http://dx.doi.org/10.21622/MARLOG.2022.11.010
Vol. 11, Iss. 1, August 2022

Investigating Maritime Accidents That Involve Dangerous Goods Using Hierarchical Clustering

Serra Patrizia [(1)i], Fancello Gianfranco [(2)], Mandas Marco [(3)], Daga Mariangela [(4)], Medda Andrea [(3)]

[(1,2, 4, 5)] DICAAR – Department of Civil and Environmental Engineering and Architecture, University of Cagliari, Cagliari, Italy,

[(3)] Department of Economics and Business, University of Cagliari, Cagliari, Italy,

E-Mail: pserra@unica.it, fancello@unica.it, marco.mandas@unica.it, mariangela.daga@unica.it, a.medda@unica.it

1. ABSTRACT: Although the important progress in terms of safety and technological advances, maritime accidents remain a critical issue in merchant shipping. A high number of accidents continue to occur every year, with negative consequences both in economic and environmental terms (with often disastrous and lasting environmental impacts for marine ecosystems) and in the loss of human life. Understanding the maritime accidents phenomenon is expedient to giving shipping practitioners a focus for tailored interventions aimed at enhancing maritime safety. Using hierarchical clustering methods, this paper analyses historical data relating to maritime accidents to highlight the potential causal relationships that can describe homogeneous groups of accidents. The study explores a database consisting of 1,079 marine accidents that occurred worldwide in the 2009-2019 decade. Accident data is taken from the International Maritime Organization (IMO) database. After illustrating a description of the data set, a non-supervised hierarchical clustering analysis is applied to identify accident patterns, thus helping to better describe the phenomenon and identify potential causal relations that repeat in various accidents. A significant distinction emerges between the accidents that occur for technical reasons and those where human factors (stress, fatigue, situation awareness, decision-making, communication, etc.) play a prevalent role. Afterwards, the clustering analysis is applied to a sub-set of accidents (153 accidents) involving ships carrying dangerous goods (gases, oils, explosives, etc.). The results of the analysis point out the role of the human factor as the prevalent (or contributing) cause of the marine accidents related to work operations. Conversely, fires and explosions, which are by far the most frequent accidents involving ships carrying dangerous goods, are mainly caused by technical problems.

Keywords: ***Dangerous goods; Sea Accidents; Hierarchical Methods; Sea Safety***

2. INTRODUCTION

According to the International Maritime Organization (IMO), nearly one million seafarers work on about 60,000 large cargo vessels every day (IMO, 2021). Among the various risks associated with the maritime transport and storage of goods, there is the possibility of breakdowns or accidents during navigation, but also of overturning packages or leaking of material from containers, which can also cause loss of life

as well as damage to the environment and properties.
When the cargo involves dangerous goods, maritime transport becomes even more complex and delicate. Dangerous goods have in fact some specific physical and chemical properties (e.g., flammability, explosion, corrosion, toxicity, radioactivity, etc.) that require greater caution and compliance with strict rules and regulations in handling. The International Maritime Dangerous Goods Code (The IMDG Code, 2020), drawn up by the IMO, is the international regulation that provides the requirements for the safe transport of dangerous goods by sea. It contains the criteria for classifying goods, packaging methods, conditions for transporting them in bulk or in tanks, indications for marking packages and transport units, criteria for drawing up transport documentation (Multimodal Dangerous Goods Form), types of tanks and vehicles suitable for transport by sea, etc.

The great volume of goods transported by sea inevitably involves significant risk of accidents. Despite several measures introduced in the last decades to increase safety levels of maritime operations (e.g., new forms of team training or new regulations), marine accidents are still a major concern. The marine accidents can be caused by a variety of factors, which are often difficult to identify. These factors can be related to the ship itself or its equipment, but can also be linked to environmental issues, navigational and operational factors, traffic aspects or human factors. They can also be caused, and often are, by more than one of the above factors (Fadda et al., 2021).

