SAR Observations and Wave Modeling Reveal Relatively High Values of the BFI in the Agulhas Current Retroflection

Article

A Detailed Identification of Classificatory Variables in Ship Accidents: A Spanish Case Study

Candela Maceiras [1], José M. Pérez-Canosa [1], Diego Vergara [2] and José A. Orosa [1,*]

1 Department of Navigation Science and Marine Engineering, University of A Coruña, Paseo de Ronda, 51, 15011 A Coruña, Spain; candelaria.maceiras@udc.es (C.M.); jose.pcanosa@udc.es (J.M.P.-C.)
2 Department of Mechanical Engineering, Catholic University of Ávila, C/Canteros, s/n, 05005 Avila, Spain; diego.vergara@ucavila.es
* Correspondence: jaorosa@udc.es; Tel.: +34-981-167-000 (ext. 4320)

Abstract: The present paper shows an original study of more than 163 ship accidents in Spain showing which of the usually employed variables are related to each type of vessel accident due to the lack of information in this region. To this end, research was carried out based on the Spanish Commission for Investigation of Maritime Accidents and Incidents (CIAIM) reports. Detailed combinatory ANOVA analysis and Bayesian networks results showed a good agreement with studies of other regions but with some particularities per each type of accident analyzed. In particular, ship length was defined as the more relevant variable at the time to differentiate types of accidents. At the same time, both the year of build and the fact that the ship meets the minimum crew members required were excellent variables to model ship accidents. Despite this, the particularities of the Spanish Search and Rescue (SAR) region were defined at the time to identify accidents. In this sense, although variables like visibility and sea conditions were employed in different previous studies as variables related to accidents occurrences, they were the worst variables to define accidents for this region. Finally, different models to relate variables were obtained being the base of deterministic dynamic analysis. Furthermore, to improve the accuracy of the developed work some indications were obtained; revision of CIAIM accidents scales, identification of redundant variables, and the need for an agreement at the time to define the classification limits of each identification variable.

Keywords: ship; identification; accidents; modelling; ANOVA; Bayesian network

1. Introduction

The traditional approach to studying or analyzing maritime safety is generally reactive rather than proactive. Hence, the importance of historic technical studies to try to foresee risk situations. Knowing which situations involve greater risk will in turn mean anticipating a possible accident. An accident theory is a collection of propositions to illustrate the principles that lead to accidents and an accident model is a simplified description of a system or process to help present the occurrence of accidents based on an accident theory [1]. According to Perrow [2], a normal accident is an intrinsic characteristic of any system. Given the characteristics of a system, multiple and unexpected interactions causing failures are inevitable. In this sense, the interactive complexity and the close characteristic coupling of the system inevitably produces an accident called "Normal Accident" or "System Accident". These accidents are the ones that must be accepted as an acceptable risk.

According to a study of Allianz, shipping in the 21st century is now safer than ever [3], probably due to a combination of factors both at technical level (technological advances), regulatory environment (international and national regulations), and fundamentally, to the influence of human factor (preparation and training). In this sense, it is in standard O-134: International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA) Recommendation on the Risk Management Tool for Ports and Restricted Waterways

(2009) [4], where it addresses the IALA Waterway Risk Assessment Program (IWRAP) and Ports and Waterway Safety Assessment (PAWSA) tools.

The aim is to make available to the Member States aid for the measurement (quantitative and qualitative) of risk. The IWRAP MK2 tool, which IALA makes available to member states, calculate risk, based on the theory of Fujii and MacDuf (1974) [5,6] and presents results based on the possibility of an accident (although only collisions and groundings). Other researchers like Hollnagel and Goteman [7], proposed the Functional Resonance Accident Model, FRAM in its English nomenclature (for Functional Resonance Accident Model). The idea suggests that accidents occur due to functional resonance within a system.

To reach this so complex objective a revision of the evolution of accidents analysis is needed. In this sense, previous research works [8,9] indicated that around 75 to 96% of maritime accidents are caused, at least in part, by some type of human error. Specifying that this human error contributes to between 89% and 96% of collisions, 75% of explosions, 79% of landings, and 75% of collisions. What is more, 56% of the approaches are caused by non-compliance or violation of Convention on the International Regulations for Preventing Collisions at Sea, COLREGs [10,11].

Despite the previous indications, no single model has the capability of being useful to satisfy all demands of the maritime industry at any moment. To solve this problem, different research works were done in the last decades. In this sense, in 2007, Ibn Awal et al. [12], proposed a dynamic model for maritime accident focused on ship-to-ship collision, using two models of simulations, pre and post-accident, with the goal being to increase the maritime transport safety. Although it can be considered one of the accidents that has the worst consequences in the maritime world (for own ships, crewmembers, cargo, and the environment), the model does not include aspects so relevant as the waves or wind forces. This fact can give us an idea of the difficulty of proposing a model that can satisfy the huge number of variables that can be present in maritime transport.

In 2005, Liu et al. [13] analyzed more than 100 marine incidents in order to improve the navigation safety system (frequencies and causes) of a specific area through the Gray Clustering Method (Gray system Theory). In this research work, it highlights the need of models that can be useful for areas where the maritime accidents/incidents database is incomplete or poor due to the unpredictability of maritime accidents, a fundamental characteristic of this Theory. Furthermore, a reliable database is not always available everywhere due to the opacity that the maritime industry sometimes presents when an accident happens. In consequence, it is shown how the knowledge of frequency and causes of maritime accidents can be very useful for shipping lines, underwriters and, in general, stakeholders in the maritime industry.

Once the main data are collected, they must be analyzed, and the probabilistic method is one of the most used in the maritime industry, in addition to simulation modelling and statistical analysis of data [11]. In this sense, in 2009, Uluscu, et al. [14] developed a paradigmatic case such as the navigation risks in the Strait of Istanbul. In this study, the authors implement a safety risk analysis with a model based on probabilistic data. In the mathematical model proposed, factors that define situations and affecting the probability of accident are grouped in vessel attributes (type, length, age, flag, tugboat, and pilot assistance), and environmental attributes. Attributes also influence consequences; in this case, vessel attributes (type and length) and shore attributes (populations, property, and infrastructure).

After slicing the Strait in several legs and collecting local information from different sources, the risk is calculated in two-tier accident types, related to a set of instigators that may cause an accident (human error, steering, propulsion, communication, navigation, mechanical, or electrical failure). As a weak point, in part of a study, authors resort to the opinion of two experts (for comparative purposes) due to the lack of a solid database, which can help to clarify the final result. As an example, the plotting of the normal probability of residuals and residuals vs. predicted variable of pair is shown, focused mainly on the human error and its influence in the collision.

At the same time, at the time to define the variables related to accidents in vessels, it is of interest to center this study on fishing vessels due to the high number of accidents. In particular, recent review works based on journal articles and reports from the maritime authorities in Poland, United Kingdom, Norway, Iceland, Denmark, United States and Alaska, and Canada [15] showed that these fishing vessels are still 50% higher at risk than on-shore workers, which is particularly high with respect to other types of vessels.

In addition, most of the previous works about fishing vessel accidents were centered on just one main variable, for instance, human factor [16–19]. Despite this, other variables were indicated as the main causes of fishing vessel accidents other than human error [19,20], e.g., weather conditions, operational status of the vessel, vessel location, seasons, and unsuitable fishing equipment [21–23]. In particular, a statistical study of accidents in fishing vessels in the northeastern United States emphasized that the probabilities of accidents in fishing vessels increase with wind speed and with placement near the shore and in the winter season [24]. In this same year, more research studies showed Marine Accident Investigation Branch (MAIB) as the better data source regarding fishing vessels accidents despite a certain tendency of under-reporting [19]. In this work, a conclusion was obtained that confirms the previous comments; the risk of accidents on fishing vessels increases as vessel length decreases. What is more, a more recent study about fishing vessel accidents in the northeastern of the United States [25] showed that the damage severity of accidents, defined in accordance with the Damage equation obtained by [26], increases with loss of stability, sinking, daytime wind speed, vessel age, and distance to shore, in agreement with previous studies. At the same time, it was obtained that the damage severity of accidents on fishing vessels increases as vessel length decreases (inversely proportional).

At the time, to identify the variables related to accidents, different research works about the analysis of the variance (ANOVA) were done. In particular, a statistical analysis centered on questionnaires and posterior ANOVA and correlations analysis was done by [16]. Its main results showed that there were clearly differentiated groups of age, vessel types, and occupation, among others. At the same time, in this work, there were assumptions of regression analysis such as linearity, normality, and multicollinearity, which were tested and found satisfactory. In consequence, it was possible to define the coefficient of the relation between related variables. From this relation it was concluded that the safety attitude is a fundamental parameter that depends on the previous experiences like people that sought medical attention (less positive attitude) and fishermen involved in accidents (positive attitude).

Once the main variables of these accidents were identified, there was a need to predict these accidents. In this sense, Mullai and Pulasson [27] developed a conceptual model based on the principles of Grounded Theory and content analysis of empirical data, in order to explain and predict maritime accidents. For that purpose, a database containing the marine accidents classified by ship and variable was used. Most of the variables were non-metric and were grouped into 11 main categories (constructs). The design of variables was based on the European Statistics guidelines, the International Maritime Organization (IMO) investigations code, and the DAMA coding system (agreed in 1990 by the Scandinavian countries for the analysis of maritime accidents). In total, 87 variables were taken into account for designing the model, reduced to 11 "sets" or "constructs", grouped as per their common properties. Therefore, the model carried out consisted of a multilevel model of 87 variables in complex relationships and permits one to obtain the correlations among a large number of independent variables as ship's age, length, Gross Register Tonnage (GRT), and the number of persons on board.

