A grounded theory model for analysis of marine accidents.

Mullai, Arben; Paulsson, Ulf

Published in:
Accident Analysis and Prevention

DOI:
10.1016/j.aap.2011.03.022

2011

Citation for published version (APA):
Mullai, A., & Paulsson, U. (2011). A grounded theory model for analysis of marine accidents. *Accident Analysis and Prevention*, *43*(4), 1590-1603. https://doi.org/10.1016/j.aap.2011.03.022

Total number of authors:
2

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Accident Analysis and Prevention 43 (2011) 1590–1603

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

A grounded theory model for analysis of marine accidents

Arben Mullai*, Ulf Paulsson

Lund University, Lund Faculty of Engineering, Industrial Management and Logistics, Ole Romer vag 1 Lund, Sweden

ARTICLE INFO

Article history:
Received 26 September 2010
Received in revised form 6 March 2011
Accepted 22 March 2011

Keywords:
Analysis model
Database
Maritime transport
Marine accidents
Risks

ABSTRACT

The purpose of this paper was to design a conceptual model for analysis of marine accidents. The model is grounded on large amounts of empirical data, i.e. the Swedish Maritime Administration database, which was thoroughly studied. This database contains marine accidents organized by ship and variable. The majority of variables are non-metric and some have never been analyzed because of the large number of values. Summary statistics were employed in the data analysis. In order to develop a conceptual model, the database variables were clustered into eleven main categories or constructs, which were organized according to their properties and connected with the path diagram of relationships. For demonstration purposes, one non-metric and five metric variables were selected, namely fatality, ship's properties (i.e. age, gross register tonnage, and length), number of people on board, and marine accidents. These were analyzed using the structural equation modeling (SEM) approach. The combined prediction power of the 'ship's properties' and 'number of people on board' independent variables accounted for 65% of the variance of the fatality. The model development was largely based on the data contained in the Swedish database. However, as this database shares a number of variables in common with other databases in the region and the world, the model presented in this paper could be applied to other datasets. The model has both theoretical and practical values. Recommendations for improvements in the database are also suggested.

1. Introduction

1.1. Background

The maritime transport system is vital for the Swedish economy and society. Marine accidents adversely affect the human, the marine environment, and properties and activities aboard ships and ashore in various forms and degree of extent. The effects of accidents vary from minor injuries to fatalities and from insignificant damage to very severe damage to the environment and property. During the period 1985–2008, thousands of marine accidents have been reported. The Swedish maritime authority spends considerable amounts of resources in order to maintain a high level of safety and protect property and the environment. The cost of accidents, including fatalities and injuries, damage to property and the environment, prevention and mitigating measures, and insurance accounts for a considerable share of transport costs.

The main purpose of every accident or risk study is to provide decision makers with valid and reliable information in order to make informed and hopefully better decisions. It is not possible to manage what is not or cannot be measured (Kawka and Kirchsteiger, 1999). The analysis makes use of large amounts of diverse datasets, the most important of which are marine accident data. There is a wide range of research strategies available, and researchers are often confronted with the questions of when and why to use a particular research strategy. The determining conditions for selecting the most appropriate research strategy are the type of research questions posed, the extent of control an investigator has over actual events and the degree of focus on contemporary as opposed to historical events (Yin, 1994). The case history research strategy is one of the most favorable strategies for accident studies. The vast majority of accident studies rely heavily on case histories (Haastrup and Brockhoff, 1991; Facchini and Brockhoff, 1992; Christou, 1999; Konstantinos and Ernestini, 2002). Studies based on historical data, which is one of the types of data most frequently used, are generally preferred (Carol et al., 2001). Furthermore, the case history has become one of the prevailing methods of representing accident knowledge (Brigitte and Carsten, 1997).

In Sweden, as in many other countries, marine accident data are recorded in a database, known in Swedish as the SOS.[1] The review of annual accident reports showed that data are analyzed based on summary statistics and results presented in a format pre-

---

* Corresponding author.
E-mail address: arben.mullai@glocalnet.net (A. Mullai).

[1] In Swedish SjöOlycksSystemet and in English "Sea Casualty System".

0001-4575/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2011.03.022

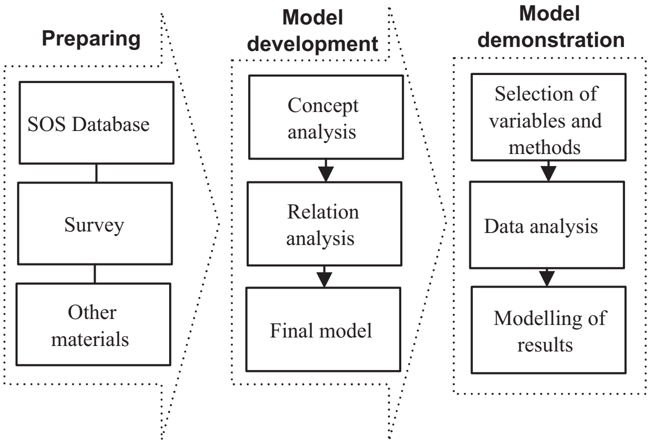

Fig. 1. Research approach.

pared by the Swedish Maritime Investigation Unit (MIU). However, no particular accident model was employed for data analysis. The database consists of a large number of variables, some of which have never been analyzed. Because of dynamic changes in maritime transport systems as well as risk and analysis methods, it is necessary and important to improve the system. According to the MIU, the SOS database has been a well-functioning system. However, the authority intends to improve the system including the employment of advanced analytical tools and integration of the SOS with other systems.

The literature recognizes the issues concerning accident models and the need to develop new models and to update or improve existing ones (Huang et al., 2004; Hollnagel, 2002; Harrald et al., 1998). Such models are a necessary foundation for accident analysis, prediction, and prevention. Inability to reduce accidents can be traced to inadequate accident models and many available models are outdated (Huang et al., 2004). In order to enhance safety and accommodate the needs and changes taking place in systems, these models need to be updated (Huang et al., 2004). Several traditional accident models are based on theories that may not take account of complex systems and phenomena. Therefore, significant progress in accident analysis for complex systems requires innovative approaches starting from a completely different theoretical foundation (Harrald et al., 1998).

Given the above context, the purpose of this study was to develop a conceptual model for description and analysis of marine accidents.

1.2. Methodology

Fig. 1 illustrates the key steps of the research process and the outline paper, namely preparing for model development, model development and demonstration.

- *Preparing for model development*: the study began with identifying the research question and choosing a sample. The question is: how to design an accident model based on empirical data? An extensive literature study was initially performed, where many relevant documents, including reports of marine accidents produced by the Swedish authority (2002–2008) and other documents (EMSA, 2007; SMA, 2009), were collected and studied. A questionnaire comprising some sixty questions concerning reporting, collection, compilation, analysis of marine accident data, and presentation and dissemination of the results was sent to the Swedish authority. The authority returned a detailed description of the system and procedures. The entire SOS database was obtained and thoroughly studied.
- *Model development*: attempts have been made to develop an accident model primarily grounded on marine accident data, “let the data talk” approach, in accordance with the principles of grounded theory and content analysis. Grounded theory is a general methodology for developing theory based on empirical data that are systematically gathered and analyzed (Glaser and Strauss, 1967; Denzin and Lincoln, 1994). Content analysis is a research tool used in a wide array of fields to determine the presence of concepts, which consists of conceptual and relational analysis or concept mapping (Krippendorf, 1980; Carley and Palmquist, 1992). The ability to abstract and understand the underlying dynamics of a complex system and phenomenon plays an important role in both methodologies. The SOS database contains detailed data on thousands of accident case histories organized by *ship* and *variable*. The selective reduction is the central idea of content analysis. The large number of variables, which represent essential properties of the maritime transport system and risk elements, were stepwise coded and organized into main categories of concepts at different levels of abstraction. Concepts were coded consistently as they appeared in the database and according to their properties. The focus of relational analysis was to look for semantic or meaningful relationships. The relationships among concepts were explored and path diagrams linking variables were designed. The process of model development is graphically described in Fig. 3 and the final model is presented in Fig. 4.
- *Model demonstration*: a model is evaluated by its consistency to the empirical data. It should also have the ability to explain and predict the phenomena under study. In order to demonstrate the model and provide answers to research questions, one categorical and five metric variables were selected and analyzed by means of summary and inferential statistics. Justifications for choices made in model demonstration are also provided. The results of the analysis replicated the model.