The attention of the scientific literature towards the topic is significant and growing. Luo and Shin's (2019) have revised 572 articles published from 1965 to 2014 showing that the focus of the research on maritime accident has shifted from naval architecture to human error and may continue to expand into socio-economic factors. To date, several studies have investigated the role of human errors in marine accidents and provided recommendations to decrease the risk associated to specific causes (Antão and Guedes Soares, 2019). According to Galieriková (2019), human error is the main cause contributing to up to 70% of marine accidents. Lecue et al. (2019) studied accidents in European ports from 1919 to present to better understand their main characteristics and typology, and consequent risk to society and the environment. Karakavuz et al (2020) analyzed chemicals transported in the Port of Houston and assigned an overall risk score based on public safety and environmental health effects, quantity transported, congestion and visibility in the port at the time of an incident. This was to assist government authorities, first responders, port workers and emergency planners in their decision-making processes. Effective risk control is a significant way to prevent accidents and provide port security, especially as regards dangerous goods storage (Chu and Lyu, 2018).

Wang et al (2021) have explored the relationship between the severity of marine accidents and influencing factors using an ordered logistic regression model that reflects the relationship between these factors and the severity of marine accidents.

The position of the vessel at the time of the accident should also be considered as the consequences may also vary according to the area involved: coastal waters, high seas, in port, at mooring or being maneuvered, at anchor, etc. For instance, when talking about human casualties and/or accidents in ports, one must consider that there are various activities in the port that imply a high human presence and therefore higher risks for humans. Such operations can include loading and unloading of goods,

maneuvering operations in ports, oil jetty operations, shipbuilding activities, presence of fishing vessels, marina activities, dredging, construction of port infrastructure, etc.
In this study, a non-supervised hierarchical clustering analysis is applied to an IMO dataset consisting of 1,079 sea accidents that occurred worldwide between 2009 and 2019. Attention is also focused on a subset of accidents that involve ships carrying dangerous goods.
The main objectives of the study can be summarized as follows:

- identify the factors contributing to accidents at sea and compare them by homogeneous groups
- investigate the influence of human factors on maritime accidents.

The paper is organized as follows: following this introduction, Section 2 describes the data while

Section 3 the methodology. The quantitative results of the application are set out in Section 4 and discussed in Section 5. The main conclusions drawn from the analysis are in Section 6.

2. DATA

According to the severity level, marine accidents can be classified into four categories of descending severity level:

-very serious marine casualties (they involve total loss of the ship, loss of life, or severe pollution)

- serious marine casualties (they result in immobilization of main engines, extensive accommodation damage, severe structural damage or pollution)

- less serious casualties

- marine incidents (of minor relevance, they are not reported in the IMO database) In this study, only the first three types of marine accidents are considered.

Information on worldwide sea accidents used in this research has been derived from the IMO Marine Casualties and Incidents Database. The IMO database, defined as the Global Integrated Shipping Information System (GISIS), is available at: https://gisis.imo.org/. In particular, the analysis carried out in this research covers the decade 2009-2019. 4,347 ships were involved in a marine accident during the period under consideration, for a total of 3,710 reported accidents. Due to missing and uncomplete information, our statistical analysis is performed only considering the accidents that report complete information (1,079 accidents). Data cleaning operations were carried out to eliminate any redundancy and make the information in the database homogeneous and functional for the objectives of the study. This operation led to a reduction in the number of variables to be analyzed from 759 down to 70, with an increase in the degree of completeness of the database from 8% to 60%. Among the 70 variables considered, 17 relate to the vessel (IMO number, flag, length, tonnage, type of ship, type of service, etc.) and 53 to the accident (date and time, position, crew on board, initial event, etc.). The main data cleaning operations carried out were:

- aggregation of identical variables: when the same information is reported by more than one variable or when information related to one variable is reported by mistake inside the field related to another one

-aggregation of similar or infrequent variables: when several similar or infrequent variables are grouped into a single variable

- replacement of nominal qualitative variables by binary variables: when more than one mode of the same qualitative variable can be assigned to the same statistical unit

- conversion of information from several variables into a single variable: when the original database reports information from a single report item divided into several variables.

In a second step, the analysis was focused only on the accidents involving ships carrying dangerous goods. Since the original database did not contain the precise information about the type of goods transported by each vessel, this information was derived by assuming that the maritime transport of dangerous goods is mainly performed by tankers. Only the accidents involving this type of vessel where thus selected, including oil tankers, chemical tankers, gas tankers, other types of tankers. The resulting subset included 153 accidents and consisted of 447 records (more than one ship can be involved in the same accident), each characterized by the same 70 variables mentioned above.