Another paper related to human factors is based on the human factors index system of ship accidents to develop a multidimensional association rules algorithm by incorporating the Reason model and classic correlation rules algorithm [28]. Other works are centered on human factor as the main cause of most serious maritime accidents [29]. Authors use the model Human Factor Analysis and Classification System (HFACS) based on the Swiss Cheese model of human error, which had been developed to provide a methodological tool

to investigating an accident in the aviation industry. The modification of this model had been implemented by the same authors to maritime accidents some years before, and for this paper, more than 150 real cases were handled. The model consists of two interrelated causal sequences: An active failure pathway and a latent failure pathway. Once coded, the causal factors are presented and divided (organizational influences, unsafe supervision, preconditions for unsafe acts and unsafe acts). Afterwards, the first-, second-, and third-tier results are depicted.

More recent works try to address the human factor and statistical procedure together to quantify the correlation path mode of the causal factor involved in marine accidents, making use of complex structural chains supported by HFACS together with the statistical method Structural Equation Modelling (SEM), used to quantitatively analyse the relationships among human factors in accidents [30]. The paper is focused on the human factor because authors consider that in the shipping industry, despite having achieved a high level of safety from the point of view of equipment, human behavior still remains the cause of many maritime accidents. In the research, a Swiss Cheese model is also produced, and a novel method to analyze causal factors in the accidents is introduced with the help of SEM, due to the difficulty of doing the basic premises of path analysis (relationship among variables).

The measurement equation of SEM describes the relationship between the observed dependent variable and the latent independent variable. Structure equations describe the relationship between the latent variable. For the study, a database of 894 accident with the presence of human error as the cause was taken. Depending on the consequences, five levels were selected: Incidents, minor accidents, general accidents, major accidents, and serious accidents. The data were integrated into 16 major accident factors and path diagrams were carried out to check the accuracy and reliability of the model (relationships of observation or indicator variables).

Other studies are not centered on the human factor and develop their predictive procedures like a risk probabilistic model [31] in order to evaluate the probability of a ship-grounding accident taking into account the causal factors. The database contained more than 200 accident and incident reports, although using only a single accident report for modelling has its own disadvantages. As in other papers, the grounding model was developed used the Human Factor Analysis and Classification System for Grounding (HFACS-Ground). In the model construction, in order to set up a qualitative causal model, the Bayesian Belief Network (BBN) modelling was used. The model can quantify the probability of grounding given a set of input parameters.

Finally, in recent works [32], this same methodology of a chi square method and Bayesian networks were employed to be a help to estimate the occurrence of accidents when the main variables of the problem change.

In the present work, based on these previous research works and the need for in-depth study of these parameters, a real case study has been conducted analyzing more than 163 accidents in the Spanish Search and Rescue (SAR) sea regions in the time frame between 2008 and 2017. For this purpose, the official reports published by the Spanish Commission for Investigation of Maritime Accidents and Incidents (CIAIM) were taken into account [33]. In particular, a procedure similar to previous research activities was employed but employing a more detailed analysis, replacing the VCH square method by the Analysis of Variance for each type of accident. This procedure will help to identify the more appropriate variable that allows us to identify the more probable type of accident in the particular sea region of the Spanish Search and Rescue (SAR). This initial study will be the base case to predict the occurrence of accidents when the main variables of the problem change based on future Bayesian networks studies.

2. Materials and Methods

2.1. CIAIM Accidents Analysis

The Spanish Commission for the Investigation of Maritime Accidents and Incidents, CIAIM [33] performs investigations into maritime accidents and incidents in Spanish waters, including inland waters, the territorial sea, and any area outside the territorial sea upon which, in accordance with international law and on application of its domestic legislation, the Kingdom of Spain exercises jurisdiction or sovereign rights. In particular, in this study, the Spanish SAR territory has been selected. That is, the waters where the Spanish State holds responsibility in terms of search and rescue missions. At the same time, the selected period (2008–2017) was in accordance with the need to circumscribe our analysis to a stable regulatory framework, essential for the coherence of the comparative conclusions. In consequence, more than 163 accidents were analyzed in this case study.

Based on both these reports and previous research works [16,31–33], the different typical variables usually employed in those reports and associated with some types of accidents were classified. In this sense, the variables selected for this study, per accident were: Type of Ship, Year of construction, Wind direction, Wind force, Sea condition, Night or day accident, Visibility, Ship Breadth, Ship Length, GT, Crewmembers, and the Minimum number of crew members required. At the same time, the causes of the accidents were also identified for future modelling applications. Finally, all these variables were codified in accordance with standard scales like the Beaufort scale for wind force or the Douglas scale for sea conditions, and other variables were codified in accordance with the typical scales employed in the own reports. All the variable and its codes are shown in Tables A1–A8 (Appendix A).

2.2. Software Resources

Different software resources were employed to develop each different statistical study. In this sense, to develop such specific statistical studies like One-Way ANOVA, the software Statistical Package for the Social Sciences (SPSS) version 22 was employed [34]. Despite this, due to the need for multivariable curve fit of response surface modelling, the software Minitab version 18 was selected [35]. To classify the information, different datasheets were employed.

2.3. Statistical Analysis

In the present paper, different statistical studies were employed. First of all, the descriptive statistic and histograms were developed to identify each variable. After that, to identify the variables actually related to accidents, a One-Way ANOVA was employed. This statistical procedure determines whether the groups analyzed (defined by the levels of the independent variable (accidents)) are different in their means respect the overall mean of the dependent variable. To apply this study, some assumptions must be considered; independence of observation (data collection), normally distributed response variable (it will be showed by histograms), and homogeneity of variance. The other important curve-fitting statistical procedure employed in this research work was the response surface technique. This is a polynomial curve fitting with more than 3 variables (which is the main limitation of the general curve fitting defined in most of the software resources). It is an initial curve fitting tool that, like a neural network, models a process in accordance with a minimum square procedure but with the advantage that gives us the mathematical model obtained (IA is based in a black box).

2.4. Methodology

The original methodology employed in this research work is based on the inference statistical procedure applied to the particular case study of ship accidents understanding. In this sense, the first step was to define the correlation between each variable. It is a linear correlation defined by the determination factor and just for linear relations will show an adequate value over 0.9.

After that, an ANOVA analysis between the accidents type and each different selected variable (selected in accordance with CIAM reports) will show the variables related to the accidents and not identified by a linear correlation, taking it a step further than previous works [16]. Besides, as it is evident, not all the variables are related in the same way with all the types of accidents so a specific ANOVA analysis between each accident with respect to the other is a more in-depth analysis that what has been done in previous statistical works. In consequence, the relevance of each variable to the detection of each different type of accident will be described and, in consequence, the adequacy of the employed scale in some variables. Finally, due to about 7 dependent variables that were related to the accident type, a response surface model will be defined to relate this influence with the type of accident (TA) as an initial general identification model. In this sense, based on the previous information obtained from the initial ANOVA study, different models that related these independent variables will be shown. This process is summed up in Figure 1.
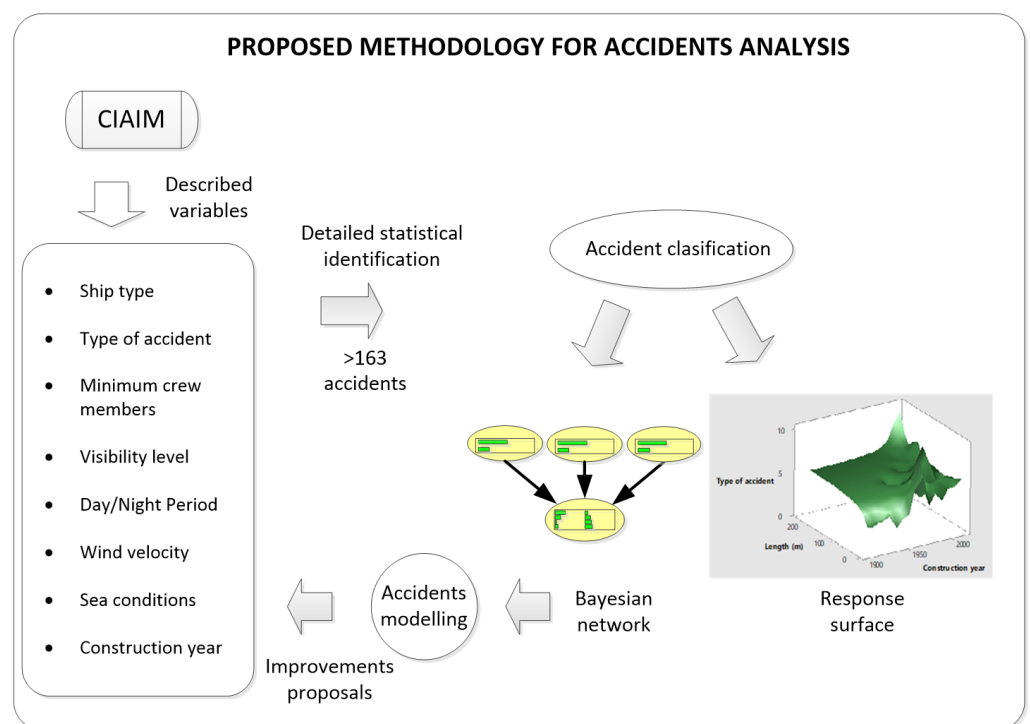
Figure 1. Proposed methodology.

3. Results

3.1. Linear Correlation Analysis

In this section, an initial correlation study between the main selected variables was done. As we can see in Table 1, the determination factor is very high when it is obtained between two variables. For instance, the highest value was obtained between length (L) and breadth (B) (0.95), which is a common-sense value due to some design criteria of these types of merchant ships. This same length variable is directly related to gross tonnage (0.79) and the number of crew members (0.78). This is another logical result due to the fact that these two new variables are usually related to the ship dimensions. A similar relation, but with lower determination factor, was obtained between the breadth and the Gross Tonnage, GT, (0.70) and between the breadth and the number of crewmembers (0.77). To sum up this initial correlation study, it can be concluded that the length of the ship is a representative variable that can replace other variables like breadth, GT, and number of crewmembers at the time to develop a mathematical model to identify the type of accident.