2. The theoretical platform – key definitions and concepts

The key definitions and concepts relevant to model design are the maritime transport system, risks, risk analysis, and accident modeling.

2.1. The maritime transport system

The maritime transport system is a very complex and large-scale (Grabowski et al., 2010) *socio-technical environment* (*STE*) system comprising human and man-made entities that interact with each other and operate in a physical environment (Mullai, 2004). The main elements of the system are objects of transport, means of transport, infrastructures, and facilities, which are linked by the information system and transport-related activities. The human is a very important element that designs, develops, builds, operates, manages, regulates, and interacts with other elements of the system. Individuals, groups, their relationships, and communication constitute organizational systems. These elements are embedded in very complex, interdependent, and dynamic relationships.

2.2. Accidents, risks and risk analysis

In essence, the concept of risk is defined as the likelihood of consequences of undesirable events (Vanem and Skjong, 2006; Hollnagel, 2008). Accidents and incidents are negative outcomes of the systems. The terms “marine accident and incident” and “marine casualty” denote undesirable events in connection with ship operations (IMO, 1996). An accident is an undesired event

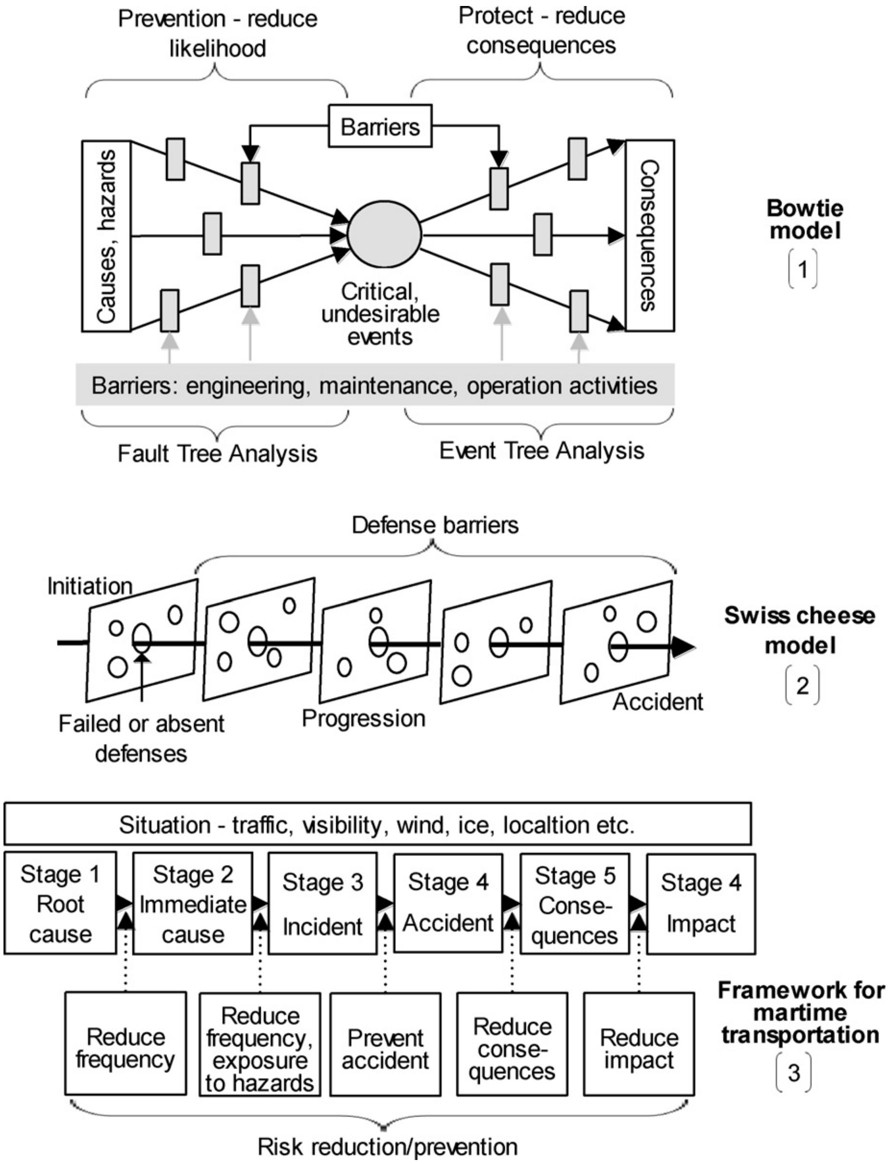

Fig. 2. Linear accident models adapted from respective sources: Bowtie model (1) (Hollnagel, 2008), Swiss cheese model (2) (Reason, 1990), framework for maritime risk assessment (3) (Harrald et al., 1998).

that results in adverse consequences, for example injury, loss of life, economic loss, environmental damage, and damage to or loss of property (Harrald et al., 1998; Grabowski et al., 2010). Accidents are due to an unexpected combination of conditions or events (Hollnagel et al., 2006). Although there is a distinction between "accident" and "incident" in terms of the magnitude of consequences, the term "accident" is most frequently used in this paper.

Risk analysis is the systematic use of available information to identify hazards and estimate the risk to people, the environment, and property (Mullai, 2004; Lars Harms-Ringdahl, 2004). In order to understand risks, risk analysis attempts to provide answers to three fundamental questions: "What can go wrong?" "What are the consequences?" and "How likely is that to happen?" – known as the "triplet definition" of risk (Kaplan et al., 2001). These questions can lead to other questions, which, in turn, require additional answers and efforts. Risks can also be measured as a combination of consequences relative to the number of risk receptors exposed to the undesirable events. This form of risk estimation has become a legal requirement in several countries (OECD, 2004). Thus, the risks can be expressed as a function (f) of frequency, consequence, and exposure (Eqs. (1) and (2)) (Mullai, 2007).

$$\sum\nolimits_{R_i} = \sum f(F_i, C_i, E_i) \tag{1}$$

$$\sum\nolimits_{R_i} = \sum f(F_i, C_i) \tag{2}$$

where R_i – individual, societal, and aggregated risks. The latter are compounded human risks (fatality, injury, and other health risks), environmental risks, property risks, and other risks. F_i – frequency – likelihood, probability; C_i – consequences for risk receptors, i.e. human, the environment, property, and other, e.g. disruption and reputation. E_i – exposure, i.e. the number and categories of risk receptors exposed to but not necessarily affected by the undesirable events.

By definition, the concepts of risk and risk analysis have a wider scope than those of accident and accident analysis. The accident is a constituent element of the risk. Risk analysis encompasses a wider range of processes than accident analysis, including exposure analysis and risk estimation and presentation.

2.3. Accident models

Different terms are used to describe accident phenomena as well as analysis tools, for example approaches, techniques, frameworks, methodologies, methods, and models. The term *accident model* is frequently used in the literature (Leveson, 2004; Grabowski et al., 2000; Nikolaos et al., 2004; Laracy, 2006).