3. METHODOLOGY

This study applies hierarchical clustering with the aim of identifying characteristic groups of marine incidents. This type of analysis allows grouping the incidents in such a way that those belonging to the same cluster have similar characteristics to each other and different from those of the other clusters.

Several clustering techniques are available in the literature (Rai and Singh, 2017), in this study a multiple correspondence analysis followed by hierarchical clustering is used.

Multiple correspondence analysis is a technique of factor analysis applicable to qualitative data (Abdi and Valentine, 2006) that allows to extract a series of factors that synthesize the information contained in the original data. This analysis makes it possible to work on a space of reduced dimensions representative of the original variables while obtaining the quantitative variables necessary to perform hierarchical clustering. We decide to keep 20 dimensions for performing the hierarchical clustering analysis. They capture 84% of the overall variability of the original data.

The analysis was carried out on a subset of the 70 variables. Some of them (Ship name, IMO number etc.) were excluded because not suitable for the purposes of this analysis. In the remaining cases, the selection process was a subjective assessment based on two criteria. The first criterion is how interesting we considered the variable for the purposes of the analysis. The second criterion relates to the percentage of completeness of the variable in the database. Since this methodology can be applied on a dataset with no missing values, the inclusion of a variable with many missing values can be costly as it reduces the total number of accidents to be included in the analysis. Based on these considerations, we obtained a complete dataset composed of the following 17 variables: Ship type, Type of casualty, Loss of life, Location of the casualty, Consequences to the ship, Pilot on board, Human errors, Human violations, Technical failure, Problem with ship's cargo, Adverse weather conditions, Navigational tools problem, Communication, Standards of personal competence or lack of training, Fatigue, stress, or excessive workload, Hardware issues, Software issues.

Hierarchical clustering was then applied to create a hierarchical decomposition of the data represented by the dendrogram which is cut at a certain height set by the researcher to define the number of clusters of interest. The length of the branches indicates how homogeneous the group is. The greater the length, the greater the homogeneity within the group.

4. RESULTS

The resulting dendrogram has been divided at the level where it decomposes into 16 branches corresponding to the same number of clusters. The decision to divide into 16 clusters was taken with the aim of selecting the smallest number of clusters with the highest homogeneity (Fig.1).

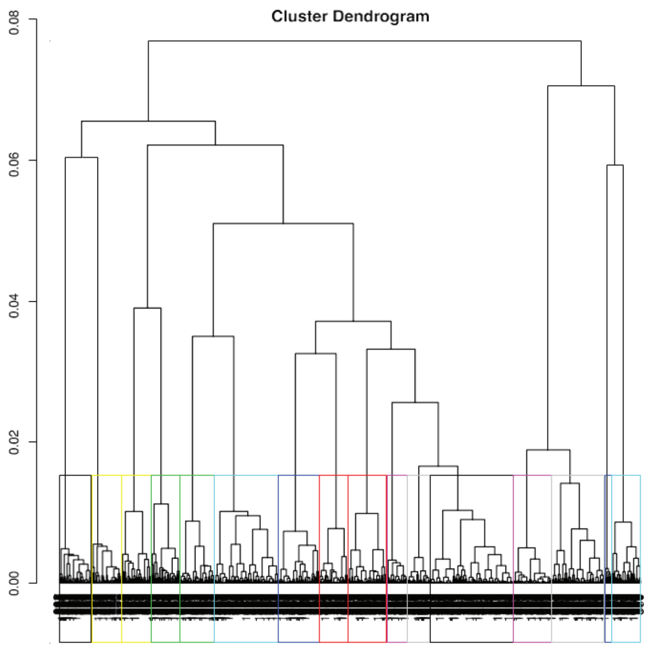

Figure 1: Dendogram – whole dataset (1,079 accidents)

Table 1 shows a description of the 16 clusters obtained. The first row shows the global composition of the variables while the remaining ones refer to each cluster and show the category of each variable significantly different from the global distribution that characterizes the most the clusters. In parenthesis, the first value is the cla/mod that is the proportion of all observations in the category belonging to the cluster and the second value is the mod/cla that is the proportion of all observations in the cluster belonging to the category.