Table 1. Linear correlation analysis.

	ST	TA	CY	WD	WF	SC	V	N	L	B	GT	CA	CM	MNC
ST	1.00													
TA	−0.02	1.00												
CY	0.00	0.22	1.00											
WD	0.02	−0.02	−0.02	1.00										
WF	0.13	0.02	0.12	−0.04	1.00									
SC	0.06	−0.01	0.12	−0.05	0.77	1.00								
V	−0.01	−0.11	−0.03	−0.12	−0.04	−0.05	1.00							
N	0.04	−0.07	0.01	0.15	−0.12	−0.08	0.08	1.00						
L	0.00	0.09	0.18	0.02	0.14	0.21	−0.04	−0.12	1.00					
B	0.08	0.08	0.20	0.04	0.15	0.18	−0.05	−0.10	0.95	1.00				
Gt	−0.09	0.08	0.12	0.03	0.03	0.15	0.01	−0.09	0.79	0.71	1.00			
CA	−0.08	−0.16	−0.14	0.02	0.00	0.08	0.11	−0.09	−0.01	−0.02	0.01	1.00		
CM	−0.12	0.04	0.12	0.04	0.07	0.10	−0.04	−0.06	0.78	0.77	0.59	0.02	1.00	
MNC	0.10	−0.46	0.07	0.00	−0.14	−0.21	0.02	0.09	−0.02	0.00	−0.06	−0.17	−0.04	1.00

ST: Type of ship, TA: Type of accident; CY: Year of build; WD: Wind direction; WF: Wind force; SC: Sea condition; V: Visibility; N: Nocturnal; L: Length; B: Breadth; GT: Gross Tonnage; CA: Cause of accidents (Human factor); CM: Crewmembers; MNC: Minimum number of crew members.

Another interesting determination factor is that which relates to the type of accident and if the ship fulfills the minimum crewmembers required in the moment of the accident (0.46). In consequence, these two variables can be defined as of interest, but cannot be derived one from another due to the not-so-high determination factor like in the previous cases.

Another high determination factor was derived from the study of the wind force and sea conditions (0.76). It is another logical result that will let us replace both variables by just one at the time to define the main variables related to a ship accident.

Finally, a reduced determination factor of 0.22 shows a slight relationship between the type of accident and the year of construction. In this sense, caution must be taken at the time to understand this so low value. As it will be shown in the next sections, the year of build is a decisive variable, but it does not exert a linear effect over accidents. In consequence, posterior, not linear, correlation between variables must be done for the variables with a low linear determination factor; for instance, wind direction, type of ship, visibility, night period, the cause of the accident, and the real number of crew members were not directly related to the type of accident.

3.2. One Way ANOVA

3.2.1. Accidents Independent Variables Recognition

As explained before, due to the determination factor shown in the correlation analysis being related to a line relation between variables, a One-Way ANOVA study must be done. In this sense, a significance of 0.05 was selected as it used to be done in this kind of technical analysis. Main results are shown in Tables 2 and 3 for an ANOVA analysis between variables and type of accidents and between variables and causes of accidents, respectively.

In these two tables, the variables with a significance below 0.05 were highlighted in grey due to the fact that they are statistically different in the categories or groups of the independent variable (the type of accident or cause of the accident, respectively). From Table 2, it can be observed that year of build, ship length, and the cause of accident have lower significant values, and, in consequence, they show great differences with respect to the type of accident; or what is more, they have more possibility to be a better variable for a curve fitting model that predicts the type of accident. As opposite variables, the number of crew members and wind direction cannot be defined as having changed for each different type of accident, so they are not good variables for a future modelling process.

Table 2. Identification of related variables with accidents (significance).

Type of Ship	ST	0.007
Year of build (years)	CY	0.002
Wind direction (degree)	WD	0.519
Wind force	WF	0.107
Sea condition	SC	0.014
Nocturnal	N	0.000
Visibility	V	0.007
Breadth (m)	B	0.979
Length (m)	L	0.002
Cause of accidents (Human factor)	CA	0.002
GT (tonnes)	GT	0.107
Crew members	CM	0.438
Minimum number of crew members	MNC	0.033

Finally, it is interesting to understand that, despite the fact that the cause of each accident is a good curve fitting variable, it cannot be employed to develop a model to predict the type of accident. What is more, as it can be observed in Table 3, the cause of the accident can just be related to the GT and wind force variables.

Table 3. Identification of variables related with the cause of the accident (significance).

Year of build	0.290
Breadth (B)	0.568
Gt (tons)	0.000
Crew members	0.951
Minimum crew members	0.563
Type of Ship	0.076
Type of accident	0.341
Wind direction (degree)	0.286
Wind force	0.000
Sea condition	0.056
Visibility	0.076
Night or day	0.341
Length (m)	0.290

3.2.2. In-Depth Analysis of Types of Accidents Identification Variables

As it was explained before, the initial classification of types of ships, sea conditions, etc., were developed as common sense and in accordance with the usual procedure in this type of accidents. Despite this, the selected a variable and its scale may not be statistically adequate to identify the different type of accidents. In consequence, it is necessary to develop an ANOVA study employing each independent variable to identify the real statistical difference to each different type of accident respect the others. To do so, the first goal of the study was to do the ANOVA analysis of the accident type a1 when it is compared with the others, employing the ship Type (Appendix B, Table A9), year of construction (Appendix B, Table A10), sea conditions (Appendix B, Table A11), the time period (Appendix B, Table A12), visibility level (Appendix B, Table A13), the ship length (Appendix B, Table A14), and the Minimum crew member required (Appendix B, Table A15) as independent variables.

From Table A9 (Appendix B), as a result of the ANOVA analysis, it can be concluded that accidents a2, a6, a8, and a10 can be identified as significantly statistically different to accident a1 by employing the type of ship as the comparative variable. So, this variable can be employed to predict future accidents and its type. In other words, there are different types of ships for each type of accident. Thus, from the data, it can be seen that accident 1 (a1) is associated with a vessel type 3, 2, 7, 8, and 5 accident 3 (a3) is associated with ship types 3, 7, and 10; accident 4 (a4) to ship type 3 and 2; and accident 5 (a5) to type 4, 3, and 2. In particular, accident type 2 (a2) is always with type 3 vessels. Finally,

it is interesting to remember that it does not mean that the other groups of column 1 of Table A9 (Appendix B) are equal, they are just not dependent groups to this a1 variable.

If we now compare each of these tables, it can be concluded that visibility level and sea conditions are the worst variables to differentiate types of accidents. At the same time, the year of build and if the ship meets the minimum crew members required are better variables to differentiate types of accidents, in agreement with the initial general ANOVA analysis. Despite this, it must be observed that ship length shows, only for the types of accidents that it can differentiate, the lowest significance level so it is a stronger differentiation parameter. What is more, the year of construction (Table A10, Appendix B) is the more useful variable to differentiate accidents, except accident type 2 (a2), to the others a1, a3, and a4. It must be related with the fact that, when employing the year of construction as identifier variable, it may be not easy to differentiate sinking, an approach, grounding, or collision due to the fact that these variables maybe not really be related to the antiquity of the ship. This same effect happens in Table A15 (Appendix B) when differentiating a2 (sinking) from the other accidents based on the minimum number of crew members required. In addition, in this particular case, the minimum crew members seems to be slightly worse in identifying accidents than the year of construction based on the difficulty to differentiate a3 (Grounding) from other accidents. Despite this, it is the second-best parameter.

From Table A9 (Appendix B), it can be concluded that ship type can be employed to differentiate the accidents a5 (Fire), a6 (Capsizing), a8 (Stranded), and a10 (operational Accident) from the other accidents and that it is not as useful of a parameter as it is usually considered. From Table A14 (Appendix B), it is observed that the ship length is a good parameter, but it has some difficultness in differentiating a2 (sinking) and a6 (Overturning), while it is better at differentiating the other accidents due to its significance being zero in most of the cells. This important information must be considered, and it can be concluded that, in accordance with the type of accident to be identified, the ship length may be the best one variable. Finally, sea conditions (Table A11, Appendix B), that used to be considered the most important variable to identify a ship accident, was the worst variable due to not allowing researchers to differentiate the types of accidents at the time to classify it.

As an illustrative example of the interpretation of these tables showing the relation between accidents and each related variable, histograms of Figures 2–4 were developed for construction year, ship length, and ship type. From these histograms, one of the requirements of ANOVA analysis can be observed, the normality, as it was done in previous research works [16]. It must be highlighted that this normality can be identified by a single mode: It is asymptotic to the abscissa axis and it is symmetric with respect to the mean. In this sense, Figure 2 shows a wave similar to a normal distribution if it is considered a decreasing frequency from 2007 until present day. In consequence, it seems logical to obtain a good accident identification based on this so characteristic ship variable. From these figures, it can also be concluded that most of the ships were built between 1999 and 2007, with a length between 24 m and 44 m, and belong to ship type 3 (Fishing vessels).

As a final analysis, it is interesting to see in Table A16 (Appendix B) that the human factor is an interesting variable that allows one to differentiate ship contact (a1) from the other accidents with maximum accuracy (sig. 0.000). At the same time, it allows one to differentiate grounding, collision, and fire accidents (a3, a4, and a5) from capsizing and groundings (a6 and a8).

After these initial conclusions, it is interesting to conduct a comparison with results from previous research works. In this sense, it was defined by previous studies that flooding, grounding, and collision are influenced by the human factor in a high percentage [8]. If we analyze this indication with our results, it was observed that, when the accidents were compared in accordance with some variable, for instance, ship type, sea conditions, and visibility, it cannot differentiate the accidents a2, a3, and a4, which implies that another factor must exert a strong influence, presumably the human factor. It is in agreement with previous works' indications, and a possible solution is highlighted in this paper

showing that just ship length allows one to successfully differentiate between the difficult identification of these types of accidents. Finally, all these tables are summed up in Table 4.

From Table 4 it can be concluded that year of construction, crew members, and vessel length are the more useful variables to differentiate between the types of possible accidents in vessels.
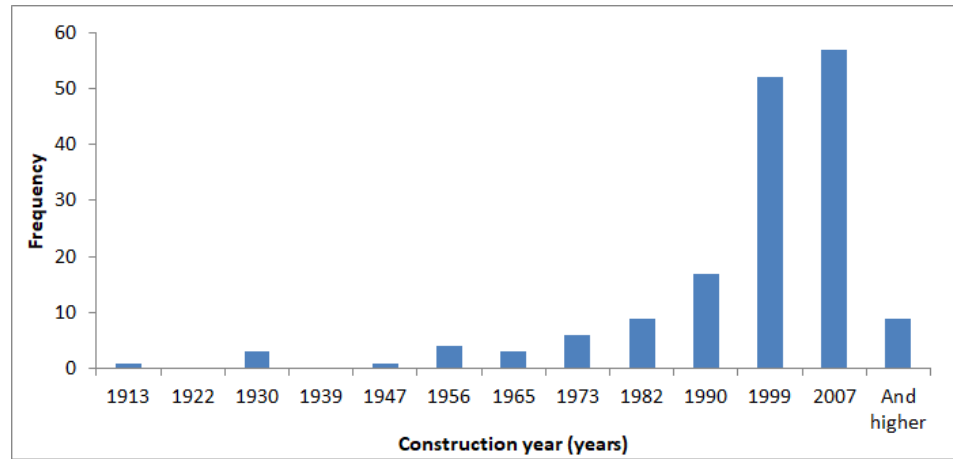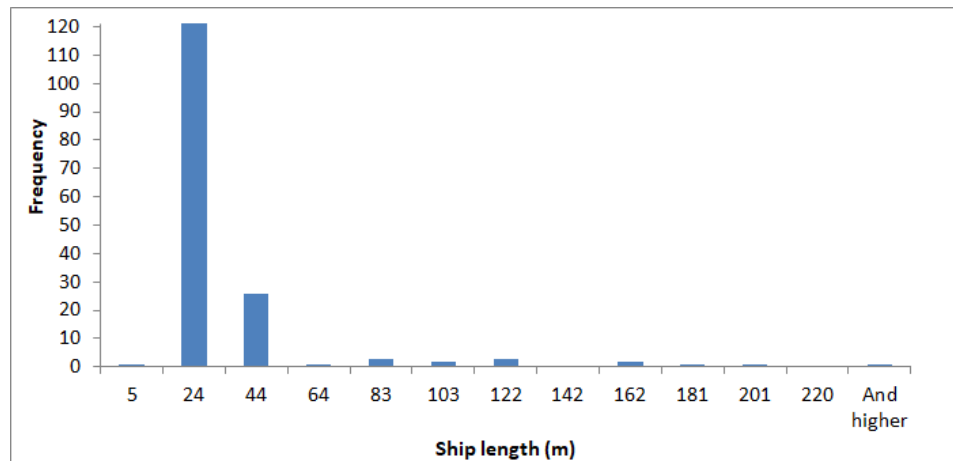

Figure 2. Construction year histogram.
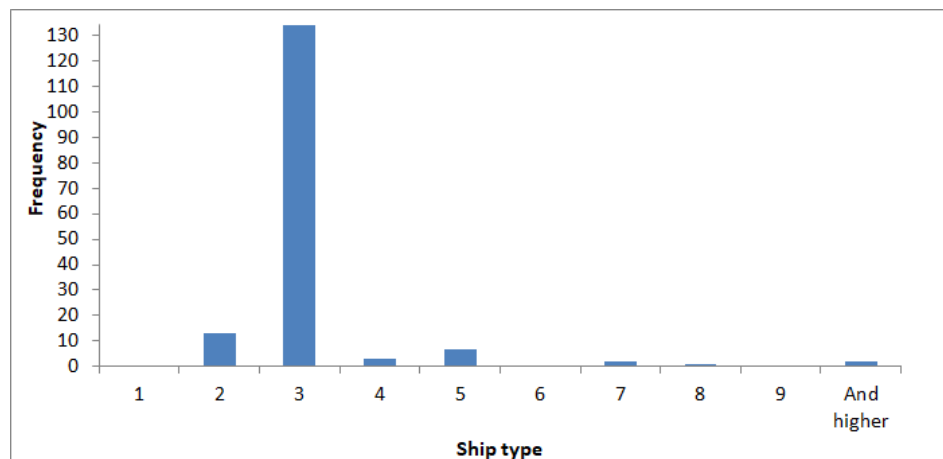

Figure 3. Ship length histogram.

Figure 4. Ship type histogram.

Table 4. Sum up of the utility of each defined classificatory variable per each type of accident.

Classificatory Variables	Indistinguishable Accidents	Differentiable Accidents
Ship type	a1, a2, a3, a4	a5, a6, a8, a10
Year of construction	a2	a1, a3, a4, a5, a6, a8, a10
Sea conditions	a1, a2, a3, a4, a5, a6, a8, a10	
Night or day period	a1, a2, a3	a4, a5, a6, a8, a10
Visibility	a1, a2, a4, a5, a6, a8,a10	a3
Vessel length	a2, a6	a1, a3, a4, a5, a8, a10
Minimum Crew members	a2	a1, a3, a4, a5, a6, a8, a10
Human factor	a2, a3, a4, a5, a6, a8, a10	a1

3.3. Bayesian Networks

In a similar way to the study recently conducted by Uğurlu et al. [32] (Figure 2), in the present research work, a Bayesian network was developed to define the probability of each type of accident based on the causal variables, which are identified in the previous sections.

Previous to develop the graphical model, it is necessary to define the ranges of the three more relevant variables identified in Table 4. In particular, as ship was identified as New if the ship was built after 1996, identified as Long if the ship had a length higher than 12 m, and, finally, it was identified as having a high number of crew members when it was higher than the minimum number required.

Once these ranges were defined, it is possible to define the initial probability of each one, as it is reflected in Figure 5. Despite this, it may be not a realistic value that allows us understand the importance of each variable, so further analysis must be done.
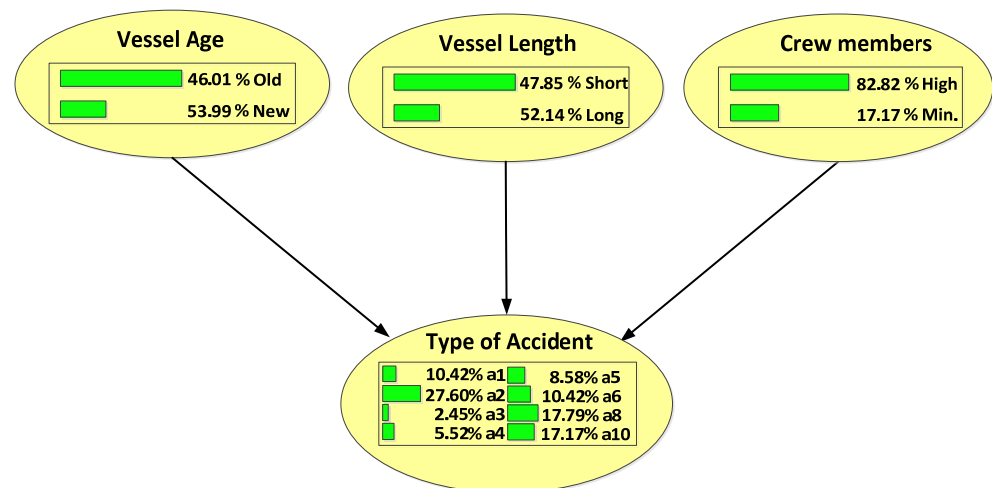
Figure 5. The initial probability of each different type of accident.

In this sense, by applying the Bayesian theorem, it is possible to define the new probability of each related variable (relevance) once it is known that one type of accident happens. In consequence, the Bayesian theorem allows us to define the new probability of vessel age, vessel length, and vessel number of crew members for the more representative accidents (a1, a4, and a10), in a similar way to how it was done in previous studies; this is reflected in Table 5.

As it was explained before, Table 5 shows the relevance of each related variable once a type of accident is identified. In this sense, from this table, an increase can be observed of the probability (from 17.17 to 58.82%) of the relationship between accident a1 (contact) and a minimum compulsory number of crew members employed. The same happens with

this same accident, being more strongly related to a high ship length, and experiencing an increment of probability from 46% to 77%.

Table 5. Probability of each different causes once it is known the type of accident happened (Posteriori).

Accident	Model	Initial Probability(%)	Final Probability(%)
a1 (Contact)	P (Age = Old/Accident = a1)	46.01	41.17
a1 (Contact)	P (Age = New/Accident = a1)	53.99	58.83
a1 (Contact)	P (Crew = Min./Accident = a1)	17.17	58.82
a1 (Contact)	P (Crew = High/Accident = a1)	82.82	41.18
a1 (Contact)	P (Length = Short/Accident = a1)	47.85	23.52
a1 (Contact)	P (Length = Long/Accident = a1)	52.14	76.48
a4 (Operational accident)	P (Age = Old/Accident = a4)	46.01	77.77
a4 (Operational accident)	P (Age = New/Accident = a4)	53.99	22.23
a4 (Operational accident)	P (Crew = Min./Accident = a4)	17.17	11.11
a4 (Operational accident)	P (Crew = High/Accident = a4)	82.82	88.89
a4 (Operational accident)	P (Length = Short/Accident = a4)	47.85	77.77
a4 (Operational accident)	P (Length = Long/Accident = a4)	52.14	22.23
a10 (Collision)	P (Age = Old/Accident = a10)	46.01	42.85
a10 (Collision)	P (Age = New/Accident = a10)	53.99	57.15
a10 (Collision)	P (Crew = Min./Accident = a10)	17.17	14.28
a10 (Collision)	P (Crew = High/Accident = a10)	82.82	85.72
a10 (Collision)	P (Length = Short/Accident = a10)	47.85	28.57
a10 (Collision)	P (Length = Long/Accident = a10)	52.14	71.43

In a similar way, when an operational accident happens (a4), a reduced ship length experiences an increment of implication in the accident by increasing the conditioned probability from 47% to 77%.

Finally, once again, it was observed that, when ship length is high, it exerts an increment of its relevance and conditioned probability in collision accidents (a10). In this particular situation, there is an increment from 52.14% to 71.43%, as it is shown in Table 5.

All these results were in agreement with previous works developed in other regions like that done by [32]. In general terms, a clear agreement with [32] was obtained in the identification of the relation between vessel length, age, and accident types.

Despite this, some differences with respect to. previous works were obtained. In this sense, sea conditions do not seem to be related to most of the accidents analyzed, in disagreement with [32] that identifies sea conditions and weather as some of the more relevant variables related to accidents. It must be related to the fact that the region that is the object of this study does not experience bad weather conditions or casualties. In consequence, there is a need to develop a recompilation of real case studies of different regions to develop a general stochastic model.

Finally, old ships (in this research work, identified as that with more than 24 years) only showed a higher relevance in operation accidents, in disagreement with the general conclusion obtained by [32]. Once again, it must be related to the particularities of each region object of study. In consequence, a ship age limitation of 20 years, as it is proposed by other researchers [32], may be not of interest in this particular fishing region.

At the same time, another difference with previous studies was obtained at the time to analyze ship length effect. The research work developed by [32] concludes that in smaller ships, the accidents increase. In our case study, this conclusion was obtained for operation accidents and the inverse effect for contact and collision accidents. In this sense, it is of interest to introduce the concept of a small ship due to it dependence on the limits fished at the time to analyze the problem. In our particular case study, the classificatory limit value of 12 m was selected and in [32], it was 24 m in accordance with different national standards.

From these results, it can be concluded that there is a need to develop particular case studies for different regions and to obtain variable conclusions per each type of accident due to the differences in accidents between regions paper.

3.4. Response Surface

3.4.1. Accident Modelling

As it was explained in previous sections, most research works try to define the stochastic probability of each variable in each type of accident [36–38], but once the main variables are identified, it is of interest to define a deterministic model that allows us to understand the dynamic risk of each situation as it used in different engineering areas.

In this sense, a modelling procedure based on the response surface may help to obtain a mathematical model that allows us to relate the type of accident and the main independent variables. As it was shown in Table A1 (Appendix A), the ANOVA study allowed us to define the main related variables identified with a significance level below 0.05. In consequence, it could be defined as a mathematical model that interrelates the independent variable with all seven dependent ones. For this objective, the response surface procedure was selected due to the useful curve fitting of multivariable models. Based on this procedure, a model with a determination factor of 54.89% could be considered as adequate due to its proceeds from real sample data in 163 case studies. The model obtained is shown in Equation (1):

$$
\begin{aligned}
TA = {}& 1085 - 62 \cdot ST - 0.75 \cdot CY + 57.8 \cdot WF - 105.9 \cdot SC - 67.7 \cdot V - 86.7 \cdot N - 14.1 \cdot L + 0.42 \cdot GT + \\
& + 18.3 \cdot NC + 0.2 \cdot MNC + 0.279 \cdot ST^2 + 0.000109 \cdot CY^2 - 0.232 \cdot WF^2 + 0.156 \cdot SC^2 + 0.353 \cdot V^2 - \\
& - 0.00479 \cdot L^2 - 0.0464 \ \cdot NC \cdot NC + 0.078 \cdot MNC^2 + 0.0300 \cdot ST \cdot CY + 1.36 \cdot ST \cdot WF - 1.96 \cdot ST \cdot SC + \\
& + 1.11 \cdot ST \cdot V - 0.89 \cdot ST \cdot N + 0.0411 \cdot ST \cdot L - 0.00110 \cdot ST \cdot GT - 0.165 \cdot ST \cdot NC - \\
& - 1.020 \cdot ST \cdot MNC - 0.0295 \cdot CY \cdot WF + 0.0545 \cdot CY \cdot SC + 0.0308 \cdot CY \cdot V + 0.0420 \cdot CY \cdot N + \\
& + 0.00730 \ \cdot CY \cdot L - 0.000222 \cdot CY \cdot GT - 0.0086 \cdot CY \cdot NC + 0.0001 \cdot CY \cdot MNC + 0.019 \cdot WF \cdot SC - \\
& - 0.377 \cdot WF \cdot V + 0.013 \cdot WF \cdot N + 0.0416 \cdot WF \cdot L - 0.00084 \ \cdot WF \cdot GT - 0.120 \cdot WF \cdot NC + \\
& + 0.047 \cdot WF \cdot MNC + 0.213 \cdot SC \cdot V + 1.178 \cdot SC \cdot N - 0.0324 \cdot SC \cdot L + 0.00133 \cdot SC \cdot GT - \\
& - 0.068 \cdot SC \ \cdot NC - \ 0.153 \cdot SC \ \cdot MNC + 1.110 \cdot V \cdot N - 0.100 \cdot V \cdot L + 0.00430 \cdot V \cdot GT + \\
& + 0.042 \cdot V \cdot NC + 0.474 \cdot V \cdot MNC - 0.144 \cdot N \cdot L + 0.00232 \cdot N \cdot GT + 0.009 \cdot N \cdot NC - \\
& - 0.061 \cdot N \cdot MNC + 0.000093 \cdot L \cdot GT + 0.0326 \cdot L \cdot NC - 0.0239 \cdot L \cdot MNC - \\
& - 0.000102 \cdot GT \cdot NC + 0.00170 \cdot GT \cdot MNC - 0.072 \cdot NC \cdot MNC
\end{aligned}
\tag{1}
$$

If we try to define a model that incorporates the human factor, a better equation is obtained with a higher determination factor of 62.51, as we can see in Equation (2):

$$
\begin{aligned}
TA = {}& 2713 + 14 \cdot ST - 2.55 \cdot CY + 59.9 \cdot WF - 112.4 \cdot SC - 102.7 \cdot V - 8.7 \cdot N + 3.90 \cdot L - 0.204 \cdot GT - \\
& - 7.3 \cdot CA - 13.9 \cdot MNC + 0.438 \cdot ST^2 + 0.000601 \cdot CY^2 - 0.150 \cdot WF^2 + 0.303 \cdot SC^2 - \\
& - 0.202 \cdot V^2 - 0.000366 \cdot L^2 + 2.111 \cdot CA^2 + 0.023 \cdot CMN^2 - 0.0087 \cdot ST \cdot CY + 0.88 \cdot ST \cdot WF - \\
& - 1.35 \cdot ST \cdot SC + 1.41 \cdot ST \cdot V - 0.81 \cdot ST \cdot N + 0.0416 \cdot ST \cdot L - 0.001512 \cdot ST \cdot GT - \\
& - 2.33 \cdot ST \cdot CA - 0.869 \cdot ST \cdot MNC - 0.0306 \cdot CY \cdot WF + 0.0581 \cdot CY \cdot SC + 0.0496 \cdot CY \cdot V + \\
& + 0.0021 \cdot CY \cdot N - 0.00165 \cdot CY \cdot L + 0.000090 \cdot CY \cdot GT + 0.0012 \cdot CY \cdot CA + \\
& + 0.0081 \cdot CY \cdot MNC - 0.236 \cdot WF \cdot SC - 0.422 \cdot WF \cdot V + 0.255 \cdot WF \cdot N - 0.0087 \cdot WF \cdot L + \\
& + 0.00062 \cdot WF \cdot GT + 0.862 \cdot WF \cdot CA + 0.298 \cdot WF \cdot MNC + 0.303 \cdot SC \cdot V + 0.596 \cdot SC \cdot N - \\
& - 0.0452 \cdot SC \cdot L + 0.00052 \cdot SC \cdot GT - 1.073 \cdot SC \cdot CA - 0.225 \cdot SC \cdot MNC + \\
& + 0.869 \cdot V \cdot N - 0.0654 \cdot V \cdot L + 0.00335 \cdot V \cdot GT + 0.284 \cdot V \cdot CA + 0.025 \cdot V \cdot MNC - \\
& - 0.058 \cdot N \cdot L + 0.00141 \ \cdot N \cdot GT + 0.829 \cdot N \cdot CA + 0.340 \cdot N \cdot MNC + 0.000057 \cdot L \cdot GT + \\
& + 0.0456 \cdot L \cdot CA - 0.1173 \cdot L \cdot MNC + 0.00040 \cdot GT \cdot CA + 0.00249 \cdot GT \cdot MNC - 0.380 \cdot CA \cdot MNC
\end{aligned}
\tag{2}
$$

In this sense, human factor allows an increase in the accuracy of the developed model and to differentiate the approach accidents to the others. Despite this, results indicate that the accuracy and, in consequence, its human factor scale, must be improved, therefore changing the simple scale usually employed in most of the CIAIM reports to a more in-depth scale allows us to improve the accident classification.

Finally, it is interesting to highlight that these models include aspects like wind forces, among other variables, as an improvement of the limitations defined in previous works [12]. It is true that, like in other studies [14], the determination factor must be improved in future analysis and it must be done paying special attention to the influence of the human factor

in some accidents, as it was concluded previously in agreement with the more recent research lines [29,30].

As an example of future study, a final model was developed. In this sense, an analysis of the more representative variable (ship length) showed that it is in fishing vessels (with a length below 40 m) where most of the values are centered. In consequence, a decrease of model definition is concluded at higher ship lengths due to a lack of a representative number of accidents of ships in this ship length range.

If we now simplify this data analysis to the most common vessel (fishing are 96 cases of 163 analyzed) just excluding two ambiguous accidents (sinking and operational accident), the model increases its precision to a determination factor of 0.83, see Equation (3).

$$
\begin{aligned}
TA = {}& -2727 + 2.99 \cdot CY + 33.1 \cdot WF - 192.2 \cdot SC - 102 \cdot V + 33.4 \cdot N + 18.2 \cdot L - 1.86 \cdot GT + 40.8 \cdot CA - 80.2 \\
& \cdot MNC - 0.000812 \cdot CY^2 - 0.296 \cdot WF^2 - 0.204 \cdot SC^2 - 0.226 \cdot V^2 - 0.0100 \cdot L^2 - 0.000273 \\
& \cdot GT^2 + 2.59 \cdot CA^2 + 0.047 \cdot MCN^2 - 0.0169 \cdot CY \cdot WF + 0.0978 \cdot CY \cdot SC + 0.0516 \cdot CY \cdot V \\
& - 0.0194 \cdot CY \cdot N - 0.00922 \cdot CY \cdot L + 0.000950 \cdot CY \cdot GT - 0.0275 \cdot CY \cdot CA + 0.0425 \cdot CY \\
& \cdot MNC + 0.443 \cdot WF \cdot SC + 0.254 \cdot WF \cdot V - 1.003 \cdot WF \cdot N - 0.076 \cdot WF \cdot L + 0.0044 \cdot WF \cdot GT \\
& + 1.044 \cdot WF \cdot CA + 1.076 \cdot WF \cdot MCN - 1.065 \cdot SC \cdot V + 1.385 \cdot SC \cdot N + 0.241 \cdot SC \cdot L \\
& - 0.0175 \cdot SC \cdot GT - 2.044 \cdot SC \cdot CA - 1.135 \cdot SC \cdot MCN + 0.295 \cdot V \cdot N + 0.152 \cdot V \cdot L \\
& - 0.0225 \cdot V \cdot GT + 0.945 \cdot V \cdot CA - 0.737 \cdot V \cdot MNC + 0.168 \cdot N \cdot L - 0.0420 \cdot N + 2.084 \cdot N \\
& \cdot CA - 0.119 \cdot N \cdot MCN + 0.00651 \cdot L \cdot GT - 0.180 \cdot L \cdot CA - 0.296 \cdot L \cdot MCN + 0.0156 \cdot GT \\
& \cdot CA + 0.0157 \cdot GT \cdot MCN + 0.13 \cdot CA \cdot MCN
\end{aligned}
\tag{3}
$$

Finally, excluding operational accidents, a determination factor of 0.90 is obtained with 76 accidents analyzed, which are half of the sampled data, see Equation (4).

$$
\begin{aligned}
TA = {}& -7088 + 7.44 \cdot CY - 122 \cdot WF + 186 \cdot SC - 138 \cdot V - 97 \cdot N + 9.8 \cdot L + 2.62 \cdot GT - 138 \cdot CA + 42.7 \\
& \cdot MCN - 0.00194 \cdot CY^2 + 0.501 \cdot WF^2 + 0.891 \cdot SC^2 - 0.229 \cdot V^2 + 0.0423 \cdot L^2 + 0.000712 \\
& \cdot GT^2 - 0.135 \cdot MCN^2 + 0.0582 \cdot CY \cdot WF - 0.092 \cdot CY \cdot SC + 0.070 \cdot CY \cdot V + 0.0474 \cdot CY \cdot N \\
& - 0.0071 \cdot CY \cdot L - 0.00093 \cdot CY \cdot GT + 0.0668 \cdot CY \cdot CA - 0.0180 \cdot CY \cdot MCN - 0.71 \cdot WF \cdot SC \\
& + 0.479 \cdot WF \cdot V - 1.86 \cdot WF \cdot N + 0.187 \cdot WF \cdot L - 0.0432 \cdot WF \cdot GT + 0.71 \cdot WF \cdot CA + 1.858 \\
& \cdot WF \cdot MCN - 1.15 \cdot SC \cdot V + 1.91 \cdot SC \cdot N + 0.285 \cdot SC \cdot L - 0.0005 \cdot SC \cdot GT - 3.14 \cdot SC \cdot CA \\
& - 1.61 \cdot SC \cdot MCN - 1.36 \cdot V \cdot N + 0.467 \cdot V \cdot L - 0.0760 \cdot V \cdot GT + 1.07 \cdot V \cdot CA - 1.134 \cdot V \\
& \cdot MCN + 0.417 \cdot N \cdot L - 0.1075 \cdot N \cdot GT + 5.80 \cdot N \cdot CA - 0.86 \cdot N \cdot MCN - 0.0107 \cdot L \cdot GT \\
& - 0.038 \cdot L \cdot CA + 0.032 \cdot L \cdot MCN - 0.0303 \cdot GT \cdot CA - 0.0156 \cdot GT \cdot MCN - 2.53 \cdot CA \cdot MCN
\end{aligned}
\tag{4}
$$

From these last models, it can be concluded that an adequate scale in accident reports and an adequate variables selection may allow researchers to define highly accurate models. What is more, the minimum number of ships that suffer an accident must be defined to prevent modelling accuracy problems. These are really interesting indications for future research works about ship accidents.

3.4.2. Relation between Variables and Scales Employed

In the present section, different models to relate implied variables are developed and shown in Table 6 with the aim to simplify the previous accidents models of Equations (1)–(4). For instance, from Table 1, it was obtained that sea conditions are related to the wind force and now a model with a determination factor of 60% was obtained. It is in accordance with the previous relation between Beaufort and Douglas scales shown in Table A8 (Appendix A) and is a good proof of an adequate variable analysis and modelling. Another interesting model is that which relates the crew member and the ship dimensions with a determination factor of 79.15, in agreement with common sense and international standards requirements.

Despite this, results indicate an inadequate model to relate the construction year and the minimum crew members to identify types of accidents. As it was commented in previous sections, the type of accident is, sometimes, a little ambiguous and the same happens with the minimum number of crew members (just meets (1) or does not meet (2)). Due to the information from CIAIM reports, a better understanding of these accidents and an improvement of these scales must be developed in the future.

Table 6. Models that relate main variables.

Related Variables	Model	Determination FACTOR (%)
SC and WF	$SC = 0.390 + 0.604\ WF + 0.0021\ WF^2$	59.12
GT, L, B and CM	$CM = -0.32 + 0.180\ L + 0.536\ B + 0.00009\ GT - 0.00353\ L^2 - 0.101\ B^2 + 0.0327\ L\ B + 0.000007\ L\ GT + 0.000127\ B\ GT$	79.15
TA, MNC	$TA = 7.86 - 2.07\ MNC + 0.135\ MNC^2$	21.18
TA, MNC, CY	$TA = 254 + 11.6\ MNC - 0.31\ CY + 0.185\ MNC^2 + + 0.000092\ CY^2 - 0.0070\ MNC\ CY$	28.00
TA, CY	$TA = 454 - 0.50\ CY + 0.000136\ CY^2$	4.99

4. Conclusions

In the present work, a case study about ship accidents in the Spanish SAR territory was conducted due to the lack of similar research works about this in this region. Despite the fact that a similar analysis was employed as that in previous case studies [16,32,33], in our case, it was done with more than two related variables. In consequence, the VCH square method was replaced with the Analysis of Variance, allowing for the definition of the relation between more numbers of variables for the particular Spanish Search and Rescue (SAR) sea regions. What is more, the importance of each of the three more relevant variables for each type of accident was defined by the Bayesian networks procedure. Finally, different deterministic models based on the variables identified were developed to show the feasibility of a deterministic model in a dynamic process. In consequence, a more in-depth analysis than before was done showing the sensibility of each implied variable to determine the more probable expected type of accident and the relevance of each variable to each type of accident. The main conclusions can be defined as:

1. The most relevant variables to identify the more probable type of accident and their average values were identified. In this sense, weather conditions do not have any kind of effect over analyzed accidents, in disagreement with previous works. Wind direction does not influence the type of accident or other parameters, and wind intensity is of interest and related to the sea condition. This effect is a consequence of the particularities of the region of study. The ship length and the year of construction were identified as other important parameters, which exert a clear effect over ships accidents (except some of type a2). This must be related to the standards of each country allowing the modification of the working life of ships higher than 20 years in this region.
2. It is possible to define different multinomial models that let us predict the type of accident in a deterministic way, that were developed based on the most relevant variables employed in accidents reports.
3. Improvements in accident reports and its scales are proposed, and common classificatory limits to define ship age, ship length, and crew members must be employed to compare results of different research works. What is more, a minimum number of accidents analyzed per each ship type must be defined.

Finally, it is of interest to highlight that this initial study will be the base case to predict the occurrence of accidents when the main variables of the problem change. In this sense, future works based on neural networks for clustering and accident modelling and prediction must be done to define the characteristics of the safest ship.

Author Contributions: Conceptualization, C.M., D.V., J.M.P.-C., and J.A.O.; methodology, C.M., D.V. and J.A.O.; software, D.V., J.M.P.-C.; validation, C.M., D.V., J.M.P.-C., and J.A.O.; formal analysis, C.M., J.M.P.-C., D.V., and J.A.O.; investigation, C.M., J.M.P.-C., D.V., and J.A.O.; resources, C.M., J.M.P.-C., and J.A.O.; data curation, D.V. and J.A.O.; writing—original draft preparation, C.M., D.V., J.M.P.-C., and J.A.O.; writing—review and editing, C.M., J.M.P.-C., D.V., and J.A.O.; visualization, C.M., J.M.P.-C., D.V., and J.A.O.; supervision, C.M., J.M.P.-C., D.V., and J.A.O.; project administration, C.M., D.V., J.M.P.-C., and J.A.O.; funding acquisition, C.M., D.V., J.M.P.-C., and J.A.O. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

a1	Contact
a10	Operational accident
a2	Flooding
a3	Grounding
a4	Collision
a5	Fire
a6	Capsizing
a7	List
a8	Stranded
a9	Loss of Stability
AI	Artificial Intelligence
B	Breadth
BBN	Bayesian Belief Network
CA	Cause of accidents (Human factor)
CIAIM	Spanish Commission for Investigation of Maritime Accidents and Incidents
CM	Crewmembers
COLREG	Convention on the International Regulations for Preventing Collisions at Sea
CY	Year of build
DAMA	A coding system agreed in 1990 by the Scandinavian countries for the registration and analysis of marine accidents.
FRAM	Functional Resonance Accident Model
GRT	Gross Register Tonnage
GT	Gross Tonnage
HFACS	Human Factor Analysis and Classification System
IA	Intelligence Artificial
IALA	International Association of Marine Aids to Navigation and Lighthouse Authorities
IMO	International Maritime Organization
IWRAP	IALA Waterway Risk Assessment Program
L	Length
MNC	Minimum number of crewmembers
N	Nocturnal
PAWSA	Ports and Waterway Safety Assessment
SAR	Search and Rescue
SC	Sea condition
SEM	Structural Equation Modelling
SPSS	Statistical Package for the Social Sciences
ST	Type of ship
TA	Type of accident
V	Visibility
WD	Wind direction
WF	Wind force

Appendix A. Variables and Its Codes

Table A1. Ship type code.

Code	Ship Type
1	Extraction Platforms and Auxiliary Supply vessels
2	Passenger and/or cargo vessels
3	Fishing vessels
4	Fishing and Aquaculture Auxiliary Boats
5	Tugboats and auxiliary boats of the Port Service
6	Commercial Yachts
7	Recreational Boats
8	Vessels of Public Organizations (National or Autonomous Communities)
9	Provisional List of Vessels under Construction

Table A2. Type of accident code.

Code	Type of Accident
1	Contact
2	Flooding
3	Grounding
4	Collision
5	Fire
6	Capsizing
7	List
8	Stranded
9	Loss of Stability
10	Operational Accident

Table A3. Code of minimum crew members' requirement.

Code	Denomination	
	One ship involved	Two ships involved
1	Meets	-
2	Do not meet	-
3	-	One does not meet
4	-	Two do not meet
5	-	Two meet

Table A4. Code of visibility level.

Code	Visibility Level
1	Very Bad (less than 1000 m)
2	Bad (between 1000 m and 2 nautical miles)
3	Moderate (between 2 nautical miles and 5 nautical miles)
4	Good (more than 5 nautical miles)

Table A5. Code of time period in which the accident occurs.

Code	Period of Time
1	Day
2	Night

Table A6. Code of human factor influence.

Code	Human Factor Conditions
1	Human factor influences the accident
2	Human factor does not influence the accident
3	Doubtful

Table A7. Code of wind velocity in accordance with the Beaufort scale.

Wind Force (Code)	Denomination	Velocity	
		(km/h)	Knots
0	Calm	0–2	Till 1
1	Light air	2–6	1–3
2	Light breeze	7–11	4–6
3	Gentle breeze	12–19	7–10
4	Moderate breeze	20–29	11–16
5	Fresh breeze	30–39	17–21
6	Strong breeze	40–50	22–27
7	High wind	51–61	28–33
8	Gale	62–74	24–40
9	Strong gale	75–87	41–47
10	Storm	88–101	48–55
11	Violent storm	102–107	56–63
12	Hurricane force	>118	>64

Table A8. Code of sea conditions in accordance with the Douglas scale.

Sea Conditions (Code)	Douglas Scale	Wave Height		Beaufort Scale
		(m)	(ft)	
0	Calm (Glassy)	No wave		0
1	Calm (rippled)	0–0.10	0.00–0.33	1–2
2	Smooth	0.10–0.50	0.33–1.64	3
3	Slight	0.50–1.25	1.6–4.1	4
4	Moderate	1.25–2.50	4.1–8.2	5
5	Rough	2.50–4.00	8.2–13.1	6
6	Very rough	4.00–6.00	13.1–19.7	7
7	High	6.00–9.00	19.7–29.5	8–9
8	Very high	9.00–14.00	29.5–45.9	10–11
9	Phenomenal	14.00+	45.9+	12

Appendix B. Results

Table A9. Identification of the accidents that can be identified as different to each other (columns) employing Ship type as dependent variable.

	a1	a2	a3	a4	a5	a6	a8	a10
a1		0.187	0.024	0.316	0.477	0.293	0.413	0.450
a2	0.000		0.871	0.000	0.000	0.206	0.088	0.319
a3	0.826	0.595		0.731	0.780	0.880	0.731	0.731
a4	0.319	0.030	1.000		0.177	0.074	0.092	0.062
a5	0.055	0.001	0.918	0.016		0.013	0.000	0.013
a6	0.018	0.000	0.838	0.023	0.009		0.000	0.006
a8	0.000	0.000	0.762	0.000	0.000	0.000		0.004
a10	0.000	0.001	0.789	0.001	0.000	0.017	0.002	

Table A10. Identification of the accidents that can be identified as different to each other (columns) employing year of construction as dependent variable.

	a1	a2	a3	a4	a5	a6	a8	a10
a1		0.212	0.000	0.000	0.000	0.000	0.000	0.000
a2	0.005		0.000	0.000	0.000	0.006	0.000	0.000
a3	0.021	0.485		0.000	0.056	0.021	0.021	
a4	0.298	0.139	0.000		0.020	0.298	0.037	0.139
a5	0.001	0.069	0.000	0.000		0.001	0.000	0.003
a6	0.006	0.055	0.000	0.000	0.000		0.000	0.000
a8	0.000	0.028	0.000	0.000	0.000	0.000		0.000
a10	0.000	0.000	0.000	0.000	0.002	0.000	0.000	

Table A11. Identification of the accidents that can be identified as different to each other (columns) employing sea conditions as dependent variable.

	a1	a2	a3	a4	a5	a6	a8	a10
a1		0.666	0.646	0.613	0.032	0.066	0.235	0.58
a2	0.469		0.731	0.972	0.066	0.865	0.835	0.85
a3	0.441	0.299		0.736	0.736	0.152	0.736	0.590
a4	0.643	0.301	0.911		0.645	0.613	0.015	0.887
a5	0.494	0.499	0.911	0.022		0.162	0.240	0.452
a6	0.007	0.593	0.586	0.081	0.136		0.932	0.454
a8	0.687	0.289	0.911	0.204	0.917	0.283		0.097
a10	0.921	0.891	0.646	0.906	0.354	0.528	0.048	

Table A12. Identification of the accidents that can be identified as different to each other (columns) employing night or day period as dependent variable.

	a1	a2	a3	a4	a5	a6	a8	a10
a1		0.832	0.49	0.075	0.446	0.086	0.399	0.230
a2	0.832		0.218	0.581	0.632	0.606	0.596	0.192
a3	0.211	1		0.160	0.063	0.160	0.063	0.141
a4	0.759	0.283	0.027		0.005	0.007	0.005	0.007
a5	0.483	0.350	0.004	0.000		0.000	0.000	0.000
a6	0.089	0.443	0.001	0.000	0.000		0.000	0.000
a8	0.482	0.463	0.000	0.000	0.000	0.000		0.000
a10	0.187	0.638	0.000	0.000	0.000	0.000	0.000	

Table A13. Identification of the accidents that can be identified as different to each other (columns) employing visibility as dependent variable.

	a1	a2	a3	a4	a5	a6	a8	a10
a1		0.299		0.645	0.63	0.677	0.677	0.533
a2	0.859			0.449	0.459	0.881	0.779	0.343
a3				0.000		0.000	0.000	0.000
a4	0.645	0.673				0.000	0.758	0.796
a5	0.866	0.882	0.000			0.054	0.886	0.790
a6	0.683	0.803		0.752	0.275		0.108	0.211
a8	0.733	0.772		0.324	0.875	0.040		0.500
a10	0.109	0.777		0.515	0.515	0.351	0.738	

Table A14. Identification of the accidents that can be identified as different to each other (columns) employing length as dependent variable.

	a1	a2	a3	a4	a5	a6	a8	a10
a1		0.424	0.000	0.000	0.000	0.900	0.000	0.000
a2	0.000		0.000	0.000	0.000	0.034	0.113	0.348
a3	0.000	0.000		0.000	0.000	0.930	0.000	0.000
a4	0.000	0.000	0.000		0.000	0.178	0.000	0.000
a5	0.000	0.160	0.000	0.000		0.377	0.000	0.000
a6	0.000	0.247	0.000	0.000	0.000		0.000	0.000
a8	0.000	0.106	0.000	0.000	0.000	0.345		0.352
a10	0.000	0.144	0.000	0.000	0.000	0.680	0.142	

Table A15. Identification of the accidents that can be identified as different to each other (columns) employing minimum crew members as dependent variable.

	a1	a2	a3	a4	a5	a6	a8	a10
a1		0.025	0.052	0.18	0.003	0.003	0.002	0.001
a2	0.031		0.314	0.009	0.226	0.227	0.401	0.346
a3	0.144	0.257		0.041	0.46	0.041	0.009	0.009
a4	0.001	0.002	0.26		0.001	0.001	0.001	0.000
a5	0.000	0.094	0.004	0.000		0.000	0.000	0.000
a6	0.000	0.091	0.004	0.000	0.000		0.000	0.000
a8	0.000	0.182	0.000	0.000	0.000	0.000		0.000
a10	0.000	0.152	0.000	0.000	0.000	0.000	0.000	

Table A16. Identification of the accidents that can be identified as different to each other (columns) employing human factor as dependent variable.

	a1	a2	a3	a4	a5	a6	a8	a10
a1		0.000	0.000	0.000	0.000		0.000	0.000
a2			0.157	0.497	0.109	1.000	0.318	0.396
a3		0.354		0.633	0.650			0.818
a4		0.528	0.633		0.477			0.285
a5		0.253	1.000	0.939			0.287	0.630
a6		0.259	0.000	0.000	0.000		0.742	0.768
a8		0.462	0.000	0.000	0.522	0.742		0.467
a10		0.446	0.506	0.939	0.630	0.500	0.365	

References

1. Awal, Z.I.; Kazuhiko, H. A Study on Accident Theories and Application to Maritime Accidents. *Procedia Eng.* **2017**, *194*, 298–306. [CrossRef]
2. Perrow, C. *Normal Accidents: Living with High-Risk Technologies*; Princeton University Press: New York, NY, USA, 1984; pp. 1–464.
3. Allianz. Safety and Shipping 1912–2012 from Titanic to Costa Concordia. 2012. Available online: https://www.agcs.allianz.com/content/dam/onemarketing/agcs/agcs/reports/AGCS-Safety-Shipping-Review-2012.pdf (accessed on 23 September 2020).
4. AISM—IALA. *IALA Recommendation O-134. on the IALA Risk Management Tool for Ports and Restricted Waterways*, 2nd ed.; AISM—IALA: Saint Germain in Laye, France, 2009. Available online: https://www.iala-aism.org/content/uploads/2016/07/o_134_ed2_iala_risk_management_tool_for_ports_and_restricted_waterways_may2009.pdf (accessed on 14 September 2020).
5. Friis-Hansen, P. IWRAP MKII Working Document: Basic Modelling Principles for Prediction of Collision and Grounding Frequencies. Technical University of Denmark. 2008. Available online: https://www.iala-aism.org/wiki/iwrap/images/2/2b/IWRAP_Theory.pdf (accessed on 13 September 2020).
6. Fujii, Y.; Yamanouchi, H.; Mizuki, N. Some Factors Affecting the Frequency of Accidents in Marine Traffic. II: The probability of Stranding, III: The Effect of Darkness on the Probability of Stranding. *J. Navig.* **1974**, *27*, 235–252.
7. Hollnagel, E.; Goteman, Ö. The Functional Resonance Accident Model. 2004, pp. 155–161. Available online: https://pdfs.semanticscholar.org/a538/fc4f7c5ebd7c083aa8474786f6886300251f.pdf?_ga=2.92429917.1335224137.1599928060-2025622813.1598193688. (accessed on 10 September 2020).
8. Hanzu-Pazara, R.; Barsan, E.; Arsenie, P.; Chiotoriou, L.; Raicu, G. Reducing of maritime accidents caused by human factors using simulators in training process. *J. Marit. Res.* **2008**, *5*, 3–18.

9.   Youn, I.-H.; Park, D.-J.; Yim, J.-B. Analysis of lookout activity in a simulated environment to investigate maritime accidents caused by human error. *Appl. Sci.* **2019**, *9*, 4. [CrossRef]

10. Liu, Z.; Zhang, Y.; Yu, X.; Yuan, C. Unmanned surface vehicles: An Overview of Developments and Challenges. *Annu. Rev. Control.* **2016**, *41*, 71–93. [CrossRef]

11. Statheros, T. Autonomous Ship Collision Avoidance Navigation Concepts, Technologies and Techniques. *J. Navig.* **2008**, *61*, 129–142. [CrossRef]

12. Awal, Z.I.; Islam, M.R.; Baree, M.S. Development of a mathematical model for analysis on ship collision dynamics. In Proceeding of the International Conference on Mechanical Engineering, Dhaka, Bangladesh, 29–31 December 2007.

13. Liu, C.P.; Ching-Wu, C.; Gin-Shuh, L.; Su, Y. Establishing marine accident classification: A case study in Taiwan. *J. East. Asia Soc. Transp. Stud.* **2005**, *6*, 952–967. [CrossRef]

14. Ulusçu, O.S.; Özbaç, B.; Altiok, T.; Or, I. Risk analysis of the vessel traffic in the Strait of Istanbul. *Risk Anal.* **2009**, *29*, 1454–1472. [CrossRef]

15. Jensen, O.C.; Petursdottir, G.; Holmen, I.M.; Abrahamsen, A.; Lincoln, J. A review of fatal accident incidence rate trends in fishing. *Int. Marit. Health* **2014**, *65*, 47–52. [CrossRef]

16. Håvold, J.I. Safety culture aboard fishing vessels. *Saf. Sci.* **2010**, *48*, 1054–1061. [CrossRef]

17. Alvite-Castro, J.; Orosa, J.A.; Vergara, D.; Costa, A.M.; Bouzón, R. A new design criterion to improve the intact stability of Galician small fishing vessels. *J. Mar. Sci. Eng.* **2020**, *8*, 499. [CrossRef]

18. Ozguc, O. Structural damage of ship-FPSO collisions. *J. Mar. Eng. Technol.* **2019**, *18*, 1–35. [CrossRef]

19. Wang, J.; Pillay, A.; Kwon, Y.; Wall, A.; Loughran, C. An analysis of fishing vessel accidents. *Accid. Anal. Prev.* **2005**, *37*, 1019–1024. [CrossRef] [PubMed]

20. Laursen, L.H.; Hansen, H.L.; Jensen, O.C. Fatal occupational accidents in Danish fishing vessels 1989–2005. *Int. J. Inj. Control Saf. Promot.* **2008**, *15*, 109–117. [CrossRef] [PubMed]

21. Jin, D.; Kite-Powell, H.L.; Thunberg, E.; Solow, A.R.; Talley, W.K. A model of fishing vessel accident probability. *J. Saf. Res.* **2002**, *33*, 497–510. [CrossRef]

22. Davis, B.; Colbourne, B.; Molyneux, D. Analysis of fishing vessel capsizing causes and links to operator stability training. *Saf. Sci.* **2019**, *118*, 355–363. [CrossRef]

23. Pitman, S.J.; Wright, M.; Hocken, R. An analysis of lifejacket wear, environmental factors, and casualty activity on marine accident fatality rates. *Saf. Sci.* **2019**, *111*, 234–242. [CrossRef]

24. Jin, D.; Thunberg, E. An analysis of fishing vessel accidents in fishing areas off the northeastern United States. *Saf. Sci.* **2005**, *43*, 523–540. [CrossRef]

25. Jin, D. The determinants of fishing vessel accident severity. *Accid. Anal. Prev.* **2014**, *66*, 1–7. [CrossRef]

26. Talley, W.K.; Jin, D.; Kite-Powell, H. Determinants of the damage costand injury severity of ferry vessel accidents. *WMU J. Marit. Aff.* **2008**, *7*, 175–188. [CrossRef]

27. Mullai, A.; Pulasson, U. A grounded theory model for analysis of marine accidents. *Accid. Anal. Prev.* **2011**, *43*, 1590–1603. [CrossRef] [PubMed]

28. Chen, D.; Pei, Y.; Xia, Q. Research on human factors cause chain of ship accidents based on multidimensional association rules. *Ocean Eng.* **2020**, *218*, 107717. [CrossRef]

29. Batalden, B.M.; Sydnes, A.K. What causes 'very serious' maritime accidents? *Saf. Reliab. Theory Appl.* **2017**, 3067–3074. [CrossRef]

30. Hu, S.; Li, Z.; Xi, Y.; Gu, X.; Zhang, X. Path Analysis of Causal Factors Influencing Marine Traffic Accident via Structural Equation Numerical Modeling. *J. Mar. Sci. Eng.* **2019**, *7*, 96. [CrossRef]

31. Azaheri, A.; Montewka, J.; Kujala, P. Towards and evidence-based probabilistic risk model for ship-grounding accidents. *Saf. Sci.* **2016**, *86*, 195–210. [CrossRef]

32. Uğurlu, F.; Yıldız, S.; Boran, M.; Uğurlu, Ö.; Wang, J. Analysis of fishing vessel accidents with Bayesian network and Chi-square methods. *Ocean Eng.* **2020**, *198*, 106956. [CrossRef]

33. Spanish Commission for Investigation of Maritime Accidents and Incidents, CIAIM. Available online: https://www.mitma.gob.es/organos-colegiados/ciaim (accessed on 20 August 2020).

34. SPSS v.22. Available online: https://www.ibm.com/es-es/analytics/spss-statistics-software. (accessed on 26 January 2021).

35. Minitab v18. Available online: https://www.minitab.com/es-mx/. (accessed on 5 September 2020).

36. Mansour Youssef, S.A.; Paik, J.K. Hazard identification and scenario selection of ship grounding accidents. *Ocean Eng.* **2018**, *153*, 242–255. [CrossRef]

37. Chen, P.; Mou, J.; van Gelder, P.H.A.J.M. Integration of individual encounter information into causation probability modelling of ship collision accidents. *Saf. Sci.* **2019**, *120*, 636–651. [CrossRef]

38. Chen, P.; Huang, Y.; Mou, J.; van Gelder, P.H.A.J.M. Probabilistic risk analysis for ship-ship collision: State-of-the-art. *Saf. Sci.* **2019**, *117*, 108–122. [CrossRef]