of this intersection. Marginal probability is the sum of two or more joint probabilities which share a same attribute. Conditional probability is the probability of an event given another event.

e.g. Probability of event A given the event B is:

$$P(A|B) = \frac{P(A\ and\ B)}{P(B)} \quad (7)$$

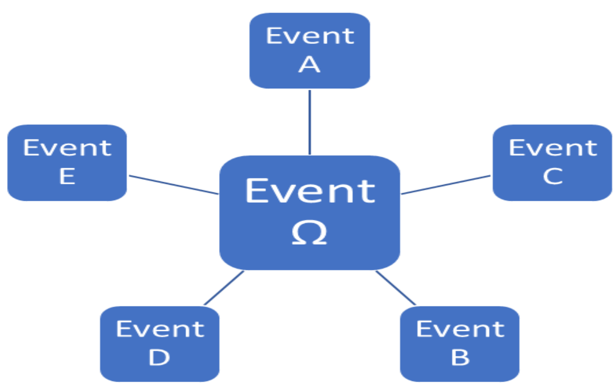

Figure 6 Related Events & Joint probability

3.3.5. Bayes's Law

In order to benchmark the relationship between two events we can use the conditional probability. However, there are situations where we have to compute the probability of possible causes of an event in order to assess the probability of the event. Bayes law is a law of logical interference. It connects the hypothesis we are interested with the corresponding evidence.

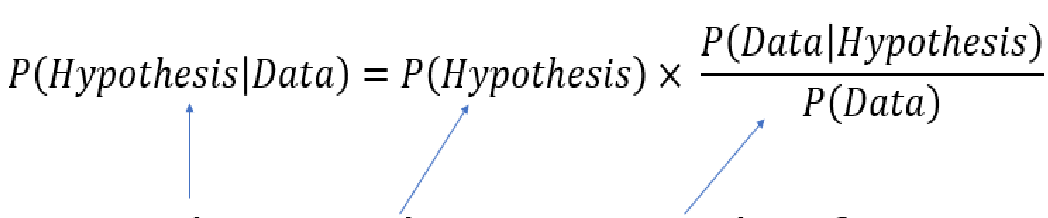

Figure 7 Bayes' Theorem

For example if we that event B is the given event and the events $V_1, V_2 \ldots V_k$ are the events for which prior probabilities are known as : $P(V_1), P(V_2) \ldots P(V_k)$. Moreover the observed likelihood probabilities are $P(B|V_1)$, $P(B|V_2) \ldots P(B|V_k)$, (the correction factor) and we are looking into for the posterior possibilities $P(V_1|B)$, $P(V_2|B) \ldots P(V_k|B)$

$$Bayes's\ law : P(V_i|B) = P(V_i) \times \frac{P(B|V_i)}{P(V_1)P(B|V_1) + P(V_2)P(B|V_2) + \cdots P(V_k)P(B|V_k)} \quad (8)$$

3.4. ***Modelling***

3.4.1. Introduction

This section has as goal to describe the theoretical framework of the modeling procedure we will follow in this research. The main steps will be:

1. Input of data
2. Describe the data via frequency tables
3. Define the structure of the BBN
4. Calculate the probabilities for two scenarios (worst case – less worse case)

3.4.2. First step:

Since the data are from different databanks, we will combine them in one data set which will be created in a excel file. Each report of maritime accident will be studied individually and then the IMO number of a vessel, which is unique, will be cross-checked in the database of Paris MoU and Tokyo MoU.

3.4.3. Second step:

In the next phase we will examine the data and create frequency tables for each type of accident and the severity of consequences.

3.4.4. Third step:

The BBN is a tool to illustrate a cause probabilistic relationship among random variables via joint probabilities distribution. It is split into two part. The first one is a graphic illustration of the network through a directed acrylic graph (DAG) and a collection of conditional probabilities distributions. The DAG is using nodes to represent the variables and edged narrow to indicate the cause probabilistic relationship. Each node has a set of possible responses which are called states. The node that the edged arrow begins from is called *parent* and the node that it ends up is called *child or descendant.* The interpretation of this edge arrow is that the parent node is causing the child node and the conditional probability of the child is defined by the possible responses of the parent node. If a node doesn’t have any connection with another node, then this node is considered as independent and irrelevant with the analysis.(Michal Horný, 2014). Another attribute of the nodes is that they are conditionally independent from all corresponding non-descendants.(Spiegelhalter, 2002) If we assume that the variables are random with a joint probability p(V) then it can be proved that the joint distribution factorizes into a set of terms that describe “local’ dependencies(Spiegelhalter, 2002)

$$p(v) = \prod_{U \in V} P(u|parents[u])(9)$$

An illustration of the Bayesian network is below in Figure 8.

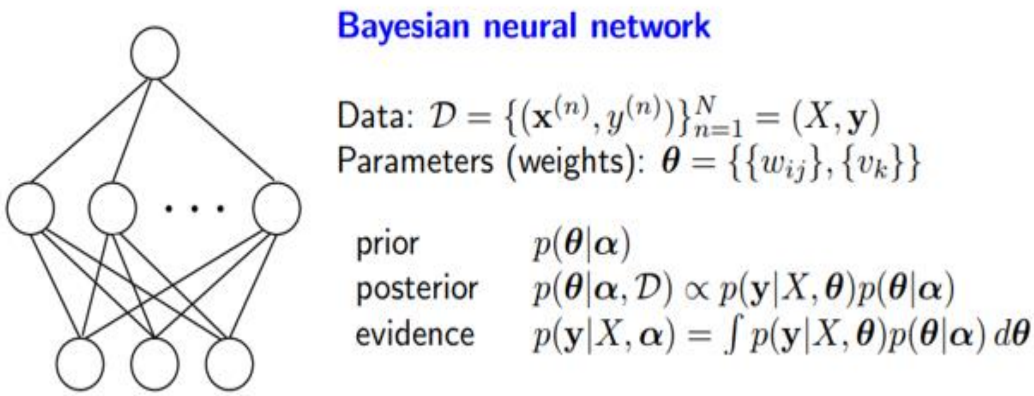

Figure 8 Bayesian neural network mathematical presentation

We selected this method because it has several advantages which will be described below. First the most important advantage is considered their ability to “learn”, in other words BBN can be constructed automatically base on a database. (Murphy K, 1998) According to Heckerman the Bayesian models have the ability to handle missing data and they provide to the analyst the opportunity to combine both his background knowledge and the data in order to study the casual relationship between the variables. (Heckerman, Geiger, & Chickering, 1993).

Another advantage is that they don’t require a large sample of data in order to provide a valid result. (Kontkanen P., 1997) Their main disadvantage is that their analyzing power for continuous data is limited, therefore the data set must be discretized before it is processed.(Uusitalo, 2007). As mentioned before construction of the BBN will be completed through the ability of the Bayesian network of autolearning. This procedure can be done with the help of an algorithm which will evaluate the available networks and choose the structure that perform better, this algorithm is called High climbing algorithm.

First, we will try to build a network without the implementation of the theoretical framework and then with the use of it. The two networks will be compared with the use of AIC score. The lower is the AIC score the better fits the model to our data. Moreover, each arc has different contribution to the network. We will measure the importance of each arc by measuring the strength of each arc with a score criterion. The value of the score represents the loss or gain that will be caused in the network’s score by removing the specific arc. The more negative is this difference the stronger is the relationship

3.4.5. Fourth step:

The result of the previous step is a DAG network which illustrate the relationships between our variables but also each node will be accompanied through a conditional probability. there for each type of investigated accident and we will compute two probabilities of the likelihood of the event and the other one for the consequence of it. As Kaplan defines it and we explained earlier in chapter 2 risk is a triplet factor:

$$R = \{(S_i, P_i(\varphi_i), P_i(x_i)\} \ (10)$$

Where S_i is the scenario, $P_i(\varphi_i)$ the likehood of it to happen and $P_i(x_i)$ represents the consequence, Their combination will provide is the risk as it is explained below in Figure 9.

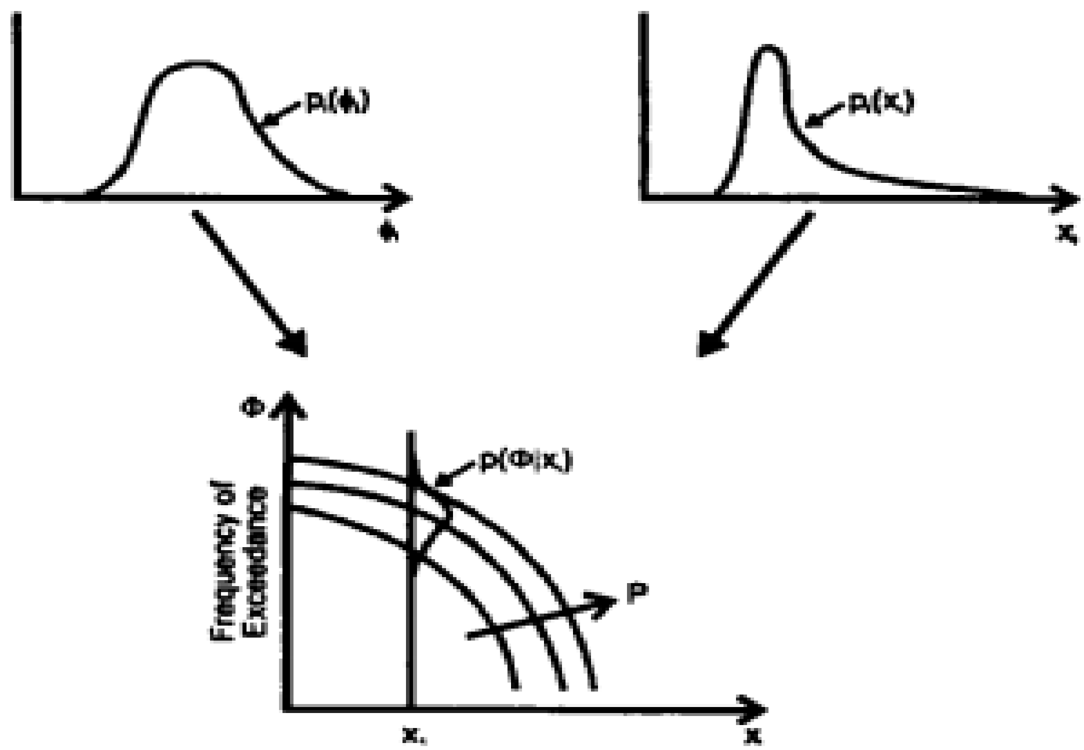

Figure 9 Graphical portrayal of risk Source: Kaplan, 1997

Each type of accident will be a scenario. We will use variables that can be affected by the ship owner in order to indicate which combination bears less risk and which variable has larger impact

3.5. ***Conclusion***

By collecting data from several databases regarding maritime accidents and with the help of statistical tools to transform the data to qualitative and categorical. A data set will be created in order to help in the construction of a BBN. This network will allow us to study the cause relationship between the variables and their behavior during maritime accidents. Next will calculate the risk profile for each type of accident and try to identify the operation strategy with the less risk.

4. Empirical analysis

4.1. ***Introduction***

In this chapter we will explain first how we treat the data and transformed them into discrete variables. Then we will present the summary of the data, specifically for each variable the corresponding frequency table will be presented. The data will be discretized in order to be better processed by the algorithm; each category of a variable will be represented by a number. With the help of an algorithm a Bayesian network will be generated which will illustrate the discovered relationships between the variables based on the data and our theoretical framework. In the last part of our analysis we will assign the probabilities of each state based on our Bayesian network and our sample and we will compare four operational strategy in terms of risk.

4.2. ***Summary of data***

4.2.1. Introduction

In this section we will explain how we created the dataset and discretized our data. Moreover, the frequency of each state of the variables will be calculated. As mentioned in previous chapter our data sources are reports of accidents and the databases of Paris MoU and Tokyo MoU organizations. After examining the maritime accident reports of the years 2016-2018 and by cross checking the IMO number of each vessel with the aforementioned organizations, we created a data frame. Since our model will be processed in R language, the state of each variable will be presented with an integer which will be quoted in column next to the name of the state.

4.2.2. Accident rate

As mentioned earlier we investigate maritime accidents in the time period 2016-2018. Totally we gathered information for 382 vessels which are split per year as it is described below in Table 9.

Year	Investigated vessels	Frequency per year
2016	175	1 accident per 2,4 days
2017	155	1 accident per 2,1 days
2018	52	1 accident per 7 days

Table 9 Accident rate per year

By the frequency of accidents, we can observe that the last three years the frequency of accidents reported is diminished. We may assume that this is a result of the Authorities to ensure that ship companies are following the necessary safety procedures during onboard operations. It is worth mentioning that during 2017 there were extreme weather phenomenon such as the Hurricanes “Harvey”, “Irma” and “Maria”. These hurricanes are considered as among the most destructive maritime incidents the last years, the losses caused by them were estimated up to $1 bn.(David, 2019)

4.2.3. Variable 1: Crew

This variable represents the number of the crew that was on board when the accident occurred as it is mentioned in the corresponding report of IMO. In case this information is missing from the report then we assumed that the vessel was fully manned, and the number of crew is as it is described at the “Number of crew on ship's certificate” in Annex 1 of IMO report. Moreover, if there are other personnel onboard such as pilots of assisting tugs or surveyors, we count them in the crew since they have active role in the operation of the vessel, or they are engaged in the accident. The summary of the analyzed data is represented in the Table 10 below.

Variable: crew	Numeral representative	count	Rel. freq.
"0-10"	1	42	11,00%
"11-20"	2	173	45,30%
"21+"	3	167	43,70%

	Sum:	382	100%

Table 10 Relative frequencies for the Variable: Crew

Most of the investigated vessels had more than 10 people on board during the accident. According to the published reports of EMSA most of the ships are well manned. Only 1% of the investigated accidents indicate manning as contribution factor.(EMSA, 2015) . Therefore, we can assume that the investigated vessels in this research were also well manned in terms of number of crew.

4.2.4. Variable 2: Type of the vessel

This variable represents the type of the vessel which is engaged in a maritime time accident during the selected time period. As mentioned in the first chapter we will study cases of cargo ships and the selected types are mentioned below as they are named in Annex 1 of IMO reports:

1	Bulk Dry (general, ore) Carrier	7	Oil Tanker
2	Bulk Dry / Oil Carrier	8	Other Bulk Dry (cement, woodchips, urea and other specialized) Carrier
3	Chemical Tanker	9	Other Dry Cargo (livestock, barge, heavy cargo, etc.) Carrier
4	Container Ship	10	Other Liquids (non-flammable) Tanker
5	General Cargo Ship	11	Ro-Ro Cargo Ship
6	Liquefied Gas Tanker	12	Refrigerated Cargo Ship

Table 11 Type of ships base on categorization of IMO

Since some categories are similar, we created six categories that will include these twelve types and are explained in Table 12

Variable: type of vessel	Numeral representative	Count	Rel. Freq
General cargo ship	1	107	28,00%
Liquified Gas Tanker	2	11	2,88%
Container Ship	3	76	19,90%
Dry Bulk/Cargo Carrier	4	97	25,40%
Tanker (Oil / Chemical/Liquid)	5	73	19,10%
Rest	6	18	4,72%
		382	100,00 %

Table 12 Relative frequencies of the variable: Type of vessel

General cargo ships (28%) suffered more accidents followed by Dry bulk (25,4%). Tankers (19,10 %) and container ships (19,90%) have almost the same number. These four categories also represent the majority of the world fleet (89%) therefore, it was expected that they will suffer more accidents than the rest fleet. Specifically, General cargo ships are often operating in short sea shipping routes. In addition, their activity is mainly concentrated in tramp trading, therefore they are involved in many port calls and they are navigating often through congested coastal waters, where the risk of accidents is higher.

4.2.5. Variable 3: Gross Tonnage of the Vessel

The variable of Gross Tonnage is an indicator of the vessel size. As Zhang et all indicates as the gross tonnage of a vessel increase, also her size increase since the length and the draught are increasing.

Variable: Gross Tonnage	Numerical representative	Count	Rel. Freq
<500	1	10	2,62%
500 - 1000	2	12	3,13%
1000 - 3000	3	57	14,90%
3000 - 5000	4	30	7,85%
5000 - 10000	5	44	11,50%
10000 - 30000	6	90	23,60%
30000 - 50000	7	76	19,90%
> 50000	8	63	16,50%
	SUM	382	100%

Table 13 Relative frequencies of the variable: Gross tonnage

The data in Table 13 indicates that vessels with gross tonnage larger than 10.000 are prone to accident than smaller vessels. Bigger ships due to their size they require more time in order to complete maneuverer moves and avoid an accident such as grounding or collision. Their structure also is more complicated since it is required bigger engines and more machinery systems to support their daily operations.

4.2.6. Variable 4: Category of vessel's flag

As explained in previous chapter every vessel is carrying a flag which is related regarding the vessel's obligation and rights in terms of safety, costs and taxes. This variable is describing the categorization of the flag of the vessel by Paris MoU. The organization is separating the flags in three categories based on their performance.

Variable: Category of the Flag	Numerical representative	Count	Rel. freq.
Black	1	24	6,28%
Grey	2	13	3,40%
White	3	345	90,32%
	SUM	382	100,00 %

Table 14 Relative frequencies of the variable: Category of Flag

Table 14 indicates that vessels carrying a flag which is listed as "white" by Paris MoU are prone to accidents. Even if "black" are known for their high detention rate and their not responsible behaviour, vessels carrying them present lower frequency of accidents than vessels carrying white flags. this indicated that there are flags among "white" flags which are reluctant to impose the proper safety standards.

4.2.7. Variable 5: Flag of convenience

This variable is indicating if the flag of the vessel in the moment of the accident was a FOC or not.

Table 15 Relative frequencies of the variable: Flag of Convenience

Variable: flag of Convenience	Number representative	count	rel. freq.
Yes	1	191	50%
No	2	191	50%
	Sum	382	100%

The data indicate that either a vessel carried a flag of convenience or not, she had the same frequency (50%) in participating in an accident. Therefore, we may assume that flag of convenience as a general category is not related with the maritime accidents.

4.2.8. Variable 6: Change of flag

This variable indicates if the vessel had changed her flag before the accident or not. This information is also enclosed in Annex 1 of IMO reports, therefore we assumed that if the corresponding part of the report is blank then the vessel had not changed her flag before.

Variable: Change of Flag	Number representative	count	rel. freq.
No	1	210	55%
Yes	2	172	45%
	Sum	382	100%

Table 16 Relative frequencies of the variable: Change of flag

The data indicates that there is not a big difference between the two states of the variables. The vessels that did not changed their flag and had an accident are slightly more. This proves that fact that a vessel changed a flag is not individually a strong indicator that it is a ship with higher risk profile.