The following list presents the types of maritime accidents identified by interpreting each cluster:

- Cluster 1: it groups the accidents where a machinery failure was registered, and the vessel was reported not able to continue navigation. In this cluster there are no registered problems related to human factors, but accidents are mostly caused by technical problems.
- Cluster 2: collisions and groundings of cargo vessels in territorial waters are grouped together. In most of these observations, the pilot was on board the vessel.
- Cluster 3: this cluster includes only accidents involving passenger vessels where no loss of life is reported. The most common type of accident in this cluster is fire and most of the vessels involved were not able to continue their navigation after the accident.

http://dx.doi.org/10.21622/MARLOG.2022.11.010

Vol. 11, Iss. 1,
August
2022

- Cluster 4: this group consists of stranded cargo ships. No loss of life is reported. Main causes include adverse weather conditions and, to a smaller part, the inappropriate choice of shipping route.
- Cluster 5: most of the accidents are related to fires caused by technical problems of the vessel. Most of the vessels involved are fishing vessels. No problems related to human factors are reported.
- Cluster 6: it groups accidents involving recreational vessels, which in most cases were involved in a collision. The accident usually took place in near-shore waters and the vessel was unable to continue navigation.
- Cluster 7: it includes accidents with damage to the vessel caused in most cases by a technical problem. In almost all observations the vessel was able to continue navigation.
- Cluster 8: it groups vessels involved in collisions. In most cases these accidents were recorded in open sea and loss of life was reported. No human factors issues were reported.
- Cluster 9: it is the largest cluster and includes collisions involving cargo ships. Most of these accidents were recorded in ports. Among the most common causes are adverse weather conditions and human errors related to lack of communication or errors in dealing with the correct operations.
- Cluster 10: all tankers are grouped together, and the most frequent accidents are related to work operations. In most cases the vessel was able to continue navigation. Communication problems and the presence of the pilot on board are reported.
- Cluster 11: it groups the incidents that report problems with protocols and/or working standards, as well as communication problems, thus relating to human factors in general. In most cases there were problems with on-board instrumentation. The most recorded incident was grounding.
- Cluster 12: it is the smallest cluster and includes all accidents involving water on board. Most of the vessels involved in this type of accident are fishing boats. The lack of adequate training of the maritime operators is reported.
- Cluster 13: it groups accidents involving capsizing. In many cases, the vessels involved are pleasure crafts and loss of life is reported. Problems related to the lack of adequate training of maritime personnel are highlighted.
- Cluster 14: it groups accidents involving work operations at sea. Loss of life is reported, but problems related to human factors are not specified.
- Cluster 15: it groups accidents related to problems with working protocols and standards. In addition, fatigue of seafarers and communication problems are reported as problems. A large proportion of the incidents were recorded in sub-coastal waters.
- Cluster 16: it groups accidents related to cargo loading operations. In most cases these accidents occurred in ports and resulted in loss of life. In this cluster, problems with working protocols and standards are evident, particularly in relation to poor communication and lack of adequate training of personnel.

Table 1: Clusters' size and features – Whole dataset (1,079 accidents)