Accident analysis, which always implies an accident model (Hollnagel, 2002), is a very important process for providing input to the development of proactive and cost-effective regulations (Psarros et al., 2010). An accident model is an abstract conceptual representation of the occurrence and development of an accident; it describes the way of viewing and thinking about how and why an accident occurs and predicts the phenomenon (Huang et al., 2004; Hollnagel, 2002).

Hollnagel (Hollnagel, 2002; Hollnagel et al., 2006) divide accident models into three main types, namely *sequential*, *epidemiological*, and *systemic and functional*. Each type consists of a set of assumptions on how the reality is viewed and how accident analysis should be performed and the theoretical foundation and limitations (Hollnagel, 2002; Hollnagel et al., 2006). Epidemiological accident models describe an accident as the outcome of a combination of factors. Such models are rarely strong, as they are difficult to specify in great detail. Systemic accident models consider accidents as emergent phenomena and are based on control theory, chaos models, stochastic resonance, and systems approach. In the latter, the system is viewed as a whole rather than individual components or functions. Systemic models are difficult to represent graphically (Hollnagel, 2002; Hollnagel et al., 2006).

Most accident models are sequential viewing accidents as a sequential chain of events that occur in a specific order (Harrald et al., 1998; Hollnagel, 2008; Leveson, 2004; Nikolaos et al., 2004; Van Drop et al., 2001; Özgecan and Ulusçu, 2009; Celik et al., 2010). Three typical sequential models, namely the Bowtie model, Swiss cheese model, and a framework for maritime risk assessment, are presented in Fig. 2. The fundamentals of many accident models lie on the Swiss cheese and Bowtie models. The latter describes how a critical event may have several precursors and consequences (Fig. 2, Model 1). The former (proposed by Reason in 1990) views accidents as outcomes of interrelations between unsafe acts and latent conditions in the system, such as weakened barriers and defenses represented by the holes in the slice of "cheese" (Fig. 2, Model 2). Harrald et al. (1998) described accidents by means of a six stage causal chain: root or basic causes and immediate causes, triggering incidents, accidents, consequences and impacts (Fig. 2, Model 3). The causes are broken down into main categories, such as human, machine, environmental (Özgecan and Ulusçu, 2009) and organizational factors (Van Drop et al., 2001). Most studies of transport related accidents have claimed that human errors cause 80% of accidents (Harrald et al., 1998; Jens, 1997).

In order to predict accidents, the conditional probabilities that link the stages in the causal chain should be estimated, but these linkages are extremely difficult to establish and require assumptions and innovative use of available data (Harrald et al., 1998). This problem is partly attributed to the accident model and data analysis methods employed. Sequential models work well for accidents caused by failure of physical components and for relatively simple systems (Leveson, 2004). But, they do not work well in all situations and systems because these models do not account for and therefore cannot correctly model complex phenomena and systems (Leveson, 2004; Trucco et al., 2008). They only provide partial answers to questions concerning safety measures (Harrald et al., 1998).

Many models such as the Swiss cheese analogy model (Reason, 1997), the Bowtie model (Hollnagel, 2008), the FSA (IMO, 2002), FTA and ETA techniques, and other models (Leveson, 2004; Van Drop et al., 2001; Li et al., 2009; Larsson et al., 2010) have been introduced or proposed by experts in the fields with no reference to or systematic analysis of empirical data. Further, many models are based on theories or concepts, for example systems theory concepts (Leveson, 2004; Laracy, 2006; Larsson et al., 2010), the Bayesian Belief Network concept (BBN) (Trucco et al., 2008; Merrick and Singh, 2003), Neural Networks (NN) concept (Hashemi et al., 1995; Le Blanc et al., 2001), fuzzy logic (Sii et al., 2001), risk-based approaches (Vanem and Skjong, 2006; Celik et al., 2010), simulation and expert judgment (Harrald et al., 1998).

Harrald et al. (1998) have pointed out that a theoretical framework is of little use in an analysis unless there is relevant data to support it. The analysts are confronted by incomplete and misleading data that make it difficult to use theoretical frameworks. Furthermore, the data are not recorded in accident databases in a form compatible with the theoretical constructs and, as a result, significant modeling assumptions have to be made in order to produce valid results (Harrald et al., 1998). The data issues are common problems in maritime safety (Celik et al., 2010) that can be partly attributed to inadequate accident models (Huang et al., 2004).

In many countries, accident data are reported, collected, and compiled over time in databases in accordance with national regulations and established codification systems (Mullai, 2004; Mullai and Paulsson, 2002). Large amounts of resources are spent building and maintaining such databases. Changes to the current systems and early data records require enormous effort, resources, and time. In addition, any inappropriate change may render many years of data records useless. A large number of incidents may not undergo thorough investigation but are simply reported by responsible individuals or organizations according to established guidelines.

Accident studies in many domains including the maritime industry are firmly rooted in probabilistic methods (Leveson, 2004; Trucco et al., 2008) and other summary statistics (Vanem and Skjong, 2006; Grabowski et al., 2000; Van Drop et al., 2001; Özgecan and Ulusçu, 2009; Konovessis and Vassalos, 2008; Merrick and Singh, 2003). The studies are largely confined to a few variables, such as vessel type, flag, and business line, total loss incidents (Li et al., 2009), and some categories of accidents, such as fire/explosion, collision (Van Drop et al., 2001; Le Blanc et al., 2001), contact and grounding (Vanem and Skjong, 2006; Özgecan and Ulusçu, 2009). Summary statistics are important but not sufficient procedures for explaining and predicting accident phenomena. Inferential statistics, including canonical correlation analysis, multivariate analysis of variance, and structural equation modeling (Hair et al., 1998; Hubert and Blalock, 1979), provides results with a high degree of confidence.

In summary, no single model has the capability of serving all systems, issues, and needs in the maritime industry at all times – "one size does not fit all". Were this the case, we would have had only a handful of accident models and studies. Many accident models are based on theories that may not account for complex phenomena and systems. It is, therefore, relevant and important to develop a model grounded on empirical data that would primarily make use of the accident data contained in the databases.

3. Empirical data – the SOS database

The SOS is one of the SMD's (Swedish Maritime Department) databases developed for website application. The database operates on the Microsoft SQL (Structured Query Language, a database computer language for data management) Server. The study of many marine accident databases shows that the SOS is one of the most complete, detailed, and well maintained databases in the BSR (Baltic Sea Region) and other parts of the world (Mullai, 2004; Mullai et al., 2009). The SOS database contains records of approximately 6000 marine accidents reported during the period 1985–2008 organized by ship and variable. The variables are designed based on the European Statistics guidelines, the IMO investigation code, and the DAMA coding system. The latter was agreed in 1990 by the Scandinavian countries for the registration and analysis of marine accidents. According to the Swedish maritime authority, one ship represents one case. Often two or more ships have been involved in a single accident, such as collision. A collision between two Swedish ships is counted as two events, while a collision between a Swedish ship and a foreign ship or pleasure boat is counted as one event. In each case, the data are compiled into 88 different variables with no particular order or categorization. The SOS database maintains records of all marine accidents and incidents that occur in Swedish territorial waters involving ships of all nationalities as well as accidents involving ships flying the Swedish flag outside Swedish territorial waters.

Certain data on properties of the ships involved in accidents were retrieved from SITS (Swedish Maritime Inspectorate Supervision System[2]), a system developed by the SMA in Prolifics/JAM7/Windows. In cooperation with the SSA (Swedish Shipowners' Association) and other actors, the MIU has also developed a database for anonymous reporting of near misses and deviations, known as INSJÖ.