4.2.9. Variable 7: Change of Name

The seventh variable indicates if the vessel had changed her name before the accident according to Annex 1 of the investigated IMO report.

Variable: Change of Name	Number representative	count	rel. f req.
No	1	152	39,80%
Yes	2	230	60,20%
	Sum	382	100%

Table 17 Relative frequencies of the variable: Change of Name

From the summary of the data in Table 17 we observe that vessels which have changed their name before are more prone to accidents. Changing the name of a vessel may be a business strategy for some ship owners to conceal controversial operational history of their vessel.

4.2.10. Variable 8: Age of the Vessel

The eighth variable of our model is the age of the vessel at the year of the accident. Since are variables are transformed to discrete, we grouped the data in four categories as it is indicated below in Table 18.

Variable: Ship Age	Number representative	count	rel. freq.
0-5	1	60	15,70%
6-15	2	172	45,00%
16-25	3	78	20,40%
25+	4	72	18,80%
	Sum	382	100%

Table 18 Relative frequencies of the variable: Age of the vessel

The data indicates that young ships are less prone to accidents. Ships with age between 6 to 15 years old (45%) tend to suffer more accidents. Generally, we can assume that the older the vessel the higher is the risk, since the operational costs are higher and ship owner tend to be less meticulous regarding the vessel's maintenance.

4.2.11. Variable 9: Season

This variable indicates the season during which the accident occurred. We had split the year in 4 seasons as it explained in Table 19 below.

Variable: Season	Number representative	count	rel. freq.
Jan - Mar	1	130	34,00%
Apr - Jun	2	86	22,50%
Jul - Sep	3	85	22,30%
Oct -Dec	4	81	21,20%
	Sum	382	100%

Table 19 Relative frequencies of the variable: Season

The data indicates that during the months January, February and March the frequency of is higher (34%). The rest of the year the frequency is roughly the same ($\approx 22\%$). At this point is worth mentioning that generally during these months the sea state is relatively rougher, the average wave height is larger, and the water velocity is stronger. This make the operation of a vessel more demanding and dangerous.

4.2.12. Variable 10: Time

The tenth variable indicates the time of the day (local time) that the accident occurred. The 24 hours of a day are divided in four equal parts as they are described in Table 20.

Variable: Time	Number representative	count	rel. freq.
(00-06]	1	113	29,60%
(06-12]	2	96	25,10%

(12-18]	3	95	24,90%
(18 - 24]	4	78	20,40%
	Sum	382	100%

Table 20 Relative frequencies of the variable: Time

The statistics of the data shows that during the night accidents have higher frequency (29,60%). The reasons behind this high frequency is first the lack of visibility, and the second the augmented fatigue and sleepiness that crew suffers during night shifts.(Price, 2011)

4.2.13. Variable 11: Location of the accident

This variable is describing the location of the vessel during the accident in relation with her distance from the land. Therefore, we created three qualitative states for this variable. The first one is called "In harbor" which refers to location such ports or terminals. The second state is called "Coastal waters", this state represent surface between the land and ocean. In our study the length of this interface is up to12 nautical from the land. The area of anchorage near the port is considered also as "Coastal waters". The third state of this variable is "Deep Sea" which describes the interface of the ocean after the 12th mile from the land. The fourth state of this variable is called "narrow waters" which describes locations such as canals, rivers or straits. We examined in detail the accident reports in order to define the location of the vessel. If there was not a clear indication about the location of the vessel within the Annex of the report or the description of the accident, we assumed that the vessel was in deep sea waters.

Variable: Location of the accident.	Number representative	count	rel. freq.
In harbour	1	85	22,30%
Coastal waters	2	125	32,70%
Deep sea	3	107	28,00%
Narrow waters	4	65	17,00%
	Sum	382	100%

Table 21 Relative frequencies of the variable: Location of the accident

The location with the higher frequency is the coastal waters which is a result of the hydrographic uncertainty that is associated with coastal waters. More specifically there are uncharted obstacles such as submerged reefs and rocks which increase the complexity of these waters.

4.2.14. Variable 12: Operational condition of the vessel

This variable is indicating the operational condition of the vessel during the accident. As previously we created four qualitative states for this variable which describe the condition of the vessel in terms of operation status. The first status is named "Sailing" under which the vessel is traveling with more than 6 knots to her target destination. The second stand state is called "Maneuvering" under this state the vessel is moving with less than 6 knots and is doing slow correcting moves such as approaching the berth or an anchorage spot In case the vessel is moving with the assistance of tugs we assume that her condition is described by this state. The third state is called "At Berth" where the vessel is anchored in a port or terminal. The fourth state is called "Loading/Discharging", during this condition the vessel is at berth and operations are occurring regarding the load/unload of the cargo. If the ongoing operations are not related with the cargo, then we assume that the vessel is at berth or in anchorage. The fifth state of this variable is named "Anchorage" where the vessel is not moving and is in the anchorage area near a port or terminal. Operations such as refueling from bunker ship in the sea without movement speed are enclosed in the last state.

Variable: Operational Condition	Number representative	count	rel. freq.
Sailing	1	217	56,80%
Manoeuvring	2	65	17,00%
At Berth	3	24	6,28%
Loading/Discharging	4	27	7,07%
Anchorage	5	49	12,80%
	Sum	382	100%

Table 22 Relative frequencies of the variable: Operational condition

The summary of data in Table 22 indicates that accident happen more frequently when the vessel is "sailing" (56,80%). Since the vessel during sailing is fully operational, we may assume that this state put the higher pressure to the structure of the vessel and her machinery. Also, crew alertness tends to drop during "sailing" due to augmented fatigue and lack of awareness of risky factors.

4.2.15. Variable 13: Performance in Paris/Tokyo MoU

The data for this variable is gathered by cross checking the IMO number of the investigated vessel with the databases of Paris MoU and Tokyo MoU. We used the IMO number because its unique for each vessel and cannot be altered. Both organizations are making surveyors in the ship and investigating if any regulation regarding the labor condition or safety standards is not followed. If there is some deficiency or non-compliance with the safety standards, then a ship is detained or marked as "high risk ship". In this variable we created two states, which describe whether if the ship was detained or marked as high risk prior to the accident or not.

Variable: Performance in Paris/Tokyo MoU	Number representative	count	rel. freq.
Yes	1	177	46,30%
No	2	205	53,70%
	Sum	382	100,00%

Table 23 Relative frequencies of the variable: Performance in Paris/Tokyo MoU"

The summary of the data in Table 23 does not show a clear trend since the frequency is almost equally divided. As explained earlier this variable is an indicator if a vessel has a responsible operational performance or not. By the first look of the data we cannot conclude that ships which are not operated by the book occur more accidents.

4.2.16. Variable 14: Initial event

The fourteenth variable is describing the initial event which was considered as the root cause of the accident. In order to define the root cause, we examined the information provided by Annex 2 of IMO report, and the description text of report. In this study we selected four main categories of root cause of an accident (Human error, Machinery/hull failure, Weather, Hazard cargo/cargo/another agent, Unknown). As Human error we describe situations where the crew made a wrong judgement or decision regarding the necessary safety procedures or the proper ship operations. Moreover, as human error is described also situations where a crew member couldn't fulfill properly his duty such as watch keep.

Finally, decisions made by the ship management company which affected the ship and put it in risk are also defined as human error. Machinery/hull failures are defined are defined as failures of the machinery or the hull structure of the ship such as engine break down. The weather is considered as root cause when it affected the operational capability of the ship due to high waves or rough sea state. "Hazard cargo/other agent" are defined as the situations where the ship was carrying dangerous cargo such as flammable substances. In some cases, the shipper had not informed properly the ship regarding the appropriate conditions in which the cargo must be transported. However, there are reports which doesn't describe what caused the accident therefore the root cause is "Unknown". Finally, the root cause of an accident may be a combination of the aforementioned factors therefore we created eight states to describe all the observed combinations.

Variable: Initial Event	Number representative	count	rel. freq.
Human Error & Machinery/hull failure	1	14	3,65%
Human Error & Weather	2	15	3,93%
Machinery/Hull Failure & Weather	3	10	2,62%
Human Error	4	159	41,60%
Machinery/Hull Failure	5	60	15,70%
Weather	6	17	4,50%
Hazard Cargo	7	14	3,70%
Unknown	8	93	24,30%
	Sum	382	100,00%

Table 24 Relative frequencies of the variable: Initial Event

The data shows clearly that the majority of Initial events are a Human Errors, which indicates that human decisions and judgement has a great contribution in maritime accidents. The adequate management of the safety on board is an outcome of teamwork and professionalism of the whole crew regardless if their position in on the bridge or in the engine room. The frequency of human errors may also indicate the lack of proper training of the crew regarding the safety procedures.

4.2.17. Variable 15: Type of Accident

This variable is describing the accident that occurred as it is described in the IMO report. In this study we will examine six group of type of accidents, each one will represent a state of this variable. The seventh state will represent accidents which are not included in the rest groups.

Variable: Accident Type	Number representative	count	rel. freq.
Occupational	1	115	30,01%
Grounding	2	48	12,60%
Collision	3	114	29,80%
Fire/explosion	4	40	10,50%
Allision/Contact	5	19	5,00%
Capsizing/ foundering	6	30	7,90%
Other	7	16	4,19%
	Sum	382	100,00%

Table 25 Relative frequencies of the variable: Type of Accident

The data shows that the majority of the accidents are occupational (30%) or collisions (30%) followed by grounding (12,6%) and fire/explosion (10,50%). Allision and Capsizing represent the least frequency which is respectively 5% and 7,90%. The high frequency of occupational accidents indicates the risk that seafarers bear every day and supports the statement of Mr. Tjibbe Joustra that working abroad is a dangerous job and seafarers deal with risks in daily basis.(Dutch Safety Board, 2009)

4.2.18. Variable 16: Severity of casualties

This variable represents the severity of the casualties as it is categorized from IMO.

Variable: Severity of the Casualty	Number representative	count	rel. freq.
Very serious	1	234	61,30%
Serious	2	113	29,60%
Less Serious	3	35	9,10%
Maritime Incident	4	0	0,00%
	Sum	382	100,00%

Table 26 Relative frequencies of the variable: Severity of Casualties

The data in Table 26 indicated that the majority of maritime accidents had very serious casualties (61,30%) followed by serious casualties (29,60%). The “very serious“ casualties are associated with total loss of the ship of human life, moreover their high frequency not only indicates how dangerous is the life on the sea but also how severe the consequences of an accident can be, not only in economics terms but also for the human life.

4.2.19. Variable 17: Vessel’s operational capability after the accident

This variable describes if the vessel was capable to continue her journey after the accident or she was not cable due to total loss of the vessel or major damages in her structure.

Variable: Capable operationally to continue?	Number representative	count	rel. freq.

Yes	1	233	61,00%
No	2	149	39,00%
	Sum	382	100,00%

Table 27 Relative frequencies of the variable: Vessel's operational capability after the accident

The data indicates that most of the vessels were capable to proceed their journey but still the frequency of major damages or total loss is relatively high (39%). From the perspective of ship management this frequency is relatively high and indicates the risk that a ship manager has to mitigate. As it is known shipping is a high capital industry therefore the loss of this capital due to an accident may have catastrophic consequences for the ship management company.

4.2.20. Variable 18: Number of injuries

The eighteenth variable describes if there was injuries and their number. This variable is described by three states. There is not distinction for the severity of the injuries. The data are collected from annex 2 of IMO report and the description text

Variable: Injuries	Number representative	count	rel. freq.
1-4 people injured	1	33	9%
>4 people injured	2	4	1%
No injuries	3	345	90%

Table 28 Relative frequencies of the variable: Number of injuries

From Table 28 we can observe that the frequency of reported injuries is relatively low (10%). As an individual figure, this is quite encouraging however it may not represent the actual number of injuries. There are injuries which are not reported by the ship manager companies if there is no need of assistance. Therefore, we may assume that the actual number of injuries on board due to accidents may be higher.

4.2.21. Variable 19: Number of Fatalities

This variable describes the number of fatalities occurred because of the accident. We are including fatalities which may not be related to the crew, but it was caused by the accidents such as fatalities related to stevedores or fishermen (in the case of collision with a fishboat). Moreover, the number of missing people in the IMO reports are considered as fatalities.

Variable: Number of Fatalities	Number representative	count	rel. freq.
1 - 5	1	152	39,80%
>5	2	10	2,60%
No fatalities	3	220	57,60%
	Sum	382	100,00%

Table 29 Relative frequencies of the variable: Number of fatalities

From the data we can observe that the frequency is fatality is relatively high (42,40%). This is another indication of the level of danger that working onboard is bearing.

4.2.22. Variable 20: Latitude

The summary of an IMO reports includes information regarding the location of the vessel during the accident in terms of coordinates. In case the information for the exact position of the vessel was missing we calculated the coordinates of the closest known location of the vessel. This Variable describes the Latitude of the vessel's coordinates. As it is known the world map is divided by imaginary horizontal lines which are expressing the distance of a point from Equator. The metric that is used is degrees, minutes and seconds. The scale is from -90 degrees to +90 degrees as it is illustrated in Figure 25 of Annex. In this study we divided the word map to 6 equal parts.

Variable: Latitude	Number representative	count	Rel. freq.
[+90,+60)	1	4	1,05%
[+60,+30)	2	176	46,10%
[+30,0)	3	149	39,00%
[0, -30)	4	38	9,95%
[-30,-60)	5	15	3,90%
[-60,-90)	6	0	0,00%

Table 30 Relative frequencies of the variable: Latitude of the vessel

The data indicated that accidents are relatively concentrated (85,1%) between specific latitudes [0,+60). This indicates a geographical pattern.

4.2.23. Variable 21: Longitude

This variable as the previous one is related with the region within the investigated maritime accidents occurred. Longitude are imaginary vertical lines which measure the distance of a point from the Prime Meridian which pass through Greenwich. The distance is measure in degrees, minutes and second. The world map is divided equally with twelve imaginary vertical lines as it illustrated in Figure 25 of Annex.

Variable: Longitude	Number representative	count	rel. freq.
[-180 , -150)	1	1	0,26%
[-150 , -120)	2	1	0,26%
[-120, -90)	3	6	1,57%
[-90, -60)	4	23	6,03%
[-60, -30)	5	11	2,89%
[-30,0)	6	17	4,46%
[0, +30)	7	93	24,30%
[+30, +60)	8	41	10,70%

[+60 , +90)	9	15	3,93%
[+90, +120)	10	92	24,10%
[+120, +150)	11	74	19,40%
[+150, +180)	12	8	2,10%
	Sum	382	100,00%

Table 31 Relative frequencies of the variable: Longitude of vessel

The relative frequencies distribution of the data indicates that most of the accidents occurred between specific Longitudes as we observed with the latitude also. Further investigation of the geographical pattern will be followed later.

4.3. *Construction of Bayesian Network*

4.3.1. Introduction

The frequency table for each variable shows that in some states there are elements which are more frequent than others based on our sample however in order to process more our data we will construct a Bayesian neural network(BNN) which will highlight the causal relationships between our variables. In this section of the chapter the structure of the Bayesian network will be illustrated with a Directed acyclic graph. This graph includes nodes which represents the variables and arc which illustrate the existence of relationship between the nodes. For our research we will use R studio which is an application that utilizes R language. The package that we will use for the process of our data and the graphical illustration are: “bnlearn”, ”grid”, “Rgraphviz”, “tidyverse”, “ggplot2”. As mentioned in chapter 3 the algorithm we will use is called High Climbing algorithm (“hc”). First part of this step of our research is to import the discretized data in R studio and create a data frame. For each variable as we explained number of states which in R language are called “levels”. The second part of this process is to use the function “hc”. (Scutari, 2010)

4.3.2. Model not based on theoretical framework

In our first attempt we did not implemented our theoretical framework as it is explained in chapter 2. The outcome of this attempt had a structure as it is illustrated in Figure 10 below. By observing the established relationships, we notice that there are relationships (arcs) which are directed wrong, such as “type of accident” -> “time”. Of course, the type of accident cannot affect the time. As Mascaro et all mention the investigation of causal relationships base on joint sample data may often produce arcs with the opposite direction, in the anti-casual direction as it is called. This is happening because the information provided by the sample is thin regarding uncovered relationships.(Mascaro, Nicholson, & Korb, 2014) For example, two parents’ nodes of a child node may not be directed related. In this case the time of accident may affect the type of the accident indirectly. In order to solve this obstacle, we utilized some prior constraints as they are explained in our theoretical framework in section 2.3.

Figure 10 Structure of the Bayesian Belief Network without implementation of the theoretical framework

4.3.3. Model based on theoretical framework

After the implementation of the constraints described in our theoretical framework, we reproduce the structure of the BNN which is illustrated in Figure 11 below. By implementing our theoretical framework, we observe that new arcs have been established or some of the old ones may have changed direction. The rest of our study will be based on the second BBN.

4.3.4. Comparison of the two models

As mentioned before the outcome of the first attempt was a BNN which included arcs with anti-casual directions. In the second attempt we included out theoretical framework. We observe that arc with anti-casual direction have disappeared and the structure of the network is similar to our theoretical framework. The second network not only is more related to our theoretical framework, but we believe that is also describes better our data.

BBN	AIC score
Not based on theoretical framework	AIC_1 = - 8017.503
Based on theoretical framework	AIC_2 = - 24256.74

Table 32 Comparison of score of the two networks

In order to compare statistically the two networks, we calculated the individual “AIC” score with the help of function “AIC”. The results are summarized in table 32 below.

Since $AIC_2 < AIC_1$ then the second model fits better than the first one.(Bertrand, Sakamoto, Ishiguro, & Kitagawa, 1988).