	Size	Type of casualty	Ship type	Location	Loss of life	Consequences to the ship	Pilot on Board	Hardware factors	Software factors	Personnel factors	Human error	Human violations	Other causes
Overall	1151	Capsizing (5%), Collision (34%), Damage to the ship (5%), Fire (15%), Flooding (1%), Grounding (14%), Machinery Failure (6%), Work Accident (20%)	Cargo (57%), Fishing (12%), Passenger (10%), Special Craft (8%), Tanker (13%)	Coastal waters (25%), Inland waters (8%), Open Sea (27%), Port (40%)	Yes (32%) No (68%)	Ship remains fit to proceed (49%) Ship rendered unfit to proceed (28%) Total loss of the ship (23%)	Yes (19%) No (81%)	Yes (19%) No (81%)	Yes (40%) No (60%)	Communication (11%), Standards of personal competence (18%), Fatigue (8%), Other (11%), No (66%)	Error in judgement (29%), Failure to respond appropriately (12%), Incorrect operations of control (11%), Inappropriate choice of route (5%), Forgetting to report information (2%), Failure to advise officer on the watch (2%), Deciding not to pass on information (4%), Failure to report due to distraction (2%), Other Errors (24%), No (46%)	Necessary (6%), Routine (12%), Other (9%), No (75%)	Problem with cargo (10%), Technical failure (33%), Structural failure (8%), Adverse weather (29%), Navigational tool problems (6%), No (40%)
1	63	Machinery failure (98%, 100%)			No (7%, 91%)	Ship rendered unfit to proceed (9%, 44%)				No (8%, 91%)	No (10%, 79%)	No (7%, 89%)	Technical failure (13%, 78%), Structural failure (21%, 30%)
2	58	Collision (9%, 64%), Grounding (11%, 29%)	Cargo (8%, 89%)	Inland waters (60%, 100%)	No (7%, 93%)	Ship rendered unfit to proceed (7%, 40%)	Yes (19%, 71%)	No (6%, 93%)	No (7%, 80%)				
3	69	Fire (15%, 38%)	Passenger (62%, 100%)	Coastal waters (9%, 36%)	No (8%, 94%)	Ship rendered unfit to proceed (11%, 52%)					No (8%, 61%)		
4	80	Grounding (50%, 100%)	Cargo (9%, 75%)	Coastal waters (13%, 45%)	No (10%, 98%)	Ship rendered unfit to proceed (10%, 39%)				Other (14%, 21%)	Inappropriate choice of route (25%, 19%)		Adverse weather (12%, 51%)
5	127	Fire (63%, 87%)	Fishing (24%, 25%)	Open sea (21%, 50%)	No (13%, 76%)	Total loss of the ship (18%, 39%)	No (12%, 91%)	Yes (17%, 30%)	No (13%, 72%)	No (15%, 87%)	No (18%, 75%)	No (13%, 91%)	Technical failure (21%, 63%), Problem with cargo (21%, 15%)
6	41	Collision (7%, 68%)	Special Craft (43%, 100%)	Coastal waters (6%, 44%)	No (5%, 90%)	Ship rendered unfit to proceed (6%, 46%)	No (4%, 93%)		No (5%, 80%)				No (6%, 61%)
7	60	Damage to ship (100%, 100%)				Ship remains fit to proceed (8%, 77%)					No (8%, 73%)	No (6%, 90%)	Technical failure (10%, 67%)
8	45	Collision (11%, 98%)	Fishing (27%, 80%)	Open sea (8%, 53%)	Yes (8%, 62%)	Total loss of the ship (11%, 67%)	No (5%, 96%)	No (5%, 96%)		No (6%, 93%)			No (8%, 78%)
9	165	Collision (40%, 95%)	Cargo (24%, 96%)	Port (18%, 49%)	No (18%, 85%)	Ship remains fit to proceed (17%, 58%)	Yes (19%, 25%)	No (17%, 96%)	No (18%, 75%)	No (17%, 78%)	Failure to respond appropriately (27%, 23%), Error in judgement (19%, 39%), Deciding not to pass on information (26%, 7%)		Adverse weather (20%, 40%)

10	76	Work accident (12%, 37%)	Tanker (50%, 100%)			Ship remains fit to proceed (9%, 70%)	Yes (12%. 33%)	No (7%, 90%)		Communication (12%, 21%)			No (12%, 74%)
11	57	Grounding (15%, 40%)		Inland waters (10%, 18%)	No (7%, 91%)				Yes (9%, 70%)		Inappropriate choice of route (25%, 26%), Error in judgement (9%, 51%), Incorrect operations of control (11%, 25%), Failure to respond appropriately (9%, 23%)	Necessary (11%, 14%)	Navigational tool problems (87%, 97%)
12	14	Flooding (100%, 100%)	Fishing (4%, 43%)	Coastal waters (3%, 64%)		Total loss of the ship (5%, 86%)			Yes (2%, 72%)	Standards of personal competence (3%, 43%)			
13	57	Capsizing (98%, 100%)	Special Craft (18%, 30%)	Coastal waters (9%, 44%)	Yes (9%, 56%)	Total loss of the ship (19%, 90%)	No (6%, 97%)	Yes (9%, 33%)		Standards of personal competence (8%, 28%)			Problem with cargo (20%, 40%)
14	76	Work accident (33%, 99%)	Cargo (11%, 92%)	Open sea (10%, 41%)	Yes (19%, 93%)	Ship remains fit to proceed (13%, 93%)		No (8%, 100%)	No (8%, 72%)	Other (13%, 21%)			No (13%, 75%)
15	58	Work accident (9%, 34%)		Coastal waters (9%, 43%)					Yes (9%, 72%)	Communication (12%, 28%), Fatigue (67%, 100%), Other (12%, 26%)	Failure to report due to distraction (33%, 14%), Other (10%, 47%), Error in judgement (8%, 48%), Deciding not to pass on information (14%, 10%), Failure to respond appropriately (9%, 22%), Forgetting to report information (18%, 7%)	Necessary (14%, 17%)	
16	105	Work accident (40%, 86%)	Cargo (13%, 83%)	Port (14%, 51%)	Yes (23%, 80%)	Ship remains fit to proceed (16%, 88%)	No (11%, 94%)	Yes (27%, 57%)	Yes (16%. 72%)	Standards of personal competence (26%, 50%), Communication (26%, 32%), Other (16%, 19%)	Incorrect operation of control (19%, 24%), Other (18%, 48%), Error in judgement (16%, 52%), Failure to advise officer on the watch (25%, 5%), Failure to respond appropriately (16%, 21%), Forgetting to report information (32%, 7%)	Necessary (30%, 20%), Routine (21%, 29%), Other (18%, 17%)	Problem with cargo (27%, 30%), Technical failure (14%, 53%),

A second cluster analysis was then applied considering only the subset of accidents involving the vessels carrying dangerous goods. The resulting dendrogram was cut to obtain a three-group partitioning (Figure 2) that proved to be the clearest and most functional for the objectives of this analysis. Table 2 lists the main features of each cluster. The first row shows the global composition of the variables while the second row the composition of the variables for tanker ships only. The remaining rows refer to each cluster and show the category of each variable significantly different from the global distribution that characterizes the most the clusters. A brief description is provided below:

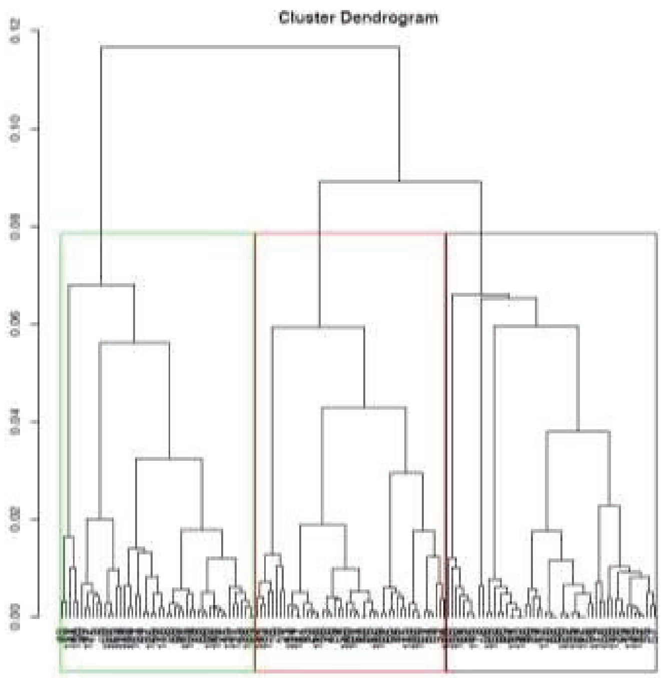

Figure 2. Dendrogram of accidents involving ships carrying dangerous goods

- Cluster 1: the main cause of these accidents is a fire or mechanical failure or damage to the ship, often (65% of cases) caused by a technical problem. The human factor is not a major factor in most of these accidents.
- Cluster 2: the cause of these accidents is a collision or grounding that occurs frequently in inland waters and with the pilot on board. Insufficient support of vessel traffic services equipment is reported in one out of six cases.
- Cluster 3: the causes of these casualties are accidents at work mostly caused by human errors or violations. In these accidents, inadequate supervision or mismanagement of resources (software factors), communication problems and/or inadequate training of the workers resulting in violations and errors of judgement are reported as the main or contributory causes of the accidents.

Table 2: Clusters' size and features – Subset of accidents involving ships carrying dangerous goods (153 accidents)

No.	Size	Type of casualty	Location	Loss of life	Consequences to the ship	Pilot on Board	Hardware factors	Software factors	Personnel factors	Human error	Human violations	Other causes
Overall	1151	Capsizing (5%), Collision (34%), Damage to the ship (5%), Fire (15%), Flooding (1%), Grounding (14%), Machinery Failure (6%), Work Accident (20%)	Coastal waters (25%), Inland waters (8%), Open sea (27%), Port (40%)	Yes (32%) No (68%)	Ship remains fit to proceed (49%) Ship rendered unfit to proceed (28%) Total loss of the ship (23%)	Yes (19%) No (81%)	Yes (19%) No (81%)	Yes (40%) No (60%)	Communication (11%), Standards of personal competence (18%), Fatigue (8%), Other (11%), No (66%)	Error in judgement (29%), Failure to respond appropriately (12%), Incorrect operations of control (11%), Inappropriate choice of route (5%), Forgetting to report information (2%), Failure to advise officer on the watch (2%), Deciding not to pass on information (4%), Failure to report due to distraction (2%), Other Errors (24%), No (46%)	Necessary (6%), Routine (12%), Other (9%), No (75%)	Problem with cargo (10%), Technical failure (33%), Structural failure (5%), Adverse weather (29%), Navigational tool problems (6%), No (40%)
Tanker	153	Capsizing (1%), Collision (28%), Damage to the ship (5%), Fire (22%), Flooding (1%), Grounding (14%), Machinery Failure (7%), Work Accident (22%)	Coastal waters (22%), Inland waters (9%), Open sea (28%), Port (41%)	Yes (33%) No (67%)	Ship remains fit to proceed (53%) Ship rendered unfit to proceed (34%) Total loss of the ship (13%)	Yes (23%) No (77%)	Yes (18%) No (82%)	Yes (35%) No (65%)	Communication (18%), Standards of personal competence (16%), Fatigue (8%), Other (11%), No (63%)	Error in judgement (27%), Failure to respond appropriately (9%), Incorrect operations of control (7%), Inappropriate choice of route (2%), Forgetting to report information (1%), Failure to advise officer on the watch (2%), Deciding not to pass on information (5%), Failure to report due to distraction (3%), Other Errors (29%), No (43%)	Necessary (4%), Routine (16%), Other (5%), No (77%)	Problem with cargo (4%), Technical failure (29%), Structural failure (8%), Adverse weather (25%), Navigational tool problems (5%), No (40%)
1	54	Fire (74%, 46%), Machinery failure (100%, 20%), Damage to the ship (100%, 15%)		No (41%, 78%)	Total loss (85%, 32%)	No (41%, 89%)		No (42%, 76%)	No (48%, 85%)	No (54%, 65%)	No (43%, 93%)	Technical failure (80%, 65%)
2	49	Collision (77%, 67%), Grounding (76%, 33%)	Inland waters (79%, 23%)	No (47%, 98%)		Yes (60%, 43%)	No (38%, 96%)					Navigational tools problem (100%, 16%)
3	50	Work accident (94%, 64%)		Yes (74%, 74%)	Fit to proceed (47%, 76%)			Yes (53%, 56%)	Communication (63%, 34%), Standards of personal competence (60%, 30%), Fatigue (100%, 26%), Other (59%, 20%)	Error in judgement (54%, 44%), Other (56%, 50%), Failure to advise officer on the watch (100%, 6%), Failure to report due to distraction (80%, 8%)	Routine (71%, 34%), Other (86%, 12%)	Problem with cargo (83%, 10%)

5. DISCUSSION

The clusters described provide a general overview of the phenomenon of maritime accidents.

From the results obtained in the analysis of the general dataset there seem to be two main variables that mostly guide the formation of the clusters. The main one is the one describing the type of accident, while the second one describes the type of vessel. Looking more closely at the clusters, the types of accident that occur most frequently are collisions and accidents during work operations. These types of accidents can be subdivided by vessel type (the vessels most involved in maritime accidents are cargo vessels), and by the location of the accident (port, open sea, inland waters, coastal waters).

As can be seen in Table 1, clusters 1 to 8 do not report problems related to human factors, unlike the remaining clusters 9 to 16. In particular, the probability of human errors and/or violations causing an incident increases when problems related to "Software issues" (problems with protocols, standards, and company policy) and "Personnel factors" (inadequate staff training, excessive workloads and communication problems) are reported.

In the specific case of accidents involving ships carrying dangerous goods, the described three-group partitioning proved to be the most functional for the objectives of this analysis. Clusters 1 and 2 highlighted how the human factor is not determining in causing these types of accidents. Fires and explosions are indeed mainly caused by technical problems. Conversely, Cluster 3 groups the accidents during working operations where the human factor plays a crucial role.

In general, the accidents involving dangerous goods do not show any noticeable peculiarity compared to the overall dataset. The only distinguishing feature is a higher percentage of fire and explosion cases than the rest of the sample (22% against 15%).

6. CONCLUSIONS

Despite recent technological and safety advances, numerous maritime accidents continue to occur every year, with causes not always known and ineffective prevention mechanisms.

This study applied clustering methods to incident data derived from an IMO database that collects the maritime accidents occurred around the globe between 2009 and 2019 in an attempt to identify causal links between the various causes of accidents.

A notable distinction emerged between the accidents due to technical causes and those where the human factor plays a predominant role. The latter, in the sense of error and violation, occurs most frequently when there is a lack in company protocols, standards and policies, or a problem with maintenance and the malfunctioning of on-board equipment.

The same methodology was applied to a subset of accidents involving ships carrying dangerous goods. The results confirmed the distinction between the accidents where the human factor is important and the accidents where it is not. Error in judgement, communication problems and low standards of competence of the personnel are pointed to be critical in causing the accidents during work operations. On the other hand, fires and explosions turned out to be more frequent for ships carrying dangerous goods, but they are mainly caused by technical problems.

ACKNOWLEDGMENTS

This research was developed in the framework of the OMD Project (Interreg It-Fr Marittimo 2014-2020)

REFERENCES

1. Abdi, H., and Valentin, D. "Multiple correspondence analysis," Encyclopedia of measurement and statistics, 2(4), 2007, pp. 651-657.

2. Antão P., and Guedes Soares C. "Analysis of the influence of human errors on the occurrence of coastal ship accidents in different wave conditions using Bayesian Belief Networks", Accident Analysis & Prevention, 133(6), 2019.

3. Chu G. and Lyu G. "Critical Assessment on Dangerous Goods Storage Container Yard of Port: Case Study of LPG Tank Container", IEEE International Conference on Industrial Engineering and Engineering Management, 16-19th December, 2018, Bangkok, Thailand, pp. 1751–1755.

4. Fadda P., Fancello G., Frigau L., Mandas M., Medda A., Mola F., Pelligra V., Porta M., and Serra P. "Investigating the Role of the Human Element in Maritime Accidents using Semi-Supervised Hierarchical Methods", Transportation Research Procedia, 52, 2021, pp. 252-259

5. Galieriková A. "The human factor and maritime safety", Transportation Research Procedia, 40, 2019, pp. 1319-1326.

6. IMO – International Maritime Organization, 2021. Frequently asked questions about how COVID-19 is impacting seafarers. https://www.imo.org/en/MediaCentre/HotTopics/Pages/FAQ-on-crew-changes-and- repatriation-of-seafarers.aspx

7. IMO, Resolution 27 november 1997, A.849(20) - Code for the investigation of marine casualites and incidents.

8. IMO, Code of International Standards and Recommended Practices for a Safety Investigation into a Marine Casualty or Marine Incident (Casualty Investigation Code), 2008 (resolution MSC.255(84)).

9. Karakavuz A., Tokgoz B.E., Marquez A., Burak Cankaya M. "Risk assessment of commonly transported chemicals in ports", International Journal of Critical Infrastructures, 16, 2020, pp. 539-544.

10. Lecue M., Darbra R.M. "Accidents in European ports involving chemical substances: Characteristics and trends", Safety Science, 115, 2019, pp.278-284.

11. Luo M. and Shin S.H., "Half-century research developments in maritime accidents: Future directions". Accident Analysis and Prevention, 123, 2019, pp. 448-460.

12. Rai, P. and Singh, S. "A survey of clustering techniques", International Journal of Computer Applications, 7(12), 2010, pp.1-5.

13. The IMDG Code, 2020 Edition (inc. Amendment 40-20). Comes into force on 1 June 2022 and may be applied voluntarily as from 1 January 2021.

14. Wang H., Liu Z., Wang X., Graham T., Wang J. "An analysis of factors affecting the severity of marine accidents", Reliability Engineering & System Safety, 210, 2021.