Selection of data analysis techniques is partly determined by the number and attributes of variables, such as type, measurement level (i.e. nominal, ordinal, and scale variables; metric and non-metric variables), and dependence (i.e. dependent and independent variables). The nature of data determines the attributes of variables. The majority of variables in the SOS database are string (56.8%) and numeric (37.5%), while the remainder (5.7%) are date and coordinate (Table 2). A string variable is a variable whose values are not numeric. In terms of the measurement level, variables in the database are nominal, scale, and ordinal, where more than half (55.7%) are nominal. The latter provide limited options for statistical data analyses. The categorical and metric variables account for 64.5% and 35.5%, respectively. Metric variables allow the highest level of precision, permitting all mathematical operations. Because of their properties and the large number of categorical values, some important variables such as coordinate, cause, type of cargo, and event description, have not been used in any accident analysis (2002–2008).

The SOS database is one of the few in the BSR and the world containing coordinate data that enable spatial data analysis and visual presentation of maritime risks. 'Cause' and 'type of cargo' have a very large number of variable values. Thus, 'cause' contains detailed variable values (94 items) codified into eight main categories covering human, technical, operational, managerial, organizational, and external factors. Codification of the causes may be a matter of personal choice and judgment, which may affect the reliability and validity of the data. In addition, the biases and practices of the investigators also affect the quality and usability of the data recorded in accident databases (Harrald et al., 1998). The 'event description' variable contains valuable information in text format concerning marine accidents, causes and contributing factors, mitigating and preventive measures, and various elements of the system. However, the information cannot be analyzed in its existing data format or by means of the traditional data analysis tools used in accident analysis.

4. The grounded theory model – model development

4.1. Variables, constructs, and path diagrams

The essential attributes and relationships of the variables in the SOS database were thoroughly studied. The database contains 88 variables representing the maritime systems and risk elements in complex relationships. One purpose of the conceptual model is to reduce the degree of detail and complexity of variables and their relationships. The total number of variables taken into consideration when designing the model was 87. As the variable 'event description' was in text format it was excluded, as in that form, neither summary nor inferential statistics was applicable to it. However, the detailed information on marine accidents that it contained provided valuable insights for conceptual and relation analysis.

The large number of variables was reduced to eleven *sets or constructs*, organized and clustered according to their common properties (Table 1 and Fig. 3) but retaining their original names. The combined constituent variables of the constructs 'consequence', 'ship property', and 'location' accounted for the vast majority (72.8%) of all variables in the SOS database. The majority of numeric and scale variables are consequence variables (Table 2). With the exception of 'cause' (X_7), all constructs were defined in conceptual terms, which, in their present form, could not be measured directly. For example, in the absence of a common measurement unit, the construct 'consequence' (or aggregated consequence) could not be measured as it is composed of a set of variables (i.e. fatality, injury, disappearance, environmental, and property damage) that are measured by different metrics. The label of each construct represents the highest level of abstraction that best suits the description of its constituent variables. A detailed list of constructs and their constituent variables and attributes is provided in the Appendix A.

The relationships among variables within each construct are explored and depicted by path diagrams. Some variables are hierarchically nested, resulting in intra-class correlations. For example, 'fatality' ($y_{2,1,1}$), 'injury' ($y_{2,1,2}$), and 'disappearance' ($y_{2,1,3}$) are variables that are hierarchically nested at a higher level, i.e. the variable human consequence. Models containing multilevel variables are multilevel models (Bryk and Raudenbush, 1991). The model developed in this study is a multilevel model that consists of many variables (87) in complex relationships. Due to space constraint and for the purpose of illustration, some examples of variables, constructs and their path diagrams are presented in Fig. 3. The circles with broken lines represent some multilevel variables at the high level of abstraction, but, for the reason of reducing the complexity of the model, their constituent variables at the lower levels are not shown.

The upper part of Fig. 3 represents the final model that is shown in Fig. 4. At the final stage of model development, the constructs are organized and connected by means of a path diagram. In terms of dependency, constructs are classified into independent (exogenous or predictor) (X) and dependent (endogenous or response) (Y) based on the properties of their constituent variables. Nine (X_1–X_9) independent and two (Y_1, Y_2) dependent constructs were determined. Marine accidents and their consequences are undesirable outcomes of the maritime transport system. The majority of consequences (Y_2) are direct results of the marine events (Y_1). In other words, the constructs 'marine event' (Y_1) and 'consequences' (Y_2) can be predicted by independent constructs (X_1–X_9). However, some categories of consequences (Y_2) have little or no relation to the marine events (Y_1) or other variables. For example, deaths due to natural causes and some injuries (e.g. intoxication and suffocation) have no relationships to the marine accident categories defined in Table 1. The path diagram is the graphic portrayal of

---

[2] In Swedish, SITS means *Sjöfartsinspektionens TillsynsSystem*, Swedish Maritime Inspectorate Supervision System.

Table 1
Constructs and variables of the SOS database.

	Constructs	Description of variables
X_1	Time	When did it happen? Year, month, day, and day of week.
X_2	Location	Where did it happen? Port, coordinate, map, geographical area, fairway, and trade area.
X_3	Ship's properties	The properties of ships involved in the accident are the IMO number, class society, nationality, year built or reconstructed, size (dwt, brt, and length), building material, and ship type.
X_4	Ship/Ship's activity	What was the ship doing at the time of accident? Activity on board and steering method.
X_5	Cargo	Type of cargo on board the ship.
X_6	Environmental condition/conditions	The environmental conditions recorded at the time of the accident: light, visibility, sea, wind, and precipitation.
X_7	Cause	The main categories are (a) external factors; (b) construction of the ship; (c) technical faults in equipment; (d) operation, management and design of equipment; (e) cargo and safety; (f) communication, organization, and procedures; (g) human factor; (o) unknown causes.
X_8	Other	Variables that cannot be classified in any other categories: pilot on board and obligation, ice breaker assistance.
X_9	Exposure	People on board the ship exposed to the accident, but not necessarily affected: crew, visitors, and the total number of people on board.
Y_1	Marine event	The main categories are collision, contact, grounding, fire, explosion, machinery breakdown, listing, capsizing, human effects, spill, and other.
Y_2	Consequence	What were the consequences? What was the magnitude? The main categories of consequences are human (fatality, injury, and disappearance – crew, passengers, pilot, and other), property (damage to ship), and the environment (amounts and types of pollutants).

the set of relationships among constructs (Fig. 4). These relationships are projections of those among variables at low levels of abstraction that are shown in the lower part of Fig. 3. Single-headed arrows depict the relationships between independent (X_1–X_9) and dependent (Y_1, Y_2) constructs. In relation to the construct Y_2, however, Y_1 is an independent construct. The single-headed arrow emanating from Y_1 and pointing to Y_2 also depicts the relationship between an independent and a dependent construct, whereas the lines connecting constructs (X_1–X_9) depict the relationships between independent constructs.