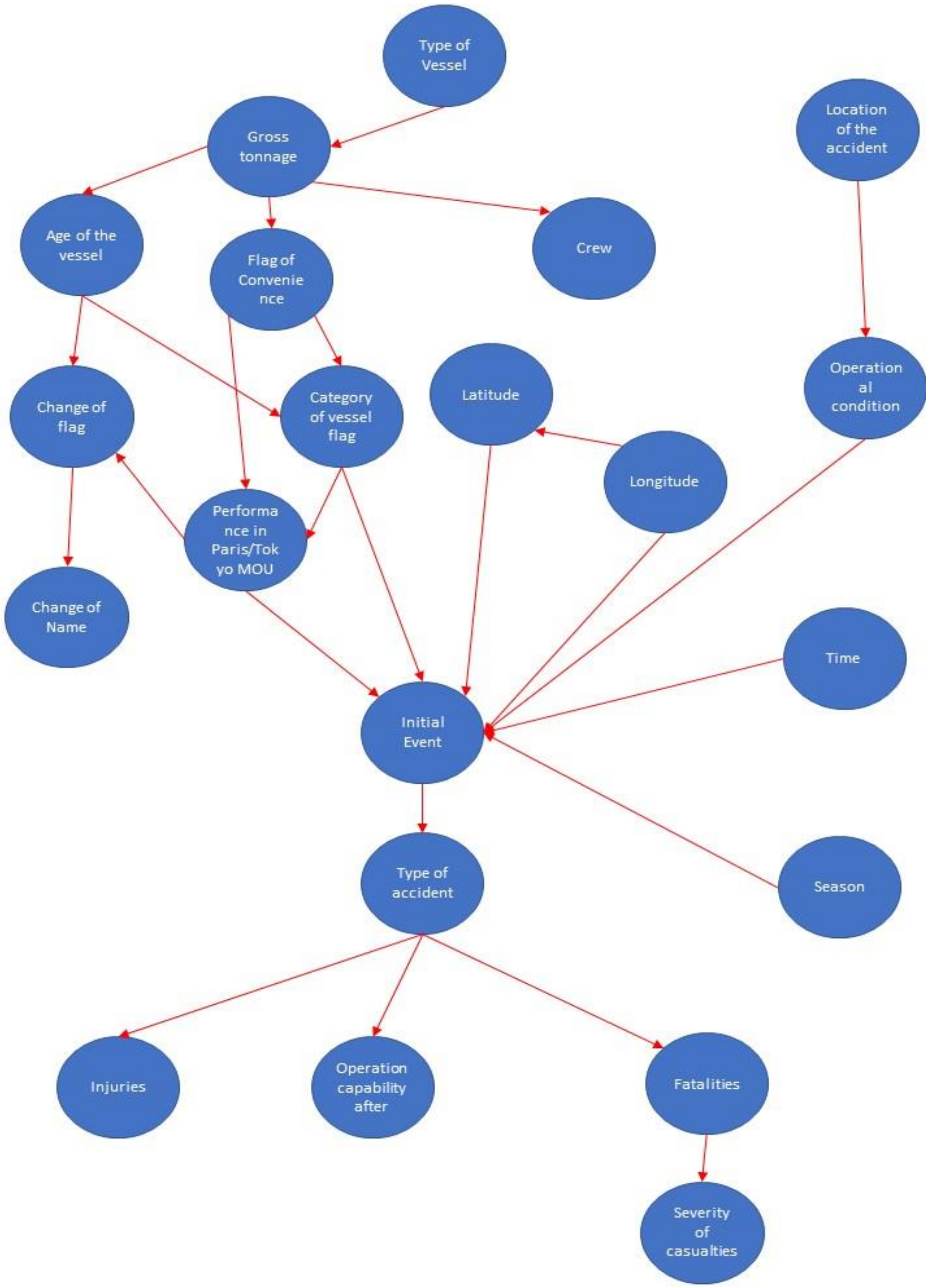

Figure 11 Structure of the Bayesian Belief Network after the implementation of the theoretical framework

4.4. ***Analysis of selected model***

4.4.1. Introduction

In the previous section we created the structure of the BBN. In this section we will study in detail each node, the established relationships and we will assign the corresponding conditional probabilities for each state of the nodes. We believe that through this analysis we will be able to make more specific conclusions regarding the investigated accidents.

4.4.2. Node 1: Type of vessel

Parents: none

Children: "Gross Tonnage", "Initial Event"

The variable of "Type of Vessel" is not affected by any other variable. But it affects two other variables which are the "Gross Tonnage" and the "initial event". In other words, it affects the size of the vessel and the initial event that caused the accident.

The Corresponding conditional probabilities table is:

General cargo ship	Liquified Gas Tanker	Container Ship	Dry Bulk/Cargo Carrier	Tanker (Oil / Chemical/Liquid)	Rest
27,98%	2,92%	19,89%	25,37%	19,10%	4,74%

Table 33 Conditional probabilities of each type of vessel

As we observed also in the section 4.2. the type of vessel that presents the higher probability of accident is the "General Cargo" ships. These vessels are mainly used in tramp shipping; therefore, they are expected to do more port calls and visit sea routes with higher traffic. We believe due to these characteristics of their operation profile they present higher probability of accidents.

4.4.3. Node 2: Gross tonnage

Parents: Type of Vessel

Children: Crew, Flag of Convenience, Ship Age, Initial event

The node of "Gross tonnage" which illustrates the size of the vessel is affected by the type of the vessel. This was expected since each type of vessel correspond to different size due to the characteristics of the cargo they transfer of the complexity of the machinery. The node of "Gross tonnage it affects the number of the crew, the bigger the vessel the bigger the number of crew. But if affects also if the flag is a flag of convenience, the age of the vessel and the initial event of the accident.

	Type of vessel					
Gross tonnage	General cargo ship	Liquified Gas Tanker	Container Ship	Dry Bulk/Cargo Carrier	Tanker (Oil / Chemical/Liquid)	Rest
<500	8,42%	0,19%	0,03%	0,02%	1,40%	0,11%
500 - 1000	7,48%	0,19%	0,03%	0,02%	5,50%	0,11%
1000 - 3000	38,28%	18,10%	3,97%	0,02%	9,60%	22,13%
3000 - 5000	16,82%	9,14%	3,97%	0,02%	5,50%	22,13%

5000 - 10000	20,55%	9,14%	7,90%	0,02%	15,06%	22,13%
10000 - 30000	7,48%	27,05%	31,54%	27,81%	31,46%	27,64%
30000 - 50000	0,95%	27,05%	21,03%	49,42%	10,96%	0,11%
> 50000	0,02%	9,14%	31,54%	22,66%	20,53%	5,62%

Table 34 Conditional probabilities of each type of vessel in comparison with size

By studying the conditional probabilities in Table 34 we observe that within categories there are size of vessel which present higher probability of accident than other. More specifically we observe that general cargo ships with Gross Tonnage less than 10000 are more prone to accidents (91,55%). This highlights the fact that this size of vessel which is mainly used in tramp shipping and have frequent port calls are facing more dangers. For the category of the Liquefied Gas Tankers the results indicated that vessel with Gross Tonnage higher than 10000 suffer more accidents (63%). This type of ship transfers flammable and explosive cargo. Moreover, as the size of the vessel grows the complexity of the vessel's operation is increasing, there are more tanks on board which may be a factor of increasing the probability of accidents. The same situation seems to be for Containerships, Dry Bulk Carriers and tankers which have Gross tonnage more than 10000 and acquire higher probability of accident, respectively 84%, 99% and 63%. For the rest type of investigated cargo ships the vessel with gross tonnage less than 10000 are more prone to accidents (66,5%). Regarding the strength of the arc we can conclude that this is an "relatively strong relationship" since its score is -1,20206E+02, thus the network's score will be decreased by the elimination of this arc.

4.4.4. Node 3: Crew

Parents: Gross tonnage

Children: None

This node is only affected by the size of the vessel. It is worth mentioning that this node based on our BNN is not connected to any other variable, therefore it is not a contribution factor to maritime accidents

Crew \Gross Tonnage	<500	500 - 1000	1000 - 3000	3000 - 5000	5000 - 10000	10000 - 30000	30000 - 50000	> 50000
"0-10"	79,42%	41,58%	36,83%	13,42%	0,09%	3,37%	1,37%	0,07%
"11-20"	20,16%	58,08%	57,84%	76,49%	68,08%	42,21%	34,21%	22,24%
"21+"	0,41%	0,34%	5,32%	10,10%	31,82%	54,42%	64,42%	77,69%

Table 35 Conditional probabilities of number of crew member in comparison with the size of vessel

By observing the data in Table 35 can observe clearly that the bigger the Bessel the bigger is the number of crew. As explained earlier this is related to the higher complexity of the ship operations on board. Regarding the strength of this arc we may say that this arc is very strong since the corresponding score is -8,44942E+01 , which indicates that the absence of this arc will affect heavily the score of the network.

4.4.5. Node 4: Age of the vessel

Parent: Gross tonnage

Children: Change of flag, Change of Name, Initial event, Category of flag

By observing the structure of the network, we can observe that the node "Age of the vessel" is affected only by the gross tonnage of the vessel and it affects the operation of the vessel regarding if it changes Name and/or flag. It is related also with the initial event of the accident and the category of the flag

Age of Vessel \ Gross tonnage	<500	500 - 1000	1000 - 3000	3000 - 5000	5000 - 10000	10000 - 30000	30000 - 50000	> 50000
0-5	0,31%	0,26%	3,56%	6,74%	2,34%	20,01 %	27,63 %	25,40 %
6-15	20,06 %	8,51%	26,31 %	33,30 %	47,66 %	41,09 %	63,10 %	60,25 %
16-25	29,94 %	25,00 %	21,06 %	26,66 %	22,73 %	31,10 %	6,61%	14,31 %
25+	49,69 %	66,24 %	49,07 %	33,30 %	27,27 %	7,80%	2,67%	0,05%

Table 36 Conditional probabilities of Age of Vessel in comparison with the size of vessel

By looking the Table 36 we may observe that for each category of vessel's size the corresponding category of age which is more prone to accident may change. Specifically, we observe that small vessels with gross tonnage less than 5000 and age above 15 years are more prone to accidents ($\approx 70\%$) than the younger one. For midsize vessels with gross tonnage between 5000 and 30000 we observe that mid aged vessels are more prone to accidents ($\approx 45\%$) On the other hand for bigger ships we observe that young and mid-aged ships show the higher probability. This may be an indication that the crew in new big ships is not well trained or experienced.

In terms of arc strength this arc has score -6.310.997, consequently, is a relatively strong arc and important for the network

4.4.6. Node 5: Flag of convenience

Parents: Gross tonnage

Children: Category of Flag, Performance in Paris/ Tokyo MoU, Initial event.

This node is only affected by the size of the vessel based on the structure of our network. However it affects three other nodes which are the category of the flag , The performance of the vessel in Paris/Tokyo MoU and the initial event of the accident.

Flag of convenience / Gross tonnage	<500	500 - 1000	1000 - 3000	3000 - 5000	5000 - 10000	10000 - 30000	30000 - 50000	> 50000

yes	10,49 %	25,26 %	28,12 %	46,68 %	47,73 %	48,89 %	65,76 %	66,63 %
no	89,51 %	74,74 %	71,88 %	53,32 %	52,27 %	51,11 %	34,24 %	33,37 %

Table 37 Conditional probabilities of the "Flag of Convenience" in comparison with the size of the vessel

The data shows that vessels with size smaller of 10000 GT and carrying a non-flag of convenience are more prone to accidents. However, as the size of the vessel grows above 1000 GT the balance shifts and those which carry a flag of convenience are more prone to accidents. It is known that as the size of the vessel is growing so is the number of the crew and the operational cost. A known strategy to reduce these costs is carrying flag of convenience which is associated with less taxes, fewer obligations and cheaper crew. But cheaper crew may be associated with less experienced.

Regarding the strength of this arc, we observe that the score is -1,14811E+01, thus is a relatively strong arc and important for the network

4.4.7. Node 6: Change of flag

Parents: Age of vessel, Performance in Paris/Tokyo MoU

Children: Change name

This node is affected by two "parents" nodes. The first one is the "Age of the vessel" and the second one is the "Performance in Paris/Tokyo MoU". It also has causal relationship with the node" Change name".

The conditional probabilities if the vessel is marked as "high risk" are:

Change of Flag/ Age of Vessel	0-5	6-15	16-25	25+
No	87,21%	56,09%	32,55%	15,50%
yes	12,79%	43,91%	67,45%	84,50%

Table 38 Conditional probabilities of the Age of vessel regarding the fact it changed flag state or not

We observe that ships with risky behavior tend to change flags as they get older. They chose to flag which allows them to keep trading even if they are aged vessels, or they are not following the corresponding safety regulations.

The Conditional probabilities if the vessel is not marked as "high-risk" are:

Change of Flag/ Age of Vessel	0-5	6-15	16-25	25+
No	86,26%	75,52%	39,51%	30,38%
yes	13,74%	24,48%	60,49%	69,62%

Table 39 Conditional probabilities of Age of vessel regarding the fact if they changed flag state or not.

Again as in Table 38 also in Table 39, the ships change flags as they get older. However, the trend is slightly smaller in this case. As main conclusion of this node we may state that vessels as they get older, they tend to change flags in order to be able to continue trading. This happens more often if the ship is marked as "high-risk" ship before.

Regarding the strength of these two arcs, the score of the arc "Age of vessel" to "Change of flag" is -3,21961E+01 and the score of "Performance in Paris/Tokyo MoU" to "change of flag" is -1,02189E+00. Both arcs seem important to the structure, but the first one is relatively stronger than the second one.

4.4.8. Node 7: Change of Name

Parents: Change of flag, Age of the vessel

Children: none

The Node of "Change of Name" is affected by "Change of flag" and "Age of the vessel". This node doesn't have any children therefore it is not connected with the causes of the accident or the type of accident.

For ship age = 0-5 years:

Change of Name\change of flag	no	yes
no	84,53%	13,08%
yes	15,47%	86,92%

Table 40 Conditional probabilities of "Change of name" in comparison with the fact the vessel changed flag or not (ship age 0-5 years)

For ship age = 6-15 years

Change of Name\change of flag	no	yes
no	67,52%	10,43%
yes	32,48%	89,57%

Table 41 Conditional probabilities of "Change of name" in comparison with the fact the vessel changed flag or not (ship age 6-15 years)

For ship age = 16-25 years

Change of Name\change of flag	no	yes
no	46,44%	4,11%
yes	53,56%	95,89%

Table 42 Conditional probabilities of "Change of name" in comparison with the fact the vessel changed flag or not (ship age 16-25 years)

For ship age = 25+

Change of Name\change of flag	no	yes
no	37,60%	5,46%
yes	62,40%	94,54%

Table 43 Conditional probabilities of "Change of name" in comparison with the fact the vessel changed flag or not (ship age 25+ years)

From the above tables of conditional probabilities, we can observe If the flag of the vessel is changed then the probability to the name to be changed also is highly increased. Moreover, as the ship get older then ship then this probability it gets higher. However, this relationship as mentioned earlier is not connected with the accident or its root causes.

Regarding the strength of the arc we can see that the arc "Change of flag" to "Change of name" has score -4,76730E+01 which indicated that it is a strong arc. In addition the arc "Age of vessel" to "Change of name" has score -4,69403E+00 which shows that this relationship is relatively strong .

4.4.9. Node 8: Category of vessel flag

Parents: Flag of convenience, Age of the vessel

Children: Performance in Paris/Tokyo MoU, Initial event

From the structure of the vessel we can observe that the Node of "Category of Vessel's Flag is affected by the nodes "Flag of convenience" and "Age of the Vessel" and is affects the nodes "Performance in Paris/Tokyo MoU" and the "Initial event" of the accident.

For Age of Vessel = 0-5 years

Category of Flag |Flag of Convenience	Yes	No
Black	0,14%	0,14%
Grey	3,46%	0,14%
White	96,40%	99,72%

Table 44 Conditional probabilities of "Category of Flag" in comparison with the variable "Flag of Convenience" (ship age :0-5 years)

For Age of vessel = 6-15 years

Category of Flag |Flag of Convenience	Yes	No
Black	1,05%	1,42%
Grey	0,04%	4,16%
White	98,91%	94,42%

Table 45 Conditional probabilities of "Category of Flag" in comparison with the variable "Flag of Convenience" (ship age 6-15 years)

For Age of vessel =16-25 years

Category of Flag |Flag of Convenience	Yes	No
Black	0,11%	15,06%
Grey	0,11%	7,58%
White	99,78%	77,36%

Table 46 Conditional probabilities of "Category of Flag" in comparison with the variable "Flag of Convenience" (ship age 16-25 years)

For Age of vessel = 25+ years

Category of Flag| Flag of Convenience	Yes	No
Black	16,75%	25,02%
Grey	8,46%	8,40%
White	74,78%	66,58%

Table 47 Conditional probabilities of "Category of Flag" in comparison with the variable "Flag of Convenience" (ship age: 25+ years)

By observing the conditional probabilities of the states of this node we can observe that as the vessel gets older the probability that a ship caries a "black" flag non-convenience is higher (from 0,14% to 25%). This is an indication that there is part of ship manager companies which prefer to manage the aged vessels with a controversial way. However, the majority of the vessels which were engaged in an accident carried a "white" flag. Which indicate that ship manager companies prefer "white" flags which allows the vessel to access all the ports without restrictions.

The Arc "Flag of convenience" to "Category of Flag" has a score -2,24920E+00, thus it is a relatively strong arc. Regarding the arc "Age of Vessel" to "Category of Vessel" the score is -1,48796E+01 which is indicates that is less strong than the first one.

4.4.10. Node 9: Performance in Paris/Tokyo MoU

Parents: Category of Flag, Flag of Convenience

Children: Change of Flag, Initial event

The node "Performance in Paris/Tokyo MoU" is affected by two nodes, which are the "Category of flag" and the "Flag of convenience". It affects also two nodes, these are the 'Change of Flag" and the "Initial Event" of the accident.

If the vessel is carrying a flag of convenience the conditional probabilities are:

Marked as High-risk | Category of flag	Black	Grey	White
Yes	79,03%	65,79%	51,91%
No	20,97%	34,21%	48,09%

Table 48 Conditional probabilities of "Performance in Paris/Tokyo MoU" in comparison with the variable "Category of flag" (ship is registered under a FOC)

If the vessel is not carrying a flag of convenience the conditional probabilities are:

Marked as High-risk | Category of flag	Black	Grey	White
Yes	89,13%	59,84%	32,73%
No	10,87%	40,16%	67,27%

Table 49 Conditional probabilities of "Performance in Paris/Tokyo MoU" in comparison with the variable "Category of flag" (ship is not registered under a FOC)

From Table 48 and Table 49 above we can observe that non-flag of convenience carrying vessels are performing better regarding the two selected authorities. This is an indication that regardless the type of flag ("Black"," Grey", "White") vessels carrying a flag of convenience have higher probability to be characterized as "high risk" ships or detained by an authority than the vessels which are not carrying a flag of convenience. Therefore, we may assume that there is higher probability of a vessel which is carrying a FOC to not follow the corresponding regulations. This probability is higher of specific flag is characterized by Paris MoU as Black or Grey. This verify the assumption that shipping companies select flag of convenience not only for smaller taxation but because they are also able to operate their vessel in a more controversial way which is related with less operational costs.

Regarding the investigated arcs, the arc "Category of flag" to "Performance to Paris/Tokyo MoU" has score -9,80015E+00 while the arc "Flag of Convenience" to "Performance to Paris/Tokyo MoU" has score -3,69305E+00. Thus, the first arc is stronger.

4.4.11. Node 10 & 11: Longitude & Latitude

The node of Longitude is not affected by any node, but it affects the type of accident and the latitude. On the other hand, the longitude is affected by the latitude and it affects the type of accident. We combined the two nodes and the corresponding probabilities in order to create the table 40 below which highlights that accidents follow a geographical pattern. More specifically the areas enclosed in the red rectangles in Figure 12 are the areas which present the higher probability of accident according to our data. We observe that the highlighted areas are sea regions with high traffic such as the straits of Malacca.

Regarding the Arc between Longitude and Latitude we observe that the corresponding score is -8,90770E+01 which show a strong arc.

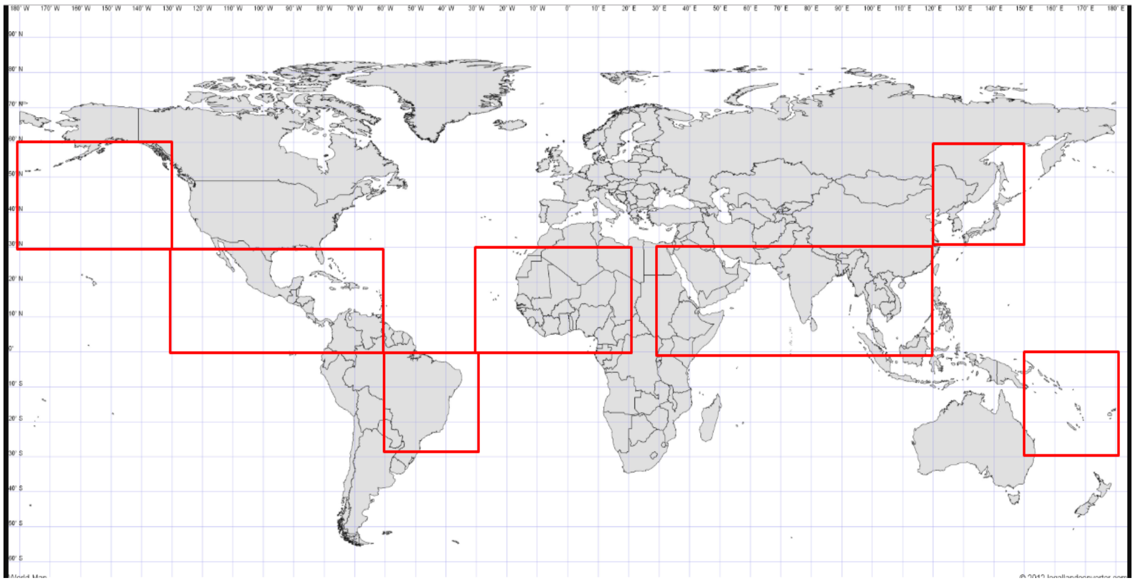

Figure 12 Areas with high probability of accident

Latitude |Longitude	[-180 , -150)	[-150 , -120)	[-120, -90)	[-90, -60)	[-60, -30)	[-30,0)	[0, +30)	[+30, +60)	[+60 , +90)	[+90, +120)	[+120, +150)	[+150, +180)
[+90,+60)	1,28%	1,28%	0,23%	0,06%	0,13%	5,93%	3,24%	0,03%	0,09%	0,02%	0,02%	0,17%
[+60,+30)	93,59%	93,59%	0,23%	21,72%	18,17%	70,33%	88,11%	36,54%	0,09%	7,62%	67,51%	12,54%
[+30,0)	1,28%	1,28%	98,86%	52,05%	18,17%	11,79%	5,39%	53,58%	99,54%	76,03%	20,27%	0,17%
[0, -30)	1,28%	1,28%	0,23%	17,39%	45,24%	5,93%	1,09%	7,34%	0,09%	16,30%	6,77%	49,66%
[-30,-60)	1,28%	1,28%	0,23%	8,72%	18,17%	5,93%	2,16%	2,47%	0,09%	0,02%	5,42%	37,29%
[-60,-90)	1,28%	1,28%	0,23%	0,06%	0,13%	0,08%	0,01%	0,03%	0,09%	0,02%	0,02%	0,17%

Table 50 Conditional probabilities of each region of sea

4.4.12. Node 12: Location of accident

Parents: none

Children: Operational condition

This node is another node which is not affected by another node, but it affects the node of “operational Condition”. This was expected since the type of operation condition of a vessel is defined by the location of the vessel.

In harbour	Coastal waters	Deep sea	Narrow waters
22,26%	32,70%	28,00%	17,04%

Table 51 Conditional Probabilities of the Variable “Location of accident”

The assigned conditional probabilities show that the probability of accident in coastal waters is higher (32,7%) followed by the “Deep sea” (28%). As mentioned earlier coastal water are usually congested waters with specific geographical characteristics since they include often sub-merged rocks.

4.4.13. Node 13: Operational condition

Parents: Location of the accident

Children: Accident Type

As mentioned previously the operational condition of a ship is defined mainly from the location of the ship and is influence the type of accident.

Table 52 Conditional Probabilities of the Variable “Operational condition” in comparison with the “Accident location”

	Accident location			
Condition Operational	In harbour	Coastal waters	Deep sea	Narrow waters
Sailing	2,40%	47,94%	99,81%	73,64%
Manoeuvring	34,08%	16,81%	0,05%	23,07%
At Berth	27,04%	0,84%	0,05%	0,08%
Loading/Discharging	30,56%	0,84%	0,05%	0,08%
Anchorage	5,92%	33,57%	0,05%	3,14%

Regarding the investigated accidents we observe in Table 52 that most of the accidents that happens in harbor area are occurred while the ship is maneuvering or is loading /discharging. This indicates that there is not adequate safety management while the ship is operating in the harbor. The crew is not judging properly the risk factors and the ship management company is not implementing the proper safety culture and training. In Coastal waters the higher probability of accident is occurred during sailing (47,94%) and Anchorage (33,57%) which shows that the complexity of coastal waters is not taken into consideration by the crew as it should. Moreover, we may assume that in the areas of anchorage ship are not communicating properly with nearby vessels on the corresponding safety measures are not followed. In the rest two categories the majority of accidents is occurred while the vessel is sailing which shows that the level of alertness on board can never drop.

Regarding the Arc between “Location of accident” and Operational Condition” the corresponding score is -1,70232E+02 which is an indication of a relatively strong relationship.

4.4.14. Node 14: Season

Parents: none

Children: Type of accident

The node of Season is not affected by any other node and it affects the type of accident.

Season	Jan - Mar	Apr - Jun	Jul - Sep	Oct -Dec
Probability	34,01%	22,52%	22,26%	21,21%

Table 53 Conditional probabilities of the variable: “Season”

The higher probability of accident is during wintertime. Generally during this period, the sea state is rough and the sea winds stronger

4.4.15. Node 15: Time of the day

Parents: none

Children: Type of Accident

This Node also is not influenced by any other node and it affects the type of accident.

(00-06]	(06-12]	(12-18]	(18 - 24]
29,57%	25,13%	24,87%	20,43%

Table 54 Conditional Probabilities of the variable “Time of the day”

According to the calculated conditional probabilities in Table 54 the time zone with the higher probability is between midnight and six in the morning. As mentioned earlier this time period has the least visibility and the crew on duty is often showing signs of fatigue or lack of sleep.

4.4.16. Node 16: Initial event

Parents: Type of vessel, Gross tonnage, Category of flag, Flag of Convenience, Age of the vessel, Performance in Paris/Tokyo MoU.

Children: Accident type

The node “Initial event” is one of the dependable variables of our model and important ones. This node has six parents’ nodes which influence it based on our theoretical framework. These nodes are the type of the vessel, her size, the type of flag she is carrying based on Paris MoU categorization, the age of the vessel and the performance based on the reports of the two selected authorities. However, since the node has six parents’ nodes there are 18342 possible combinations, therefore, we decided to study these combinations which present very high probability such as 99.99%. By studying these combinations, we discovered the finding explained below.

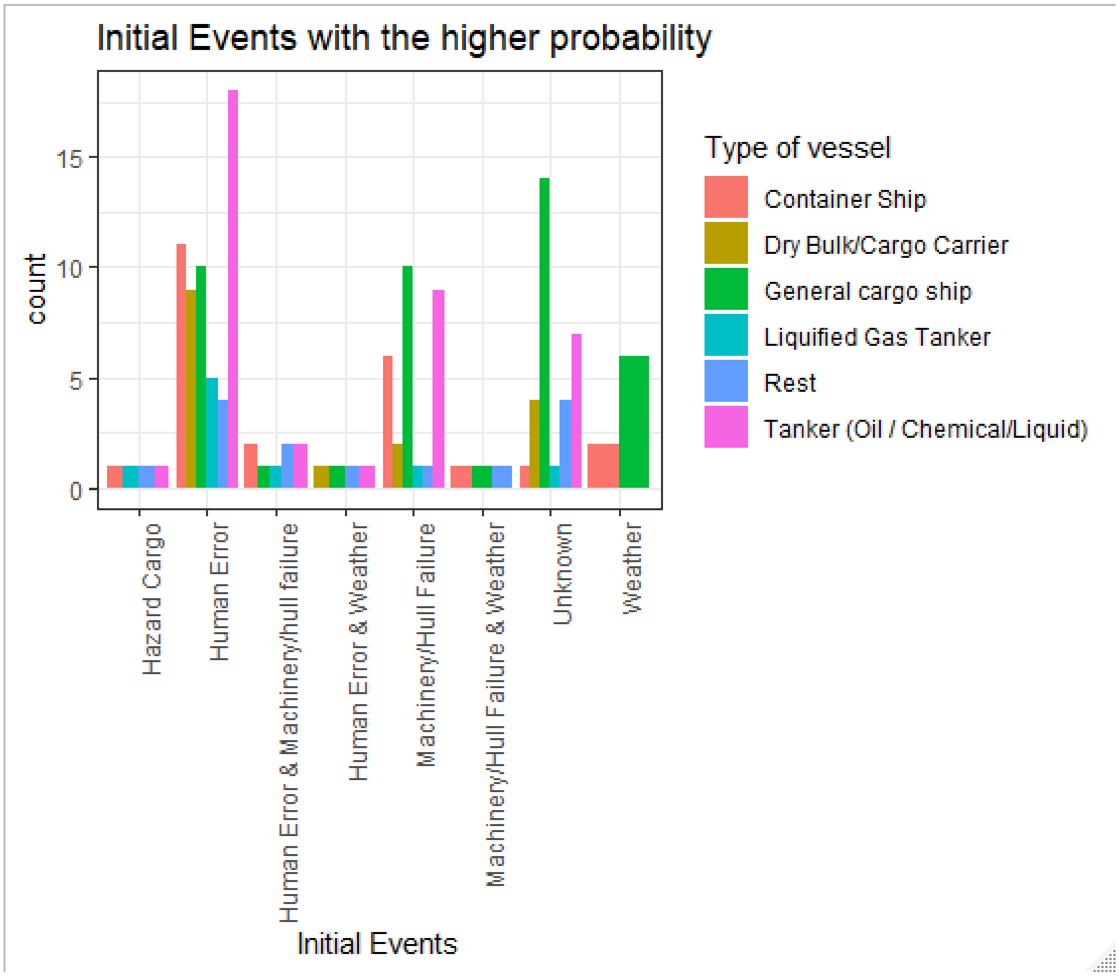

Figure 13 Initial Events with the higher probability in comparison with the "type of Vessel"

By observing the Figure 13 above we can see that there is higher concentration to some type initial events but also at the same time specific types of vessel seem to be more susceptible to some initial events than other. More specifically human error and machinery/hull failure seem to present the higher frequency. In the case of human error tankers are presenting the higher frequency. This is an indication of the complexity of the operations onboard of tanker ship. Second in frequency at the same category is the container ship which also type of vessel with many sophisticated operations. In terms of machinery/hull failure general cargo ship are presenting the higher frequency followed by the tankers. Cargo ships as mentioned before are usually used in tramp trading, therefore the have a lot of port calls and many loading/discharging operations, as a result the machinery and hull of the vessel is stressed more than other type of vessels. Due to that reasons we believe their probability of machinery/hull failure is calculated higher. The second type of vessel in frequency of machinery/hull failures is tankers. This one is also is a type of vessel which is used in tramp shipping.

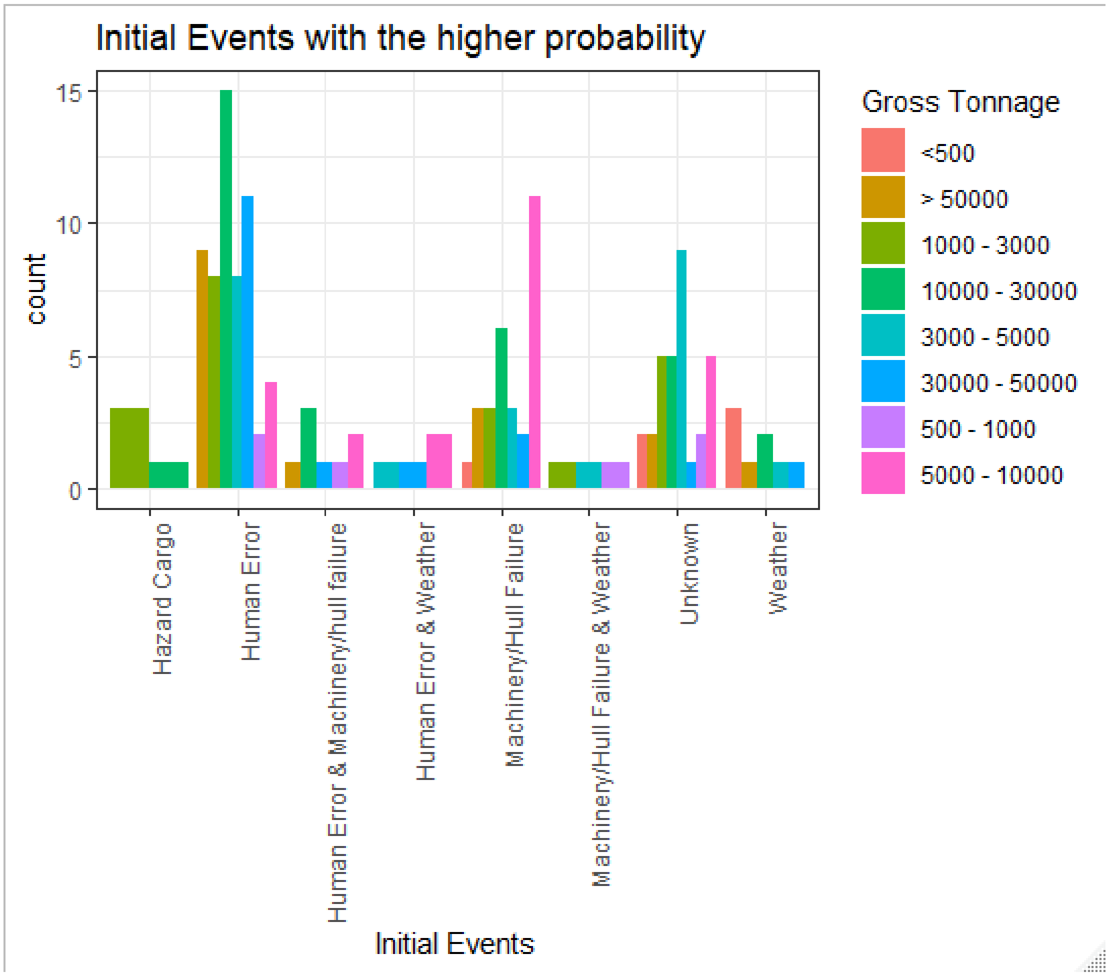

Figure 14 Initial Events with the higher probability in comparison with the "Gross Tonnage"

By studying the initial events based on the size of the vessel we observe that in the case of human error the vessels with gross tonnage 10000-50000 present the higher frequency. These are medium size and large vessels. This is an indication that as the size of the vessel get bigger the probability of human error is also higher. Since the vessels are bigger, they transfer bigger quantity of cargo and heavier therefore the operations on board have bigger duration and the equipment used is more sophisticated. On the other hand, in the case of machinery/hull failures the higher frequency is presented by vessels with gross tonnage 5000-10000, which are mid-sized vessels. The hull and machinery of these vessels seem to be often under stress more than they can afford, thus they present more frequent failures.

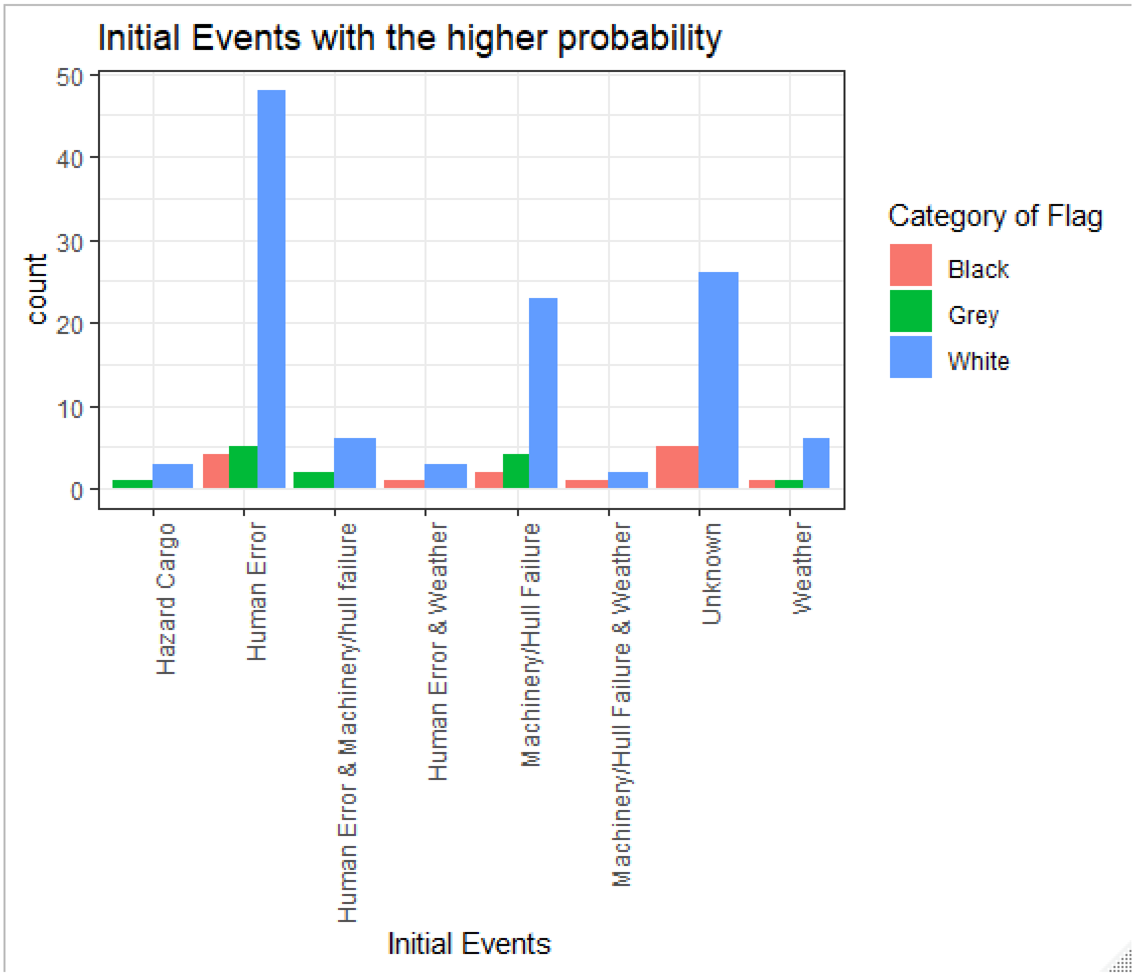

Figure 15 Initial events with the higher probability in comparison with the "Category of the flag"

For all types of initial events the higher frequency is presented by vessels carrying "white" flags. This is a clear indication that the belief that black or grey flags are more susceptible to accidents it is not true. It worth mentioning that there is a high frequency of reports which do not indicate the initial events of the accident. This demonstrates that flags are not obligating the ship management companies to fill in detail the reports of accidents. Therefore, many accidents have occurred under unclear circumstances.

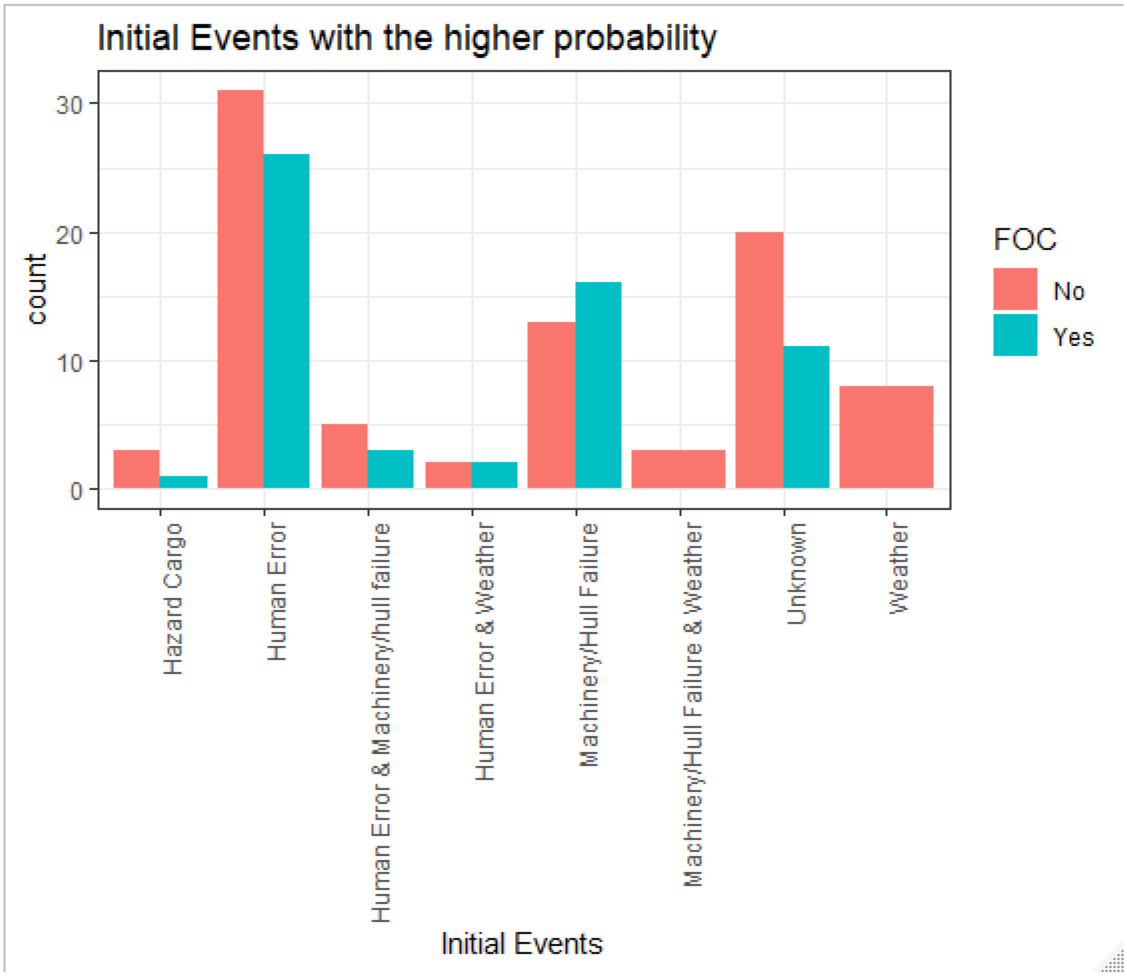

Figure 16 Initial events with the higher probability in comparison the "Flag of Convenience"

By examining the results of Figure 16 we observe that human error is occurred more often to vessel which are not carrying a flag of convenience than to vessels carrying flag of convenience. Even though a crew under a flag of convenience is expected to be less expensive therefore less experienced, the results demonstrate that this fact is not related with the frequency of human errors. On the other hand, Machinery/hull failures are more regularly happening to vessels under a flag of convenience. However usually this may be related to the fact that aged vessels tend to be registered into some specific flags of convenience which allow then to keep trading despite their age. In addition, aged vessels are presenting higher probability to hull and machinery failures due to their aged equipment and materials as it is explained below.

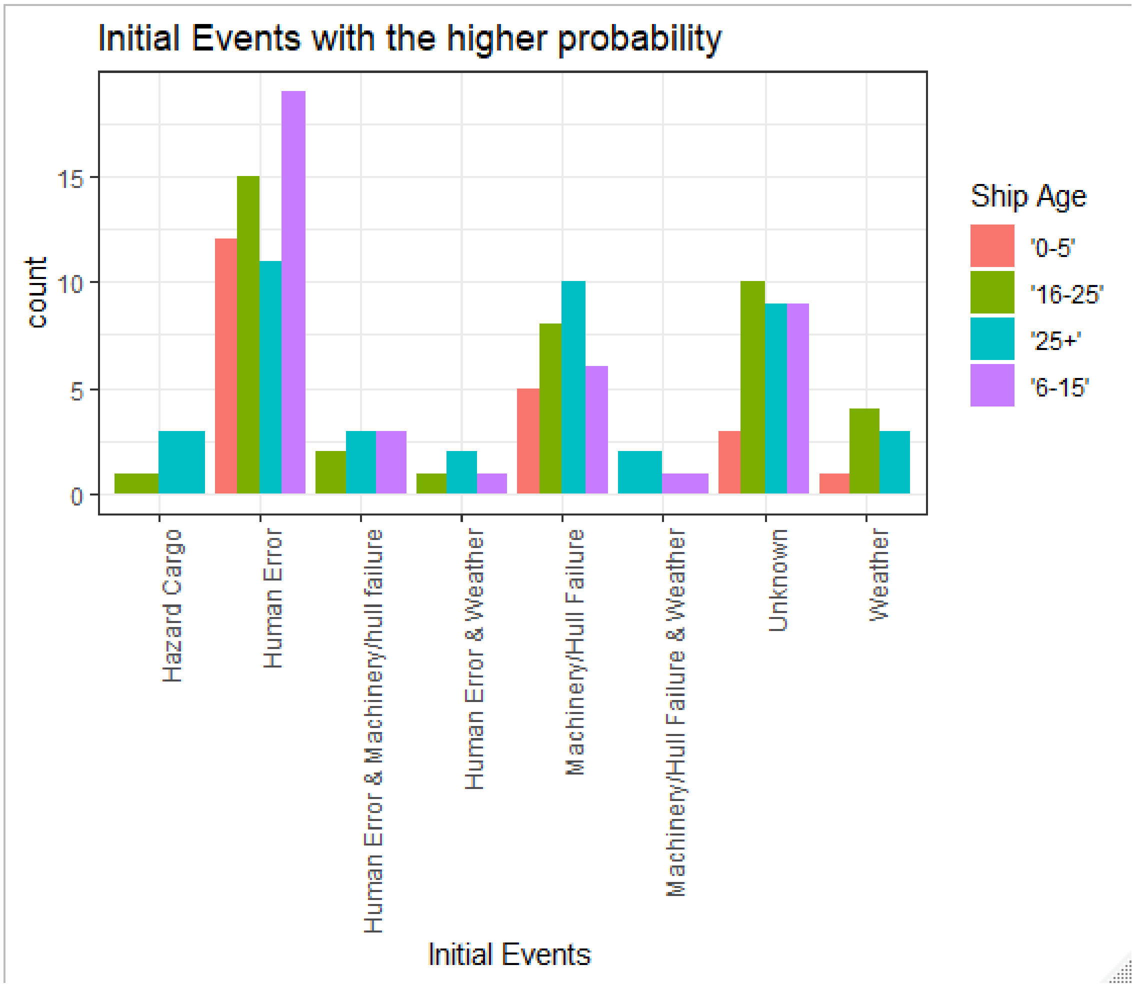

Figure 17 Initial events with the higher probability in comparison with the "Ship Age"

Regarding the initial events and the age of the vessel we observe in Figure 17 that vessels with age between 6 to 15 years old are presenting the higher frequency of human errors while ships above 25 years old demonstrate the smaller frequency. It is known that young crew prefer younger ships because they have advanced IT systems and connection with internet through VSAT technology.(Deloitte, 2011). Therefore, we may assume that older ships have more experienced crew therefore the frequency of human error is diminished. However, in the occasion of machinery/hull failures older ships are presenting higher frequency since the structure and equipment of the vessel is more used and strained compared to the younger ones.

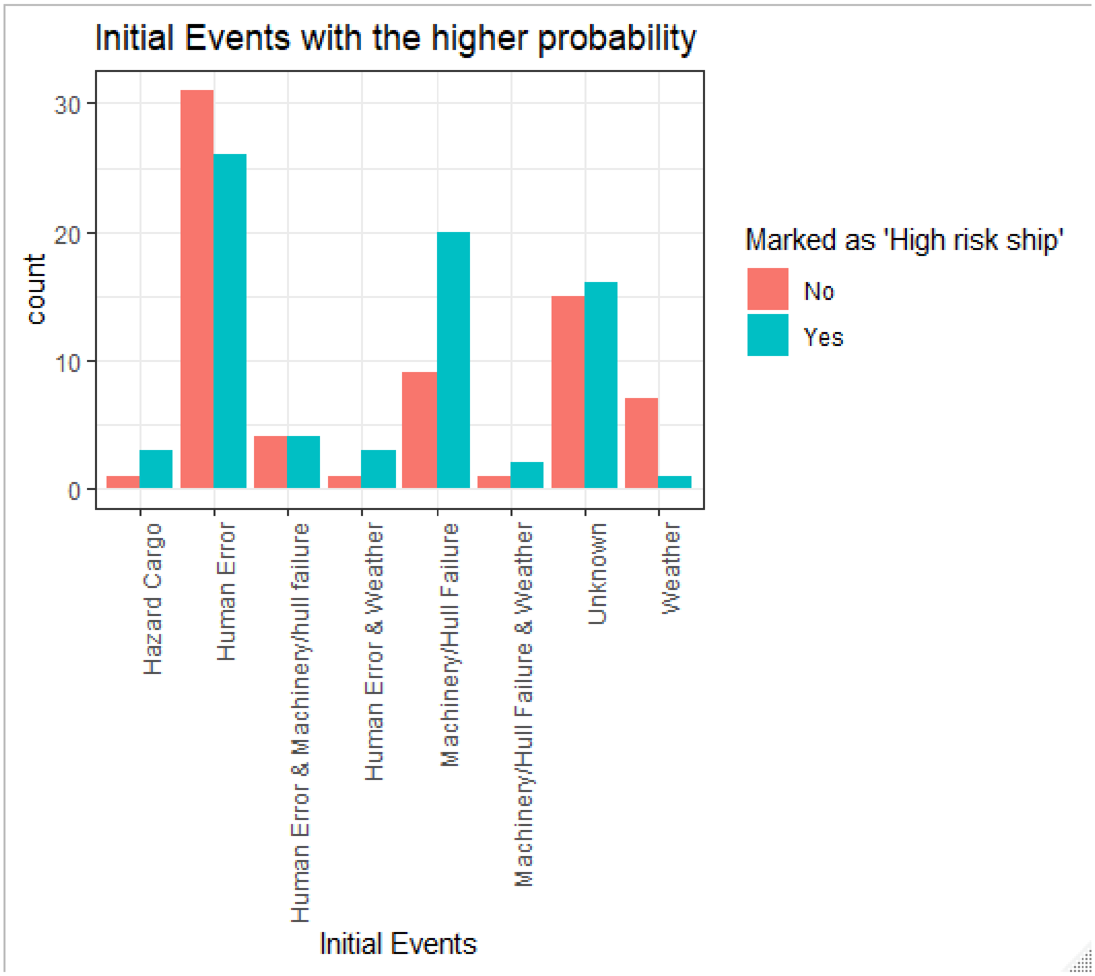

Figure 18 Initial events with the higher probability in comparison with “Performance in Paris/Tokyo MoU”

Regarding the distribution of initial events in relation with the performance of the vessel in Paris/Tokyo MoU we observe in Figure 18 that vessels that perform better are presenting slightly higher frequency of human errors than vessels which do not perform well. This shows that performance of a vessel regarding the Port State regulations are not related with the performance of the crew. On the hand vessels which are marked high-risk or detained demonstrate a much higher frequency of machinery/hull failures. This observation is an indication that vessel with low performance and controversial operation behavior regarding the Port State regulations have higher probability of accidents caused by machinery/hull failures.

Above we examined the relationships between the “parents” nodes of initial event and the node. In terms of arc strength, we present the corresponding scores in the Table 55 below.

Category of Flag	Initial Event	1,07329E+04
Flag of Convenience	Initial Event	8,00014E+03
Gross tonnage	Initial Event	1,39441E+04
Performance in Paris/ Tokyo MoU	Initial Event	7,99681E+03
Age of vessel	Initial Event	1,19634E+04
Type of vessel	Initial Event	1,33127E+04

Table 55 “Parents” nodes of Initial Event and the corresponding “arc strength score”

The score of arc’s strengths shows that the selected nodes are connected with relatively “weak” arcs. However, they are important for our study and the results demonstrate their ability to act as indications of higher probability of accidents.

4.4.17. Node 17: Accident type

Parents: Season, Time, Operational condition, Initial event, Longitude, Latitude.

Children: Operational capability after, Injuries, Fatalities

This node is another “hub” node for our model and is affected by many “parents” nodes while it has causal relationship with the consequences of the accident. The consequences are divided in three categories, the first one is related with the damages on the vessel and the rest two are related with the personnel on board and the extent of the harm the accident caused to them. Consequences have impact not only to the vessel and crew but also to the ship managing company. This impact is not only related with economical losses but with reputation and market share losses. Therefore, they are important for the ship managing company.

Since this node has many parents the number of possible combinations is very large. Therebefore we selected to study 303 observations which present conditional probability above 90%. Based on that Sample we observed the distribution of frequencies explained below.

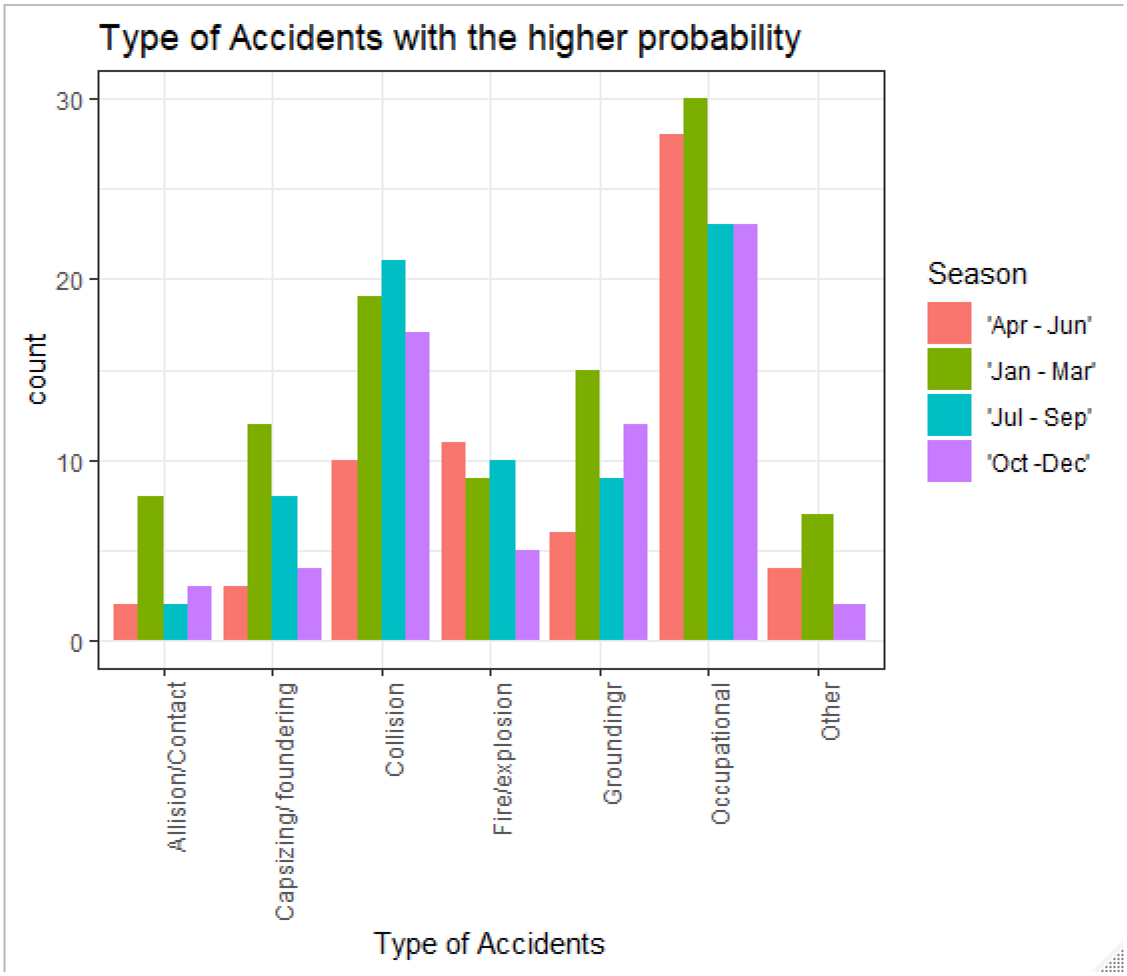

Figure 19 Type of accidents with the higher probability in comparison with "Season"

By examining the Figure 19 above first we observe that higher frequency is presented by occupational accidents and the second higher frequency is observed by collision. Generally, most types of accidents are occurred during the months January, February and March which we believe it is related with the fact that weather conditions and sea condition is rougher and make the operation of a vessel more demanding. However, for some types of accidents the frequency is higher in different season. Specifically, in the case of collisions we observe that the corresponding frequency is higher during summer season (jul-sep). Generally during this period, the temperature is higher, and the crew's fatigue is augmented. These conditions we believe that are causing the higher frequency. Also, in the case of fire/explosions we observe that the frequency is higher during spring and summer where the average temperature is higher.

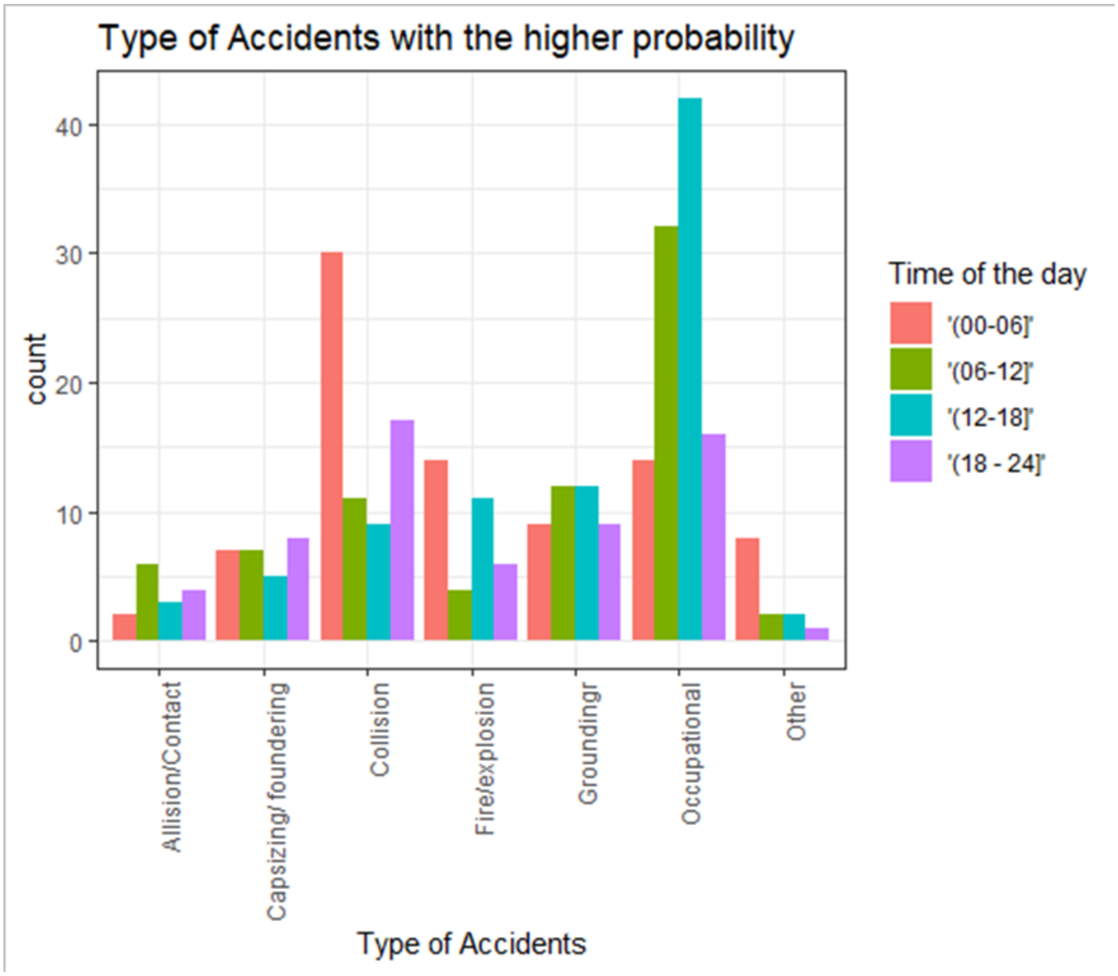

Figure 20 Type of accidents with the higher probability in comparison with the "Time of Day"

Regarding the time of the accident the corresponding frequencies follow different distribution for each type of accident. More specifically during the night (00-06) the visibility is low, the number of crew working is relatively small, and capability of reaction seems to be diminished. These facts in combination with the augmented fatigue and lack of sleep because of night shifts seem to cause the high frequency of Collisions and Capsizing. During the morning (06-12) part of the crew is resting or finishing its night shift therefore the crew alertness is relatively low. Moreover, some of the onboard operations are ongoing. As a result of these conditions we observe during these time period a high frequency of allisions and a relatively high number of occupational accidents. The time slots between 12 and 18 are the part of the day when most operations on board are ongoing. Thus, the frequency of occupation accidents is higher. The last part of the day (18-24) is the time where most of the crew is resting, the daily onboard operations have just ended thus the crew's fatigue is augmented and the level of alertness is low. As a result, we observe during that period relatively high frequency of Allision/contact, Capsizing/foundering and collisions.

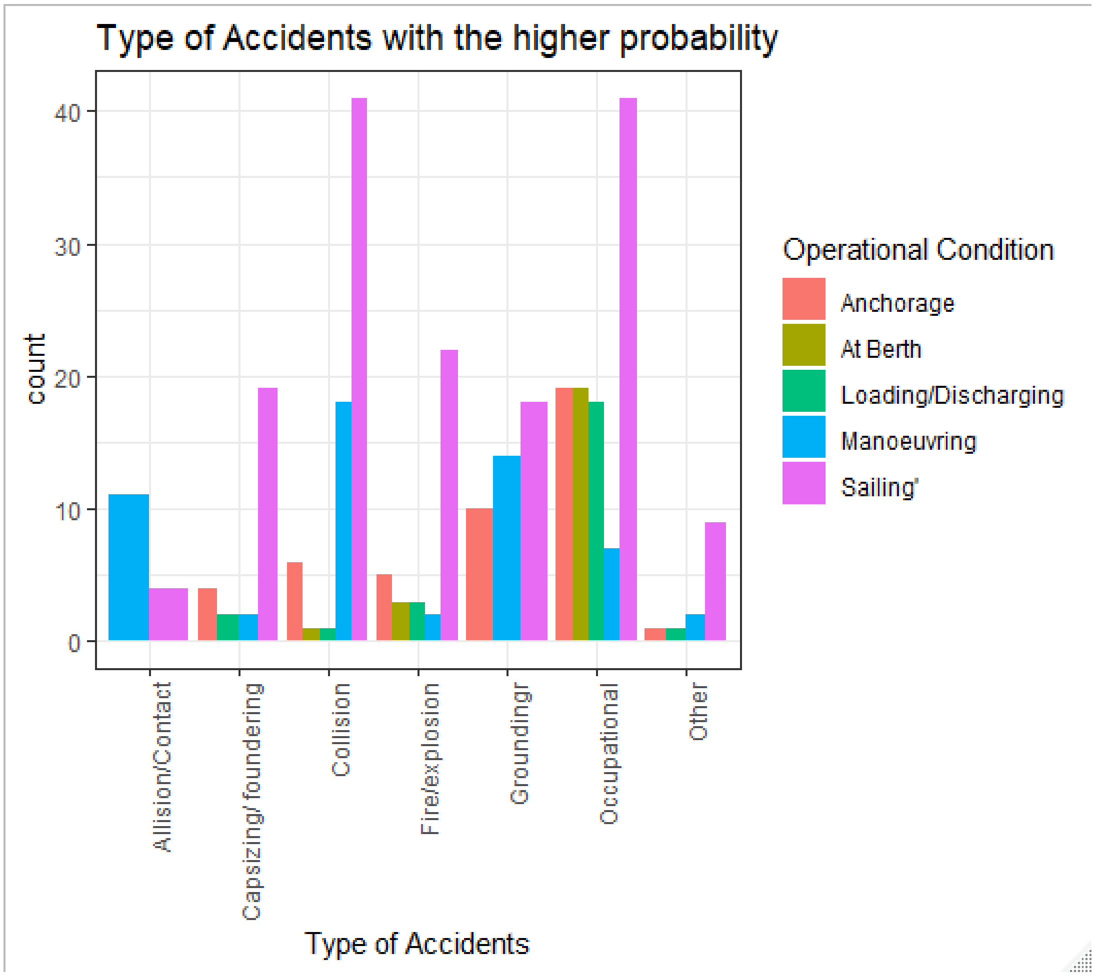

Figure 21 Type of accidents with the higher probability in comparison with the "operational condition"

Regarding the operational condition of the vessel during an accident, Figure 21 indicates that the higher probability is demonstrate while the ship is "sailing" towards her target destination for all types of accidents Allision/contact which present higher frequency during "Maneuvering". This was expecting because a ship is maneuvering in narrow waters or harbor areas where it is surrounded by constructions which she may have contact with. During "sailing" the is traveling with relatively higher speed than in the other conditions and most of the operations are ongoing therefore we may assume that the dangers it is facing the vessel and the crew are more.

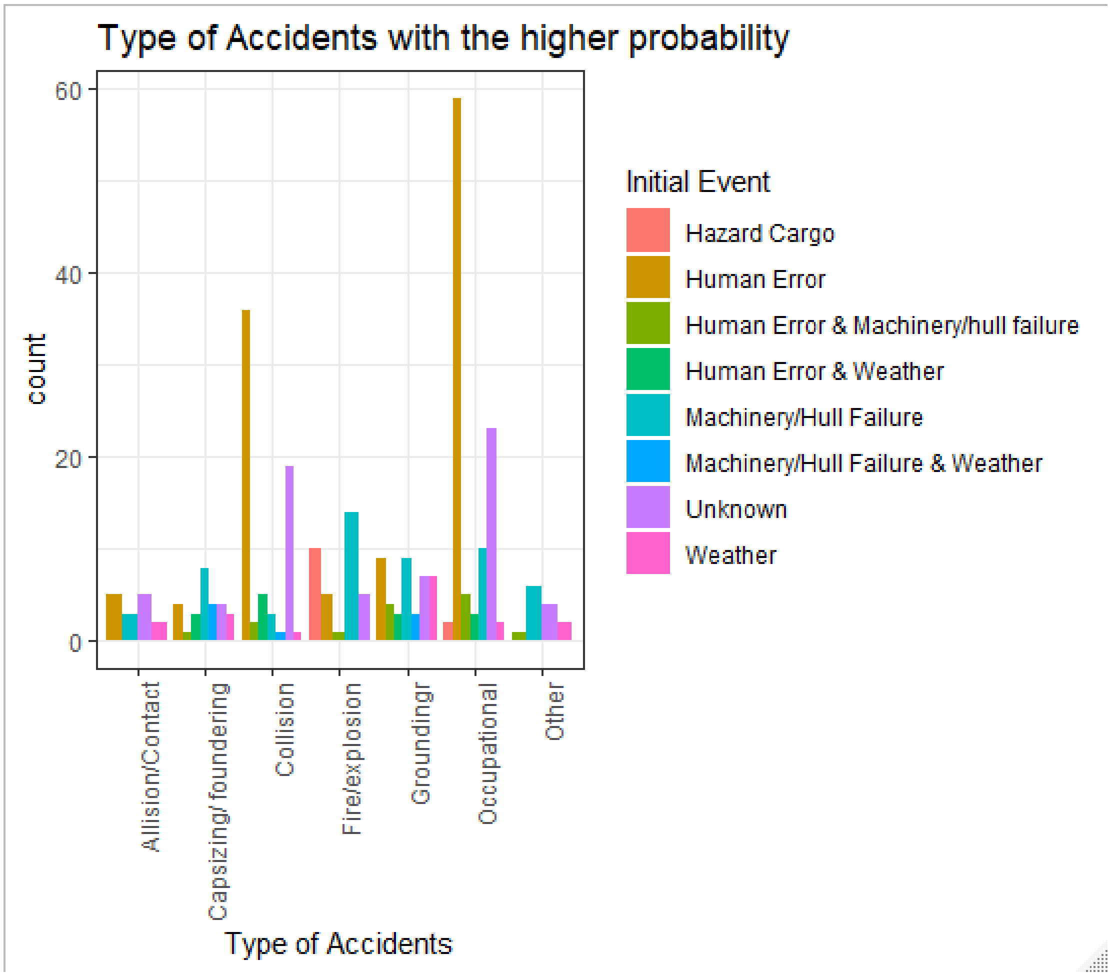

Figure 22 Type of accidents with the higher probability in comparison with the "Initial Event"

As explained before the most common root cause of an accident is the human error however by examining more the root causes of the accident per type of accident, we observe in Figure 22 that in some types the main root cause is machinery/hull failure. These types are the capsizing/foundering and fire/explosion. This indicate that the proper maintenance of a ship is necessary and important in order to avoid failures which may cause a serious accident with severe impact. However, the results highlight the importance of human factor. Especially in the case of collision and occupational accidents the frequency of human errors. This is a concerning finding since it indicates that the crew's training is not sufficient. Most of these accidents are happening due to lack of proper judgment and inadequate safety measures.

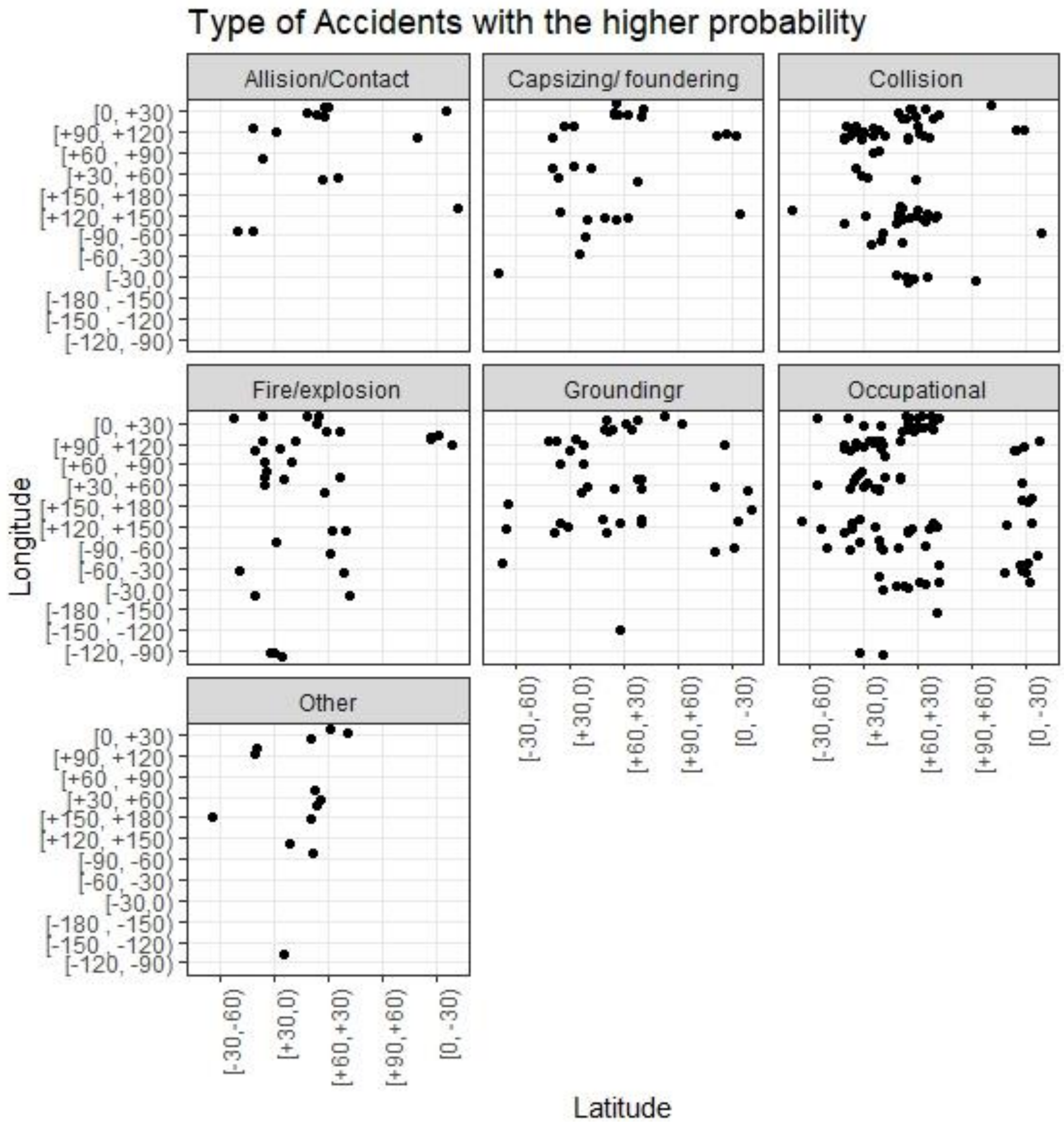

Figure 23 Type of accidents with the higher probability in comparison with the coordinates of the location of the accident

The above set of graphs in Figure 23 is showing which groups of Latitudes and Longitudes are presenting the higher frequency per type of accident. In the case of Allision/contact, "fire/explosion and "other" there is not high concentration in some specific region. In the case of capsizing we observe that in the region enclosed by Longitude = [0,30] and Latitude = [+30,+60] which enclose the North Sea and the Mediterranean Sea. For collision accidents the areas with higher concentration are area1=[longitude=(+90,120),Latitude=(0,+30)], area 2=[longitude =(0,+30),Latitude(+30,+60)], area 3=[Longitude=(+120,+150),Latitude(+30,+60)] which are congested waters. In the category of grounding there a minor concentration in the

area with longitude [0,+30] and latitude (+30,+60) which encloses the English channel. Regarding the occupational accidents the area with longitude (+90,+120) and latitude (0,+30) and the area with Longitude [0,+30] and latitude[+30.+60] are presenting the higher frequency.
Regarding the strength of the arcs, the corresponding score are:

Operation Condition	Type of Accident	2,21135E+05
Initial Event	Type of Accident	2,41833E+05
Latitude	Type of Accident	2,30379E+05
Longitude	Type of Accident	2,53379E+05
Season	Type of Accident	2,07299E+05
Time	Type of Accident	2,07298E+05

Table 56 "Parent" nodes of Type of Accident and the corresponding "arc strength score"

Since the scores are positive with may say that the relationship, they are representing is not strong enough.

4.4.18. Node 18: Capability after the accident

Parents: Type of Accident
Children: none
The capability of the vessel to continue her journey after the accident is influenced by the type of accident.

	Capable to continue the journey?	
Type of Accident	Yes	No
Occupational	65,20%	34,80%
Grounding	70,77%	29,23%
Collision	49,12%	50,88%
Fire/explosion	72,42%	27,58%
Allision/Contact	63,06%	36,94%
Capsizing/ foundering	53,32%	46,68%
Other	68,58%	31,42%

Table 57 Conditional probabilities of "Capability after the accident" in comparison with the "Type of Accident"

In term of severity of the damage's collision seems to bear the more risk since the probability of a vessel to not be able to continue after the incident is 51%
Since the corresponding score -4,98561E-01 we may say that this arc is relatively strong.

4.4.19. Node 19: Injuries

Parents: type of Accident
Children: None

Another type of consequence which is affected by the type of the accidents is the number of injuries.

	Injuries		
Type of Accident	1-4 people injured	>4 people injured	No injuries
Occupational	14,81%	0,04%	85,15%
Grounding	0,10%	0,10%	99,80%
Collision	2,67%	1,79%	95,54%
Fire/explosion	22,54%	5,10%	72,36%
Allision/Contact	5,47%	0,25%	94,28%
Capsizing/ foundering	6,79%	0,16%	93,05%
Other	6,49%	0,29%	93,22%

Table 58 Conditional probabilities of "Injuries" in comparison with the "Type of Accident"

The corresponding probabilities indicate that the accident with the higher probability of injuries is the Fire/explosions (27,64%) followed by the occupational (15%). The score of this arc is -6,38609E+00 which suggests that this arc also is relatively strong.

4.4.20. <u>Node 20: Fatalities</u>

Parents: Type of Accident

Children: Severity of Casualty

Finally, the last type of consequence which is affected by the type of accident is the fatalities.

	Fatalities		
Type of Accidents	"1-5"	>5	No fatalities
Occupational	94,71%	0,04%	5,25%
Grounding	0,10%	0,10%	99,80%
Collision	14,94%	5,30%	79,77%
Fire/explosion	32,50%	0,12%	67,38%
Allision/Contact	15,92%	0,25%	83,83%
Capsizing/ foundering	16,75%	13,43%	69,83%
Other	31,27%	0,29%	68,44%

Table 59 Conditional Probabilities of "Fatalities in comparison with the "Type of Accident"

The accident that bear the higher risk of fatalities is the Occupational Accident. This illustrates that the operations on board are quite dangerous if not adequate safety measures are not taken. Moreover, despite the technological advance of the onboard mechanisms, accident such as fire/explosion and capsizing bear a significant probability of loss of life (30%). Regarding the strength of the arc the score is -1,23858E+02 which suggests that this arc also relatively strong.

4.4.21. Node 21: Severity of Casualties

This node is expressing the severity of casualties as they are characterized by IMO and our algorithm showed that this categorization is affected by the number of fatalities however the severity of casualties is a term which may interpreted differently by personal criteria. Therefore, we decided to not examine the conditional probabilities since they are related with the number of fatalities. The bigger the number is the higher is characterized the severity.

4.5. *Sensitivity analysis through scenarios*

4.5.1. Introduction

Some variables cannot be affected by the ship owner therefore in this section we will focus on the variables which are affected by the managing strategy of shipowner or a ship management company. These variables are flag of convenience and the performance according to Paris/Tokyo MoU. The type of the vessel and the size can me influenced by many factors which are not study in this thesis such as the freight rates, the access to capital, the preferences of the ship owner. Thus, we will exclude these variables from our sensitivity analysis. Moreover, we are assuming that the shipping company had selected a white flag which allows the vessel to approach every port. Based on that conditions we will create 2 scenarios. The first scenario is the worst-case scenario where the vessel suffers serious damages and there are fatalities (1-5). The second scenario will be the less bad scenario where there will be severe damages to the vessel but no fatalities. For each scenario there will be four strategies which will be the all the possible combinations for the two aforementioned variables.

Another assumption for our scenarios will the frequency of accidents. We are assuming that the upcoming year there will be 128 accidents (the average of the last three years). According to UNCTAD the total fleet of merchant ships for 2018 were 1924. If we assume that the fleet increase next year by 5% then the probability of accident will be equal to 128/2021 =6,33%

As mentioned to Chapter 2 the Risk is the probability of the accident multiplied with the probability of consequence. We will calculate separately the probability for each type of accident and the sum of it will be multiplied with 6,33% in order to calculate the risk that each strategy carries.

4.5.2. Scenario 1: Worst case consequences

Severe damages or total loss, not able to proceed
There are fatalities

Flag of convenience	yes	yes	no	no
Listed as High risked	yes	no	yes	no
Occupational	31,65%	33,00%	32,00%	32,00%
Grounding	0,00%	0,00%	0,00%	0,00%
Collision	7,90%	7,40%	8,30%	7,80%
Fire/explosion	8,20%	7,90%	8,50%	9,00%
Allision/Contact	6,80%	4,30%	7,90%	5,90%
Capsizing/ foundering	6,08%	6,25%	9,00%	8,00%

Other	9,55%	9,30%	10,00%	0,00%
Sum	70,18%	68,15%	75,70%	62,70%
Probability of Accident	6,33%	6,33%	6,33%	6,33%
Risk	4,44%	4,32%	4,79%	3,97%

Table 60 Total risk of each operational strategy for "worst case" scenario

The strategy that bears the least risk is the fourth strategy where the vessel is not carrying a flag of convenience and is not marked as highly risk ship. The variable "performance in Paris/Tokyo MoU" seems to have higher influence. As proved in the previous section the performance of a vessel in the selected authorities is an indication of how well it is maintained and is following the regulations. A vessel with good performance and history is bearing less risk not only for her prosperity but also for the health of the crew.

4.5.3. Scenario 2: Less worse case consequences

Severe damages or total loss of the ship, not able to proceed

No fatalities

Flag of convenience	yes	yes	no	no
Listed as High risked	yes	no	yes	no
Occupational	1,90%	1,60%	2,05%	1,60%
Grounding	30,00%	29,34%	28,50%	2,99%
Collision	39,00%	42,27%	43,23%	41,83%
Fire/explosion	17,98%	18,75%	19,65%	18,34%
Allision/Contact	30,00%	32,00%	32,77%	30,73%
Capsizing/ foundering	30,55%	32,27%	31,46%	34,40%
Other	21,00%	20,97%	24,04%	22,22%
Sum	170,43%	177,20%	181,70%	152,11%
Probability of Accident	6,33%	6,33%	6,33%	6,33%
Risk	10,79%	11,22%	11,51%	9,63%

Table 61 Total risk of each operational strategy "less worst-case" scenario

As the previous Scenario the strategy that seem to bear the less risk the fourth one. Which indicates that the existence of flag of convenience is affecting more the probability of casualties related to damages on the ship.

Through this sensitivity analysis we concluded to two main findings. First, all ships under no FOC and with good performance in the inspections are bearing less risk. Regarding the risk related to the crew, the performance of the vessel in the inspection seems to be a better indicator. However, in terms of risk of damages in the vessel the existence of flag of convenience is a more suitable indicator. We believe that these findings could be a useful decision-making tool for the interested parties and the corresponding variables can act as indicators of risk.

4.6. *Conclusion & main findings*

The first part of our empirical analysis was describing the summary of the data through frequency tables. In some variables there was clear indication that they follow a pattern, however we will focus on the results of the rest analysis. In the second part of our analysis we tested our data without utilizing the theoretical framework and we concluded that the data are thin to describe the casual relationship with the proper direction. As a result, we constructed a second BBN based on the theoretical framework we explained in chapter 2. In the third part of our analysis we calculated the conditional probabilities for each state of the variables, and we examined the established relationships with the “parents” nodes. In the last part of our analysis we conducted a sensitivity analysis based on a group of selected variables which we believe demonstrate the operation strategy and performance of a ship management company.

The main findings of our empirical analysis were:

Regarding the hypothesis we made in Chapter 2 that each variable is influencing the probability of accident directly or indirectly, we conclude that we accept it for all the variables except of the Number of crew, If the vessel changed Flag , and if the vessel changed name before the accident. Based on our BBN these variables are not affecting the dependent variables. However, we may conclude that aged vessels tend to change their name and their flag.

Moreover, based on the assigned strength score on each arc which is summarized in Table 62 below. We observe that there are relationships which are negative therefore the corresponding arcs are important for the network and they describe a relative strong relationship. However, there are also positive scores which indicate that these relationships based on our sample are relatively less strong, the positive score can be also a result of absence of other unknown variables which are not selected in this study, but they are affecting the network.

	Parent Node	Child Node	Strength score
1	Location of the accident	Operational condition	-1,70232E+02
2	Location of the accident	Fatalities	-1,23858E+02
3	Type of vessel	Gross tonnage	-1,20206E+02
4	Longitude	Latitude	-8,90770E+01
5	Gross Tonnage	Size of Crew	-8,44942E+01
6	Fatalities	Severity of casualties	-7,38031E+01
7	Gross tonnage	Vessel’s Age	-5,71488E+01
8	Change of flag	Change of name	-4,76730E+01
9	Vessel’s Age	Change of flag	-3,21961E+01
10	Vessel’s Age	Category of flag	-1,48796E+01
11	Gross tonnage	FOC	-1,14811E+01

12	Category of flag	Performance in Paris/Tokyo MoU	-9,80015E+00
13	Type of Accident	Injuries	-6,38609E+00
14	Vessel's Age	Change of name	-4,69403E+00
15	FOC	Performance in Paris/Tokyo MoU	-3,69305E+00
16	FOC	Category of flag	-2,24920E+00
17	Performance in Paris/Tokyo MoU	Change of flag	-1,02189E+00
18	Type of Accident	Operational capability after the accident	-4,98561E-01
19	Performance in Paris/Tokyo MoU	Initial Event	7,99681E+03
20	FOC	Initial Event	8,00014E+03
21	Category of flag	Initial Event	1,07329E+04
22	Vessel's Age	Initial Event	1,19634E+04
23	Type of vessel	Initial Event	1,33127E+04
24	GT	Initial Event	1,39441E+04
25	Time of the day	Type of accident	2,07298E+05
26	Season	Type of accident	2,07299E+05
27	Operational condition	Type of accident	2,21135E+05
28	Latitude	Type of accident	2,30379E+05
29	Initial Event	Type of accident	2,41833E+05
30	Longitude	Type of accident	2,53379E+05

Table 62 Summary of arcs' strength score of the examined BBN

The types of vessels which present the higher probability of accident in relation with their size are:

- General cargo, refrigerator ships, RO-RO with gross tonnage less than 3000
- Liquified Gas tankers, Dry bulk Carriers, Tankers, container ship with more than 10000 grt

The age is also a factor. More specifically Aged vessels with small size and young vessels with big size are demonstrating the higher probability of accident.

Another important finding is that vessels carrying a flag of convenience tend to perform worse during inspection, and they present higher frequency of machinery/hull failures therefore the flag of convenience is related with the level of maintenance of a vessel and is a strong indicator. Another indicator of the level of maintenance the performance of the vessel during inspections.

When it comes to the initial event of an accident, human error seems to be the main root cause especially for tankers. Though in the case of general cargo ship machinery/hull failures shows the higher frequency. In terms of age, old ships tend to present a higher frequency of machinery failures too. Concerning the type of accident for fire/explosion, machinery failures are the most common main reason.

Furthermore, there is a season and time pattern when it comes to the type of accident. Collisions are more frequent during Jul-Sep, fire/explosion are more often during summer months and for the rest types the frequency is higher during winter. During night the probability of most type of accidents is higher except the occupational which are more frequent between 12-18 hour.

The condition of vessel with the higher estimated probability is the "Sailing" while the location with the higher probability is "coastal waters".In terms of consequences we find that the two more risky accidents for the human life is the occupational and fire/explosion while for the ship is collision and fire/explosion.

Finally, a general remark of our results is that there is a relatively large number of reports which does not elaborate on the causes or the circumstances of the accident. This suggests that there is asymmetry of information and the element of "unknown" in the reports is relative strong. As we will explain in the Section 5.4. below this may act as a limitation of the study.

5. Conclusion

5.1. *Introduction*

This is the last chapter of our thesis where we will first reply to the main research question. Next we will compare our main findings with the results of the presented in the section 2.8. After this discussion we will describe the limitations of this study and we will make some suggestions for future research.

5.2. *Answer to Research question*

The main research question of this study was:

Does the vessel's characteristics and operational profile relate with the probability of participating in an accident and is there a modelling feasibility of predicting this probability?

After the conduction of this study we believe that some of vessel's qualitative and quantitative characteristics as well as their operational profile/history are related with the probability of participating in an accident as well as the related consequences. Therefore, it is feasible to assess conditions under which this probability is higher, but we cannot predict with certainty when a maritime accident will happen.

In order to conclude to this answer, first we did an extended literature research. Through this literature research we answered the first sub-question regarding what risk is, how it is defined and how it can be measured quantitatively. Second, we conducted a literature review to prior studies which help us identify our methodology and the preferred variables. By constructing a Bayesian network based on our data we were able to test our theoretical framework and answer the second sub-question regarding the main root causes of a maritime accident. Also, through the Bayesian inference we were able to answer the main research question regarding the relationship of the independent variables with the probability of accident. Finally, through our sensitivity analysis via scenarios we tried to validate the framework and answer the third sub question about which of the selected operation strategies bears less risk for the vessel and the crew.

Particularly, we proved that characteristics such as the type, size and age of the vessel affect her risk profile. Moreover, qualitative features related with her operation profile such as the type of flag she is carrying and her history of performance during inspections from authorities are also related with the investigated probability and can acts as indicators of risk. However, accidents appear to be the outcome of high complex factors and coincidence which are rarely foreseen by the people involved. The large number of factors and coincidences are creating an uncertainty regarding the prediction of an accident. On other hand the selected method of this study proved that we can estimate conditions where the combination of these factors is producing higher probability of accident. Ship management companies can adjust their safety procedures or crew's training based on our findings and we proved that the method of BBN can act as a decision-making tool for interested parties regarding the risk profile of a vessel based on public data.

5.3. *Discussion*

In this section we will compare our key findings with prior studies. Particularly we proved that there is not only one type of vessel which is more prone to accidents as Silveira et al and Yang et all suggested in their studies. We proved that by analyzing the type of vessel in conjunction with the size, General cargo ships and RO-RO ships with small size are prone to accidents while Tankers, Container ships, Dry bulks or Liquefied gas tankers with relatively big size are presenting a high possibility of accident. Furthermore, we rejected the assumption that there is relationship between the size of crew, the fact that a ship has changed flag and/or name with the root causes of an accident and the probability of accident. This is in contrast with the statement of Mr Ami Daniel that “vessel who change flag and/or name are more prone to accidents”(Corbett, 2018). Regarding the age of the vessel we proved that vessels with small aged ships and big young vessel are prone to accidents. This finding is partially in align with the findings of Chalikias et all who proved that old vessels and with big size are prone to accidents.

In terms with the root causes we find human error is the main root cause for most accidents as was states also in the study of Ozdemir and Guneroglu. Thus, in the case of General cargo ships and for the case of fire/explosion Machinery/hull failures are presenting the higher probability. Hanninen and Kujala quoted that the probability of collision is higher during conditions with low visibility or when the crew is tired. We partially confirmed it, since we proved that most of the maritime accidents occurred during the night where the visibility is low however the majority of occupational is during the day where most of the onboard operations are ongoing. In terms of season patterns Yang et all cited that during spring collision are more frequent, we demonstrated that during summer months collisions are presenting higher probability and so are fire/explosions. Partially we had the same findings with Pagiaziti. More specifically we showed that coastal waters are presenting the higher probability of accident however we also presented that collisions and fire/explosion are producing the most severe consequences and not the groundings.

5.4. *Limitations*

The results of this study quite profoundly suggest that the selected variables are influencing either the initial event or the type of accident, therefore the probability of accident. However, these results are based on some assumptions which may affect our results. As we explained in chapter 4 the positive score of some arc strength test indicates that there are unknown variables which are not studied in this research but they affect the network. Second, our data were based mainly on the maritime reports of IMO, a dataset combined from different data sources may demonstrate different findings. According to an article of Mrs. Bakhsh the reports of serious maritime accidents in the last four years are not yet published by the flag states or they are not accessible by the public. There is the possibility that these specific detailed reports from the flag states may alter the initial findings.(Nidaa, 2019)

As we saw in many of our findings the circumstances under which the accident was caused were “unknown”. This is an indication that the reports are not filled in detail. Another observation based on our results is that none of Maritime incident were reported in IMO especially in the case of occupational accidents. This is an indication that injuries and less minor accidents may have not be reported, giving perhaps a skewed dataset.

5.5. *Further Future Analysis*

This thesis has encompassed some conditions and factors which increase the probability of accident. however, the complexity and diversity of shipping industry calls for a more integrated research that it will take into consideration the limitation we discussed in section 5.4. More a new type of risk is emerging as the technology of communications on vessels is developing. This is the risk of Cyber-attacks whose impact may affect the shipping industry as well as the maritime accidents. The London P&I club is alerting the shipping industry that the industry is vulnerable to cyber risks. The risk is higher for newbuilding's which contain systems more vulnerable to cyber-attacks.(Forgey, 2016)

Bibliography

AIMS. (2015). What is a Marine Surveyor. Retrieved June 2, 2019, from https://www.amsa.gov.au/vessels-operators/incident-reporting/what-marine-incident

Allianz Global. (2018). ALLIANZ GLOBAL CORPORATE & SPECIALTY SAFETY AND SHIPPING REVIEW 2018 An annual review of trends and developments in shipping losses and safety. Retrieved from www.agcs.allianz.com/

Ancuţa, C., Stanca, C., Andrei, C., & Acomi, N. (2017). Behavior analysis of container ship in maritime accident in order to redefine the operating criteria. IOP Conference Series: Materials Science and Engineering, 227(1). https://doi.org/10.1088/1757-899X/227/1/012004

Asariotis, R., Assaf, M., Benamara, H., Hoffmann, J., Premti, A., Rodríguez, L., … Visser, D. (2018). Review of Maritime Transport 2018. In UNCTAD 2018 (p. 49). Retrieved from https://unctad.org/en/PublicationsLibrary/rmt2018_en.pdf

Aven, T. (2010). On how to define, understand and describe risk. Reliability Engineering and System Safety, 95(6), 623-631.

Aven, T. (2011). On Some Recent Definitions and Analysis Frameworks for Risk, Vulnerability, and Resilience. Risk Analysis, 31(4), 515–522. https://doi.org/10.1111/j.1539-6924.2010.01528.x

Aziz, A., Ahmed, S., Khan, F., Stack, C., & Lind, A. (2019). Operational risk assessment model for marine vessels. Reliability Engineering and System Safety, 185(January), 348–361. https://doi.org/10.1016/j.ress.2019.01.002

Baumberger, C. (2014). Art and Understanding. In Defence of Aesthetic Cognitivism. Bilder Sehen. Perspektiven Der Bildwissenschaft, (March), 1–24. Retrieved from https://s3.amazonaws.com/academia.edu.documents/30333716/art_and_understanding_homepage.pdf?response-content-disposition=inline%3B filename%3DArt_and_Understanding_In_Defence_of_Aest.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIWOWYYGZ2Y53UL

Bertrand, P. V., Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1988). Akaike Information Criterion Statistics. Journal of the Royal Statistical Society. Series A (Statistics in Society), 151(3), 567. https://doi.org/10.2307/2983028

Bitner-gregersen, E., & Monbaliu, J. (2006). Towards the identification of warning criteria : Analysis of a ship accident database, 27(2005), 281–291. https://doi.org/10.1016/j.apor.2006.03.003

Board, S., & Accident, M. (2017). Chapter 5 Marine accident and incident investigations Japan Transport Safety Board Annual Report 2017 65, (c), 65–82. Retrieved from http://www.mlit.go.jp/jtsb/annualreport2017e/E09chapter5.pdf

Caridis, P. (1999). Casualty Analysis Methodology for Maritime Operations. Final Report of the European Research Project CASMET.

Chalikias, M., Ntanos, S., & Milioris, K. (2015). Data analysis on maritime accidents over 1000 grt : The case of Greece Data analysis on maritime accidents over 1000 grt : The case of Greece. In 4th International Conference in Quantitative and Qualitative Methodologies in the Economic and Administrative Sciences.

Chatzaras, A. (2018). Statistical Analysis and Modeling of the Baltic Dry Index (BDI) by. Erasmus University.

Committee, S., & Senate, T. (2009). Chapter 1 Introduction and conduct of inquiry, (March), 18–21.

Corbett, A. (2018). Could data have predicted the Kerch Strait tanker tragedy? | TradeWinds. Retrieved August 26, 2019, from https://www.tradewindsnews.com/insurance/could-data-have-predicted-the-kerch-strait-tanker-tragedy-/2-1-529330

D.D., H., B.D, P., R.G., B., K.H., R., & R.B., W. (1998). Safety Management Assessment System (SMAS): a process for identifying and evaluating human and organizational factors in the marine systems operations with field test results.

David, O. (2019). Bigger ships , bigger losses. Lloyd's LIst, 44(0), 65–67.

Deloitte. (2011). Challenge to the industry - Securing skilled crews in today's marketplace, 16. Retrieved from http://www.deloitte.com/assets/Dcom-Greece/Local Assets/Documents/Attachments/Press/Crewing Survey-Full Report.pdf

Det Norske Veritas. (2001). Marine risk assessment. London.

DNV GL. (2016). Efficient updating of risk assessments.

Dutch Safety Board. (2009). Shipping Occurrences Report Lessons learned and priorities.

EMSA. (2015). Annual Overview of Marine Casualties and Incidents 2015.

EMSA. (2017). Equasis Statistics from 2005 to 2017. The World Merchant Fleet, 101. Retrieved from http://www.emsa.europa.eu/equasis-statistics/download/5046/472/23.html

Esbensen, P. J. (1985). The importance of crew training and standard operating procedures in commercial vessel.

Faghih-Roohi, Shahrzad, Xie, M., & Ng, K. M. (2014). Accident risk assessment in marine transportation via Markov modelling and Markov Chain Monte Carlo simulation. Ocean Engineering, 91, 363–370. https://doi.org/10.1016/j.oceaneng.2014.09.029

Faturachman D, M. S. (2012). Transportation Accident Analysis in Indonesia. . The 2012 International Conference on Asia Pacific Busines Innovation and Technology Management.

Faulkner, D. (2013). Shipping safety. Ingenia.

Fernandes, R., Braunschweig, F., Lourenço, F., & Neves, R. (2016). Combining operational models and data into a dynamic vessel risk assessment tool for coastal regions. Ocean Science, 12(1), 285–317. https://doi.org/10.5194/os-12-285-2016

Forgey, B. (2016). The fear factor. Lloyd's LIst, 69(4), 4–10.

Graziano, A. (2018). The Maritime Commons : Digital Repository of the World Flag state performance and the implementation of port state control in the European Union - A mixed methods approach.

Gurning, R. O. S. (2011). Maritime Disrutions in the Australian Indonesian Wheat Supply Chain; an Analysis of Risk Assessment and Mitigation Strategies, 363. Retrieved from https://eprints.utas.edu.au/12929/2/Whole_thesis_excluding_Appendix_I_Gurning_thesis.pdf

Haimes, Y. Y. (2009). On the complex definition of risk: A systems-based approach. Risk Analysis, 29(12), 1647–1654. https://doi.org/10.1111/j.1539-6924.2009.01310.x

Hamad, H. B. (2016). Flag of Convenience Practice: A Threat to Maritime Safety and Security. IJRDO-Journal of Social Science and Humanities Research, (September).

Hänninen, M., & Kujala, P. (2009). The effects of causation probability on the ship collision statistics in the Gulf of Finland. Marine Navigation and Safety of Sea Transportation, 4, 267–272.

Hassel, M., & Lars Petter, H. (2009). Marine Databases - A comparative study, from a risk management perspective. Norwegian University of Science and Technology Department of Marine Technology.

Heckerman, D., Geiger, D., & Chickering, D. M. (1993). Learning Bayesian Networks : The Combination of Knowledge and Statistical Data Metrics for Belief Networks : Machine Learning, 20(3), 197–243. https://doi.org/10.1023/A:1022623210503

ICS | Shipping and World Trade. (n.d.). Retrieved May 2, 2019, from http://www.ics-shipping.org/shipping-facts/shipping-and-world-trade

IMO. (2002). GUIDELINES FOR FORMAL SAFETY ASSESSMENT (FSA) FOR USE IN THE IMO RULE-MAKING PROCESS.

IMO. (2008). MSC-MEPC.3/Circ.3. CASUALTY-RELATED MATTERS* REPORTS ON MARINE CASUALTIES AND INCIDENTS.

International Transport Workers' Federation. (2019). Flags of convenience | ITF Global. Retrieved July 22, 2019, from https://www.itfglobal.org/en/sector/seafarers/flags-of-convenience

ISO 31000:2018 - Risk management -- Guidelines. (n.d.). Retrieved July 17, 2019, from https://www.iso.org/standard/65694.html

Isotalo, J. (2006). Basics of Statistics. The American Statistician, 48(1), 59. https://doi.org/10.2307/2685091

Kaplan, S. (1997). The Words of Risk. Risk Analysis, 17(4), 407–417. https://doi.org/10.1111/j.1539-6924.1997.tb00881.x

Kaplan, S; Garrick, J. (1981). On The Quantitative Definition of Risk Stanley Kaplan' and B. John Garrick2 Received July. The Annals of Occupational Hygiene, 24(I), 245–248.

Keller., G. (2014). Statistics for Management and Economics. Gengage Learning.

Kontkanen P., M. P. (1997). A Bayessian approach of discretization. European symposioum on Inteligence techniques, (pp. 265-268). Bari,italy.

Ladan, M., & Hänninen, M. (2012). Data Sources for Quantitative MarineTraffic Accident Modeling. Aalto University, Department of Applied Mechanic (Vol. SCIENCE +).

Liuzzo, G., Bentley, S., Giacometti, F., Bonfante, E., & Serraino, A. (2014). The term risk: etymology, legal definition and various traits. Italian Journal of Food Safety, 3(1). https://doi.org/10.4081/ijfs.2014.2269

Lützhöft M, G. M. (2011). Information Environment, Fatigue and Culture in the Maritime Domain. Reviews of Human Factors and Ergonomics.

Maritime, E., & Agency, S. (2018). Annual Overview of Marine Casualties and Incidents 2018. European Maritime Safety Agency, 1–135. Retrieved from http://emsa.europa.eu/emsa-documents/latest/item/3156-annual-overview-of-marine-casualties-and-incidents-2017.html

Mascaro, S., Nicholson, A., & Korb, K. (2014). International Journal of Approximate Reasoning Anomaly detection in vessel tracks using Bayesian networks, 55, 84–98. https://doi.org/10.1016/j.ijar.2013.03.012

Mazaheri, A., Montewka, J., & Kujala, P. (2014a). Modeling the risk of ship grounding—a literature review from a risk management perspective. WMU Journal of Maritime Affairs, 13(2), 269–297. https://doi.org/10.1007/s13437-013-0056-3

Mazaheri, A., Montewka, J., & Kujala, P. (2014b). Modeling the risk of ship grounding—a literature review from a risk management perspective. WMU Journal of Maritime Affairs, 13(2), 269–297. https://doi.org/10.1007/s13437-013-0056-3

Michal Horný. (2014). Bayesian Networks. Technical report (Vol. 5).

Mitroff, I., & Alpaslan, M. C. (2003). Preparing for. Harvard Business Review, 81(October), 109–115.

Montewka, J., Ehlers, S., Goerlandt, F., Hinz, T., Tabri, K., & Kujala, P. (2014). A framework for risk assessment for maritime transportation systems - A case study for open sea collisions involving RoPax vessels. Reliability Engineering and System Safety, 124, 142–157. https://doi.org/10.1016/j.ress.2013.11.014

Montewka, J., Goerlandt, F., & Kujala, P. (2014). On a systematic perspective on risk for formal safety assessment (FSA). Reliability Engineering and System Safety, 127, 77–85. https://doi.org/10.1016/j.ress.2014.03.009

Murphy K. (1998). A Brief Introduction to Graphical Models and Bayesian Networks.

Nidaa, B. (2019). Flag states failing to meet casualty reporting rules. Lloyd's List Investigates, 44(0), 1–13.

Norway, P. S. (2016). Clarification of the risk concept.

O'Neil, W. A. (2003). The human element in shipping. World MaritimeUniversity Journal of Maritime Affairs.

ÖZDEMİR, Ü., & GÜNEROĞLU, A. (2015). Strategic Approach Model for Investigating the Cause of Maritime Accidents. PROMET - Traffic&Transportation, 27(2), 113–123. https://doi.org/10.7307/ptt.v27i2.1461

Pagiaziti. (2014). Thesis on maritime accidents. National Metsovio Polythechnic.

Pazara HR, B. E. (2008). Reducing of Maritime Accidents caused by human factors using simulators in training process. Journal of maritime research.

Price, M. (2011). The risks of night work. Retrieved August 16, 2019, from https://www.apa.org/monitor/2011/01/night-work

S, G., K, R., & M, O. (1997). Impact of human elements in marine. Pergamon.

Scutari, M. (2010). Learning Bayesian Networks with the bnlearn R Package. Journal of Statistical Software, 35(3), 1-22. Retrieved from http://www.jstatsoft.org/v35/i03/.

Shaughnessy, T. T. (2007). Flag of Convenience: Freedom and Insecurity on the High Seas. Journal of International Law and Policy, 1-31.

Silveira, P. A. M., Teixeira, A. P., & Soares, C. G. (2013). Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal. Journal of Navigation, 66(6), 879–898. https://doi.org/10.1017/S0373463313000519

Spiegelhalter, D. J. (2002). Bayesian graphical modelling: a case-study in monitoring health outcomes. Journal of the Royal Statistical Society: Series C (Applied Statistics), 47(1), 115–133. https://doi.org/10.1111/1467-9876.00101

The United Nations Convention on the Law of the Sea. (1982). United Natios, p. Article 91.

Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. Ecological Modelling, 203(3–4), 312–318. https://doi.org/10.1016/j.ecolmodel.2006.11.033

Vanem, E., Anta, P., Østvik, I., Del, F., & Comas, C. De. (2008). Analysing the risk of LNG carrier operations, 93(March 2006), 1328–1344. https://doi.org/10.1016/j.ress.2007.07.007

Ventikos, N. P., Stavrou, D. I., & Andritsopoulos, A. (2017). Studying the marine accidents of the Aegean Sea: critical review, analysis and results. Journal of Marine Engineering and Technology, 16(3), 103–113. https://doi.org/10.1080/20464177.2017.1322027

Veritas, B. (2006). CASUALTY ANALYSIS OF TANKERS.

Wagenaar, Willem, A., & Groeneweg, J. (1987). Accidents at sea: Multiple causes and impossible consequences. International Journal of Man-Machine Studies, 27(5–6), 587–598. https://doi.org/10.1016/S0020-7373(87)80017-2

Xia, Y., Xiong, Z., Dong, X., & Lu, H. (2017). Risk assessment and decision-making under uncertainty in tunnel and underground engineering. Entropy, 19(10). https://doi.org/10.3390/e19100549

Yang, B., Zhao, Z., & Ma, J. (2018). Marine accidents analysis based on data mining using K-medoids clustering and improved A priori algorithm. IOP Conference Series: Earth and Environmental Science, 189(4). https://doi.org/10.1088/1755-1315/189/4/042006

Zealand, M. S. (1996). Maritime accidents.

Annex

List of annexes

ANNEX 1: SHIP IDENTIFICATION AND PARTICULARS
Indicates the information to be submitted in all casualty reports.

ANNEX 2: DATA FOR VERY SERIOUS AND SERIOUS CASUALTIES
Indicates information to be supplied on "very serious" and "serious" casualties.

ANNEX 3: SUPPLEMENTARY INFORMATION ON VERY SERIOUS AND SERIOUS CASUALTIES
Additional information required for "very serious" and "serious" casualties.

ANNEX 4: INFORMATION FROM CASUALTIES INVOLVING DANGEROUS GOODS OR MARINE POLLUTANTS IN PACKAGED FORM ON BOARD SHIPS AND IN PORT AREAS
This form may be applicable for marine casualties as defined as well as marine incidents.

ANNEX 5: DAMAGE CARDS AND INTACT STABILITY CASUALTY RECORDS
This form may apply to "very serious" and "serious" casualties.

ANNEX 6: FIRE CASUALTY RECORD
This form may apply to "very serious" and "serious" casualties.

ANNEX 7: QUESTIONNAIRE RELATED TO THE GLOBAL MARITIME DISTRESS AND SAFETY SYSTEM
This form may apply to "very serious" and "serious" casualties.

ANNEX 8: FATIGUE AS A CONTRIBUTORY FACTOR TO MARITIME ACCIDENTS – FATIGUE FACTORS DATA COMPILATION SHEET
This form will apply where fatigue is deemed to be a contributory factor in the casualty.

ANNEX 9: INCIDENTAL SPILLAGES OF HARMFUL SUBSTANCES OF 50 TONNES OR MORE
This form relates to incidents involving harmful substances. The report is considered necessary when investigating a casualty or an incident (MARPOL, articles 8 and 12), however this does not replace the one-line entry report required by the annual mandatory report under MARPOL, article 11 (MEPC/Circ.318, Part 1).

ANNEX 10: LIFE-SAVING APPLIANCE CASUALTY RECORD
This form is for all casualties involving life-saving appliances, adding any other information which would provide lessons to be learned concerning the use of this equipment.

Figure 24 List of Annexes in IMO reports

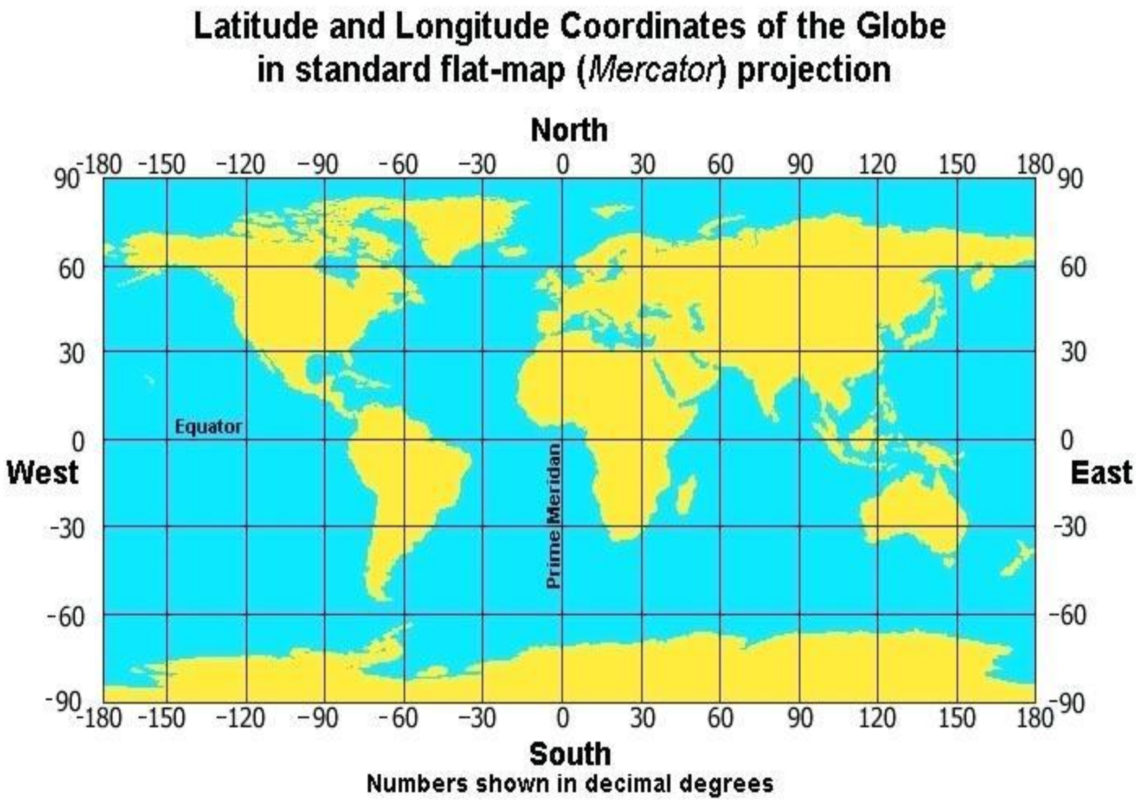

Figure 25 World map with latitudes and longitudes