4.2. *Guide to model application*

One purpose of the accident model is to explain and predict marine accidents. In order to manage maritime risks in an efficient manner, it is important to identify the most influential variables that contribute to accidents and their consequences. The model will assist risk/accident analysts to generate and address a wide range of research questions. One relationship or combination of relationships in the model may form the basis for a research question. Some examples of research questions that could be considered acci-

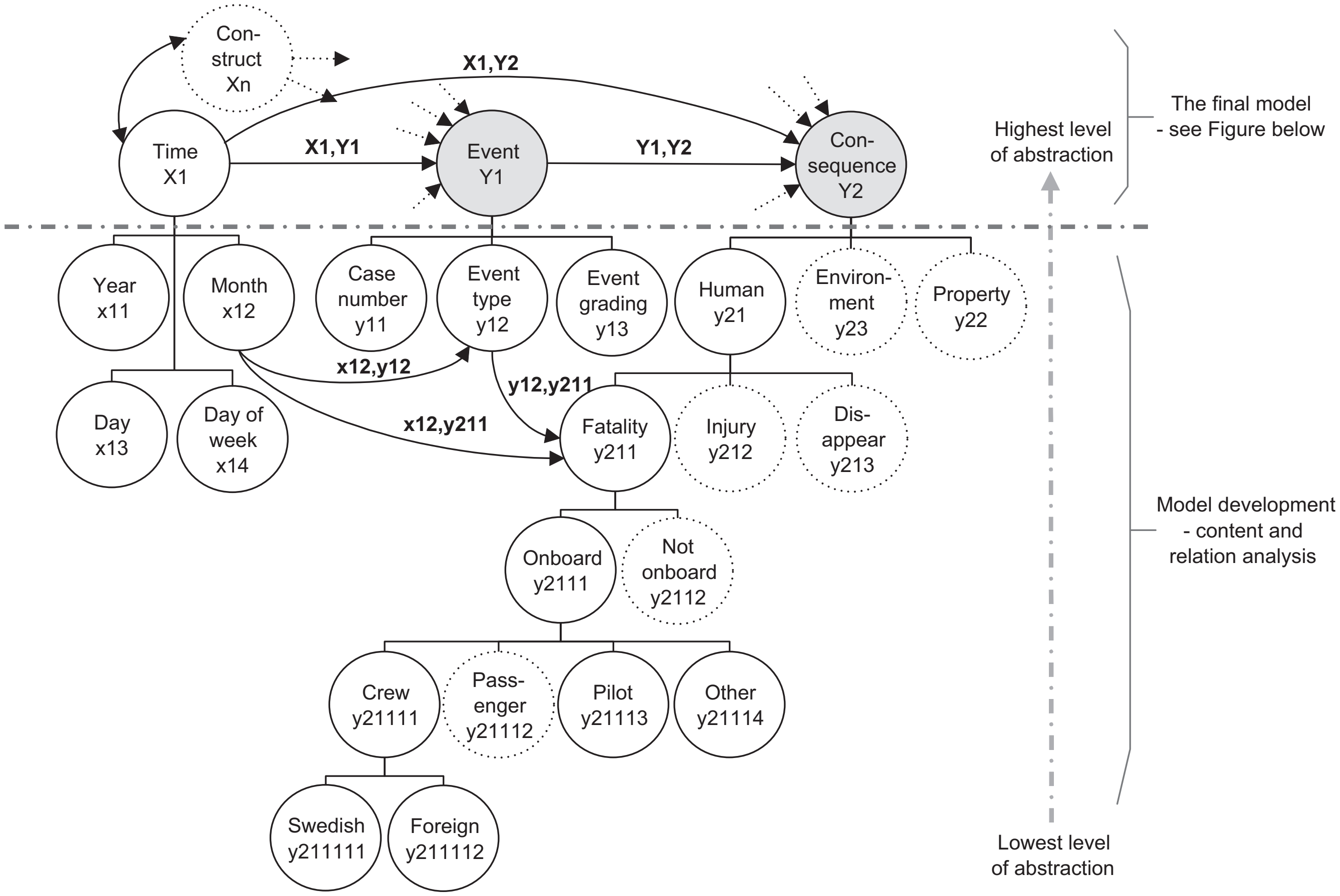

Fig. 3. Example of model development processes.

Table 2
Constructs, variables, and their properties.

	Constructs	Variables								
		Nr	%	Variable type[a]				Measurement level[a]		
				String	Numeric	Date	Coordinate	Nominal	Ordinal	Scale
X_1	Time	3	3.4			3		3		
X_2	Location	12	13.8	10			2	12		
X_3	Ship	16	18.4	13	3			13		3
X_4	Ship's activity	3	3.4	3				3		
X_5	Exposure	3	3.4		3					3
X_6	Cargo	2	2.3	2				2		
X_7	Event	3	3.4	3				2	1	
X_8	Cause	1	1.1	1				1		
X_9	Other	3	3.4	3				3		
Y_1	Environmental conditions	6	7.2	5	1			3	2	1
Y_2	Consequence	35	40.2	9	26			6	3	26
		87	100	49	33	3	2	48	6	33

[a] A nominal variable is one whose values represent categories with no intrinsic ranking. An ordinal variable is one whose values represent categories with some intrinsic ranking. A (interval or ratio) scale variable's values represent ordered categories with a meaningful metric (Statistical Package for the Social Sciences, Version 17.0 for Windows, 2009).

dent or risk studies are: what types and how many accidents have happened? What are the most frequent types of accidents? When and where did they happen? What are the trends? What are the causes of marine accidents? What are the relationships among variables? How strong are these relations? To what extent are marine accidents explained by independent variables? What are the most influential factors? What types of risks are involved? What are the levels of individual and aggregated risks? How significant are they?

Given the number and attributes of variables and the number of accident cases, both non-inferential and inferential statistics are applicable to the data. The model will facilitate application of various data analysis techniques, from simple univariate and bivariate to complex multivariate analysis. Based on Hair et al. (1998) and the properties of variables in the SOS database, the following multivariate data analysis techniques could be considered as the most appropriate techniques in application of the model: canonical correlation, multivariate analysis of variance, multiple discriminant analysis, multiple regression analysis, conjoint analysis, structural equation analysis, factor and cluster analysis.

The analysis should be performed at the second and third levels of the model resolution, as all constructs are unobserved and have no variable values, except for 'cause'. Descriptive statistics can be used to reduce the large amount of data, for example the purposes of risk estimation and presentation. This statistical procedure is very useful because many variables in the database are non-metric. Inferential statistics should be employed in the case of metric variables for inferences about the entire population of marine accidents. Those recorded in the SOS database do not represent the entire population, as many minor and near-miss incidents go unreported. Furthermore, the dataset is, to some extent, biased because data are not collected from random observations, but in connection with marine accidents. However, given the size of data, the number of accidents recorded in the database is a fair representation of the population. Both non-probability and probability samplings, such as simple, systematic, stratified, cluster, and repeti-

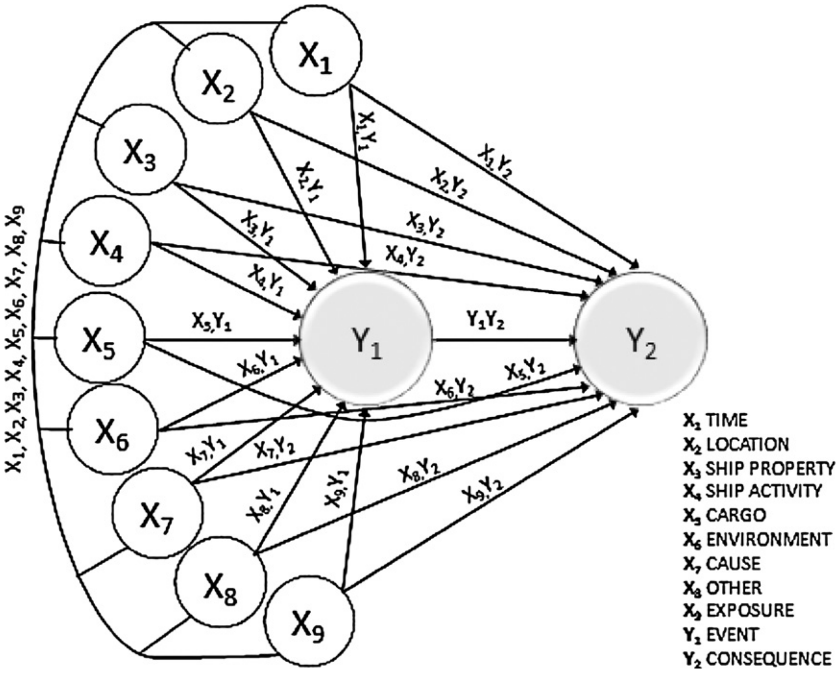

Fig. 4. The constructs and the path diagram of the model – the highest level of abstraction.

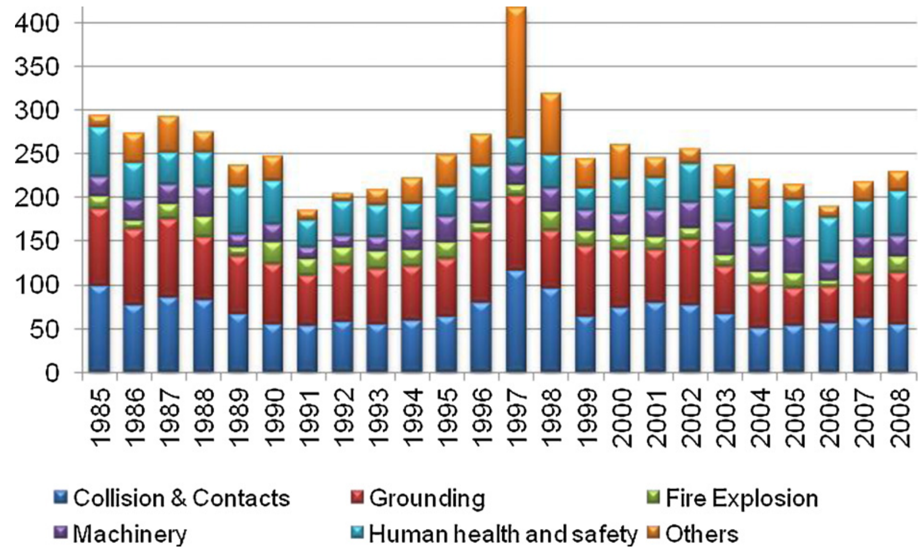

Fig. 5. Marine accidents by category and year.

tive sampling techniques, can be employed for inferences about the SOS data.

The number of variables determines the appropriateness of individual analysis techniques. The ratio between the number of accident cases and independent variables affects the generalization of results. A general rule is that this ratio should never fall below five to one, which means that there should be five cases for every independent variable (Hair et al., 1998). The desired level of the ratio is between 15 and 20 to 1. The ratio of cases and variables in the SOS database (approx. 6000 cases/88 variables = 68) is in accordance with the desired level. The summary statistics procedures are, however, not constrained by the number of cases and variables.

Simultaneous analysis of all 88 variables using any of the multivariate analysis techniques presented above and presentation of the results would be very complex and difficult, if not an impossible procedure. However, application of a comprehensive risk analysis allows the use of all variables as well as the presentation of the results in a systematic manner. The model facilitates exploration and quantification of the system and risk elements.

5. Model demonstration – an illustrative example

The SOS database contains a large number of metric and non-metric variables. For the purpose of model demonstration, some relevant research questions are addressed: What are the most frequent accidents occurring in the Swedish waters and involving ships flying Swedish flag? What is the trend? What categories and how many people have been killed? Can the variable ‘fatality’ be predicted by independent variables? If so, how strong is the prediction power? What are the fatality risks? In order to provide answers to these questions, one non-metric or categorical and five metric variables were selected (see Table 3).

The selected variables are among the most important for measuring the performance of the maritime transport system. All metric variables were selected for model demonstration. Non-metric variables should be converted into metric prior to any inferential statistical analysis. Many non-metric variables contain large numbers of values. Conversion will result in a huge number of dummy variables. Due to space constraints, one important non-metric variable was selected for the purpose of analysis, i.e. marine accidents. The marine accident categories were converted into metric variables with values of 1, 0 (see Table 3).

Both descriptive and inferential statistics were employed for the selected variables. Inferential statistics are performed by means of the AMOS (Analysis of MOment Structures) program, which implements the structural equation modeling (SEM) data analysis approach. The SEM, also known as causal modeling analysis, is a very effective method for analyzing marine accidents. In addition, the method provides a graphical presentation of directions and number of relationships among variables (see Figs. 7 and 8).

In total, 6007 marine accidents and incidents, an average of approx. 250 events per year, were reported during the period 1985–2008 (Fig. 5). These represent the main categories of initial events, which are often followed by one or more subsequent events. The number of accidents declined during this period. The increase in marine accidents during the late 1990s (1996–1998) is explained by the large number of incidents reported in the area of the Sound, i.e. between Sweden and Denmark. At that time, the bridge between the two countries was under construction, which interfered with the dense vessel traffic in the area. Collisions, contacts, and groundings were the most frequent accidents reported in the Baltic Sea and accounted for more than half (54%) of the total number. These accidents are largely due to weather and navigation hazards as well as dense vessel traffic in the region.

During the period 1985–2008, some 1507 fatalities were reported to the SOS. Fig. 6 shows the main categories of fatalities with their respective frequencies – crew members (24.5%), passengers (74%), and other (1.5%). More than one third (36%) of crew

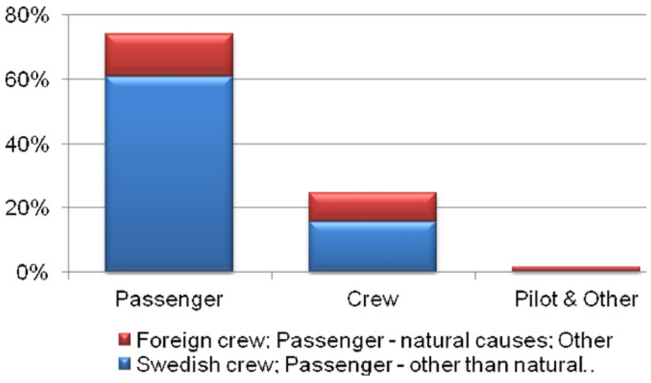

Fig. 6. Categories of fatalities.

Table 3
The name, description, and value of the selected variables.

Name	Description	Value
Fatality_T_I	The total number of fatalities (Swedish and foreign crew, passengers, pilots, and other) including deaths due to natural causes.	Metric
Age	The age of the ship in years.	Metric
BRT	Brutto Register Tonnage or Gross Register Tonnage (GRT) represents the total internal volume of a ship, where a register ton is a measure of the size or carrying capacity of ships and is equal to a volume of 100 cubic feet (2.83 m^3), which volume, if filled with fresh water, would weigh around 2.8 tons.	Metric
Length	The length of the ship in meters.	Metric
People_OB	The total number of people (i.e. crew, passengers, pilots and others) on board the ship exposed to accidents, but not necessarily affected. This variable represents the element of exposure.	Metric
Marine accidents or events		Non-metric
Coll_Contact	Collision with ships, pleasure boats, drilling platforms, and floating objects and contact with quays, bridges, and other fixed objects.	1, 0[a]
Grounding	Grounding and bumping on the bottom/sea floor.	1, 0
Fire_Explo	Fire and explosion in machinery, cargo, and other spaces and electrical installations.	1, 0
Machinery	Machinery breakdown: the main engine, auxiliary engines, propulsion, and steering systems.	1, 0
Human_S_H	Human safety and health includes personal injury, death due to natural causes (e.g. heart attack) and poisoning, suicide, and disappearance.	1, 0
Other	Other events include hull/watertight failure (e.g. leakage, weather, and ice damage); listing/capsizing and stability problems without listing; and other (e.g. ship missing), near accidents, cargo shifting and damage, spills.	1, 0

[a] Dummy variable: 1 for collision and contact accidents and 0 for all other.

fatalities has involved non-Swedish crew. Approximately 18% of passenger deaths were due to natural causes that have often caused disruptions to the system. The category 'other' included pilots and other people on board. The large number of passenger fatalities is explained by two catastrophic accidents involving two passenger ferries. A fire broke out on one ferry, while the other sank due to failure of the hull (the bow door). In the latter case, severe weather conditions contributed to the accident and the seriousness of the consequences. The bow door partly failed due to heavy seas. Many people who survived the capsizing and sinking of the ship reportedly died in lifeboats due to hypothermia because of the low temperature and delays in rescue operations. The severe weather conditions hampered search and rescue operations. These two accidents accounted for more than two thirds (68%) of the total number of fatalities reported to the SOS database during the period 1985–2008.

The 'fatality' variable is an important indicator of the maritime risks. In relation to 'fatality', which is a dependent variable, 'marine accidents', 'ship's age', 'BRT', and 'number of people on board' are independent variables. The variables 'BRT' and 'length', which represent the size of a ship, can also serve as elements of exposure.

The calculations performed to fit the model generated a series of scalar estimates; see Table 4. The scalar estimates encompass some essential measurement parameters, such as covariance and correlation between exogenous variables, variances of exogenous variables, standardized regression weights, intercepts for predicting endogenous variables, and squared multiple correlations. Due to space constraints and complexity, Figs. 7 and 8 present two graphic regression models, where the dependent variable 'fatality' is predicted as linear combinations of the independent variables. Single and double-head arrows represent regression and correlation weights respectively. A variable labeled 'other' was inserted into each model to represent measurement errors and a composite of other variables on which the variable 'fatality' is dependent, but which were not included. In order to identify the regression models, the scale of the latent variable 'other' was defined. Such constraints were imposed in order to make the models identifiable.

The prediction power of the independent variable 'marine accidents' is relatively weak and the combination of all categories of marine accident only accounted for 1%, i.e. the squared multiple correlation, of the variance of the variable 'fatality'. The mean values of collisions/contacts and groundings are larger than those of other categories of accident. However, the 'human health and safety' category has shown a higher prediction power than the other categories (Table 4 and Fig. 7). The correlation between the categories of 'collision/contact' and 'grounding' is relatively strong (−.39) and stronger than the other correlations. Both categories of accident are the most frequent accidents in the Baltic Sea. The most plausible explanations for the strong correlation between these two categories of marine accidents are the effects of weather and environmental conditions (e.g. winds, seas, currents, visibility, ice conditions) in the region.

The ships involved in accidents were relatively small (mean 7598 BRT) and old (mean 20 years) (Table 4). The mean value of number of people on board was 68, indicating that many ships involved in accidents were ferries. The Baltic Sea including Swedish territorial waters is among the areas with the heaviest amount of ferry traffic in Europe as well as the world. The correlation between the variables 'age' and 'BRT' has a coefficient of −.30 (Table 4), which means that as the age of ships involved in accidents increases, the size decreases. In other words, smaller ships were older than larger ships involved in accidents. The number of people on board is partly proportional to the size of the ships (BRT). Today, however, large ships have smaller crew than before. A general view is that old ships are unsafe. The results show that the variable 'age' has the lowest prediction power (2%) in relation to fatalities. Many of the old ships involved in accidents were flying the Swedish flag, which means that they were generally well operated and maintained. In addition, the majority of marine accidents were reported in the Baltic Sea, which is calmer and safer for the navigation of small ships than other areas. Many foundering and hull/watertight failure accidents were attributed to heavy sea conditions. Both categories represent the least frequent accidents in the Baltic. However, any negative effect of old ships on the marine environment and property is not excluded.

The results show that the prediction power of 'fatality' by the exposure element (i.e. the number of people on board ships) alone is quite strong (85%), which means that the fatality risk increases significantly as the number of exposed people increases. The prediction power of the combined independent variables 'ship's properties' and the 'number of people on board' accounted for 65% of the variance of the variable 'fatality'. Although these results may

Table 4
Scalar estimates – maximum likelihood estimates.

Marine accidents & fatality			
Estimate category/variable	Estimate	S.E.[a]	C.R.[b]
Covariances between exogenous variables			
Human_S_H ↔ Others	−.008	.001	−7.692
Machinery ↔ Human_S_H	−.016	.001	−10.818
Machinery ↔ Others	−.005	.001	−5.864
Fire_Explo ↔ Human_S_H	−.01	.001	−8.684
Fire_Explo ↔ Others	−.003	.001	−4.695
Fire_Explo ↔ Machinery	−.006	.001	−6.624
Grounding ↔ Fire_Explo	−.018	.001	−12.318
Coll_Contact ↔ Grounding	−.078	.003	−27.531
Coll_Contact ↔ Others	−.015	.001	−11.011
Coll_Contact ↔ Human_S_H	−.045	.002	−19.955
Coll_Contact ↔ Machinery	−.028	.002	−15.406
Coll_Contact ↔ Fire_Explo	−.018	.001	−12.413
Grounding ↔ Others	−.014	.001	−10.927
Grounding ↔ Human_S_H	−.044	.002	−19.809
Grounding ↔ Machinery	−.027	.002	−15.290
Variances of exogenous variables			
Coll_Contact	.202	.004	53.768
Grounding	.199	.004	53.582
Fire_Explo	.061	.001	53.768
Machinery	.089	.002	53.768
Human_S_H	.134	.002	53.768
Others	.049	.001	53.768
Standardized regression weights			
Fatality_T_I ↔ Coll_Contact	−.022		
Fatality_T_I ↔ Grounding	−.025		
Fatality_T_I ↔ Fire_Explo	.036		
Fatality_T_I ↔ Machinery	−.017		
Fatality_T_I ↔ Human_S_H	.053		
Fatality_T_I ↔ Others	−.013		
Means of exogenous variables			
Coll_Contact	.281	.006	47.577
Grounding	.277	.006	47.203
Fire_Explo	.065	.003	20.109
Machinery	.099	.004	25.144
Human_S_H	.159	.005	33.052
Others	.052	.003	17.787
Correlations between the exogenous variables			
Human_S_H ↔ Others	−.102		
Machinery ↔ Human_S_H	−.144		
Machinery ↔ Others	−.077		
Fire_Explo ↔ Human_S_H	−.115		
Fire_Explo ↔ Others	−.062		
Fire_Explo ↔ Machinery	−.087		
Grounding ↔ Fire_Explo	−.164		
Coll_Contact ↔ Grounding	−.389		
Coll_Contact ↔ Others	−.146		
Coll_Contact ↔ Human_S_H	−.272		
Coll_Contact ↔ Machinery	−.207		
Coll_Contact ↔ Fire_Explo	−.165		
Grounding ↔ Others	−.145		
Grounding ↔ Human_S_H	−.270		
Grounding ↔ Machinery	−.206		
Intercepts for predicting endogenous variables			
Fatality_T_I	.122	.109	1.122
Squared Multiple Correlations			
Fatality_T_I	.007		

Ship's properties, exposure & fatality			
Estimate category/variable	Estimate	S.E.[a]	C.R.[b]
Covariances between exogenous variables			
BRT ↔ People_OB	1.43E+6	93.66E+3	15.268
People_OB ↔ Length	−1.69E+4	1.87E+3	−9.017
BRT ↔ Length	8.17E+5	51.58E+3	15.842
BRT ↔ Age	−7.08E+4	3.31E+3	−21.378
People_OB ↔ Age	−395.415	117.675	−3.360
Length ↔ Age	−497.893	64.895	−7.672
Variances of exogenous variables			
BRT	181.81E+6	3.47E+6	52.343
People_OB	18.04E+4	4.39E+3	41.039
Length	76.01E+4	1.44E+3	52.829
Age	300.001	5.702	52.612
Other	1.576	.043	36.342
Standardized regression weights			
Fatality_T_I ↔ People_OB	.848		
Fatality_T_I ↔ Age	−.015		
Fatality_T_I BRT ↔	−.226		
Fatality_T_I ↔ Length	.183		
Means of exogenous variables			
Age	20.75	.233	89.202
BRT	7598.03	181.62	41.835
Length	88.85	3.69	24.092
People_OB	68.72	6.67	10.298
Correlations between the exogenous variables			
BRT ↔ People_OB	.250		
People_OB ↔ Length	−.144		
BRT ↔ Length	.220		
BRT ↔ Age	−.303		
People_OB ↔ Age	−.054		
Length ↔ Age	−.104		
Intercepts for predicting endogenous variables			
Fatality_T_I	.004	.043	.105
Squared Multiple Correlations			
Fatality_T_I	.645		

[a] Standard error of regression weight.

[b] Critical ratio for regression weight, i.e. dividing the regression weight estimate by the estimate of its standard error.

not seem very surprising, all studies in the reference list as well as many other studies in the field have overlooked the fact that the exposure element plays an important role in maritime risks. In addition, as mentioned earlier, many accident studies have stated that the human factor is responsible for approximately 80% of accidents, which means the other factors combined should account for 20%. But, the results of this study suggest that this statement may not be quite accurate. The earlier studies have been largely based on summary statistics (Vanem and Skjong, 2006; Grabowski et al., 2000; Van Drop et al., 2001; Özgecan and Ulusçu, 2009) and accident models that do not account for complex phenomena and systems (Leveson, 2004; Trucco et al., 2008). Prevention programs are often proposed to reduce the incidence of human error (Harrald et al., 1998). Large amounts of resources were spent on reducing risks, sometimes at a margin, by mainly focusing on the human factor as well as technical, managerial, and operational factors.

The variable 'fatality' is, in part, predicted by the independent variables 'marine accidents', 'ship's properties', and 'number of people on board'. On the other hand, the variable 'fatality' affects maritime systems in the medium or long-term. Some major changes in these systems have taken place in response to catastrophic accidents involving many fatalities. At present, however,

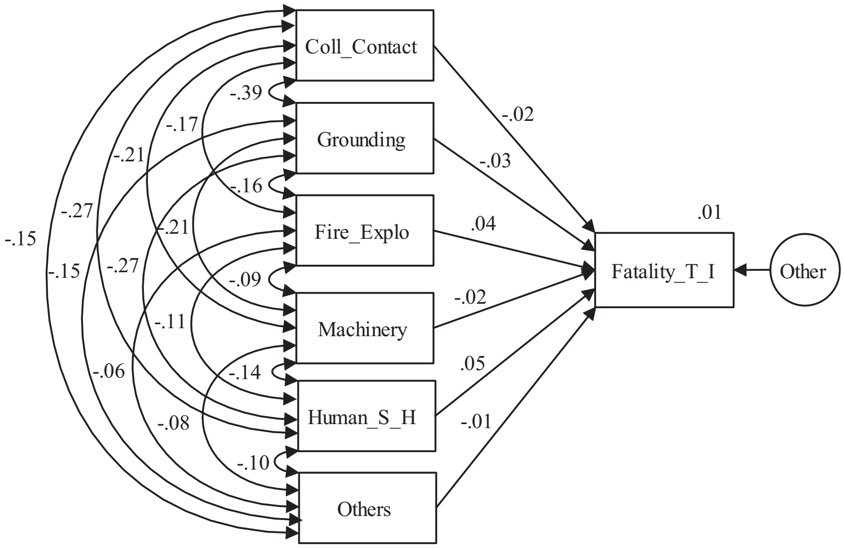

Fig. 7. The linear regression model (standardized estimates) – 'fatality' predicted by marine incidents.

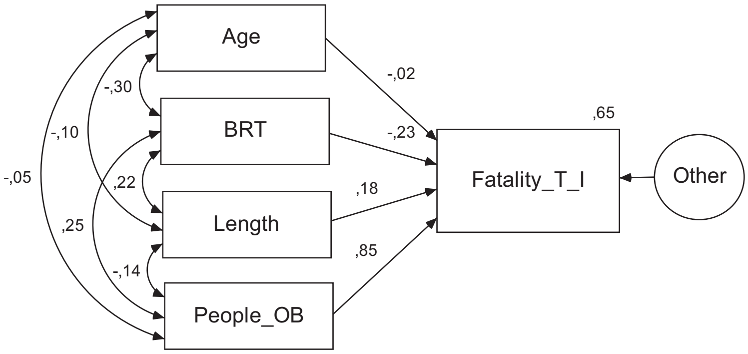

Fig. 8. The linear regression model (standardized estimates) – 'fatality' predicted by ship's properties and number of people on board.

the SOS and many other databases have no data to statistically test the prediction of changes in the systems by 'fatality'.

The frequencies of the magnitudes of 'fatalities' are obtained from the entire dataset, after which the risks are estimated and presented as the FN-curve (Fig. 9), illustrating the relationship between the cumulative frequencies per year and magnitudes of fatalities. The magnitude of fatalities in a single accident varies between 1–6, 15, 158 and 853. The slope or degree of the FN-curve inclination changes significantly after the magnitude of 10, i.e. from a normal distribution to unusually severe magnitudes extending close to 1000 fatalities per accident. As mentioned earlier, the risks were increased due to two catastrophic accidents.

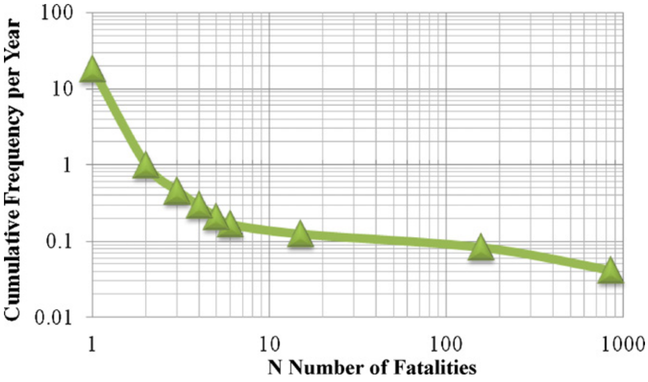

Fig. 9. The FN-curve of fatality risks.

The Swedish maritime authority has, at present, no criteria for benchmarking the fatality risk, which makes it difficult to judge its significance. However, compared to the results of another study (Mullai and Larsson, 2008) concerning risks in the USA's supply chain (1990–2004), the fatality risk in the Swedish maritime system and Swedish territorial waters is higher than that in the USA. These results may serve as the basis for establishing human risk criteria for the Swedish maritime transport system.

6. Conclusions and recommendations

This paper addressed a relevant and important subject. With reference to the purpose of our study, a conceptual model was developed for analysis of marine accidents. The model consists of eleven main constructs and a path diagram. The results of the analysis replicated the model, which indicates that it is valid and reliable. The model has both explanatory and predictive abilities. The model was developed based on a different approach, i.e. grounded theory. Unlike theoretically based accident models, the model presented in this paper was primarily grounded on a large amount of empirical data (ca. 6000 marine accidents). In addition, the model will serve the analysis of the similar data from which it emerged, i.e. the marine accident data. The SOS database shares many variables in common with other databases in the BSR and the world, and therefore the model can be applied to other datasets. Based on the results of our study, we propose the following:

- The SOS database records marine accidents involving ships of all nationalities in Swedish territorial waters as well as ships flying the Swedish flag outside Swedish territorial waters. Therefore, for the purpose of data filtering or sampling, certain variables should be designed.
- The text format variable ‘event description’ contains very valuable information. With some adjustments, certain variables could be identified from the text and designed accordingly. The variable ‘event description’ could also be analyzed by means of text analysis tools.
- A large number of variables in the database are categorical, where the values have no real numerical relationship to each other. In order to make use of all variables in the database and employ inferential statistics, design dummy variables must be designed by assigning numbers (1, 0) to variable values, thus transforming categorical into metric variables. Some variables, for example ‘cause’, ‘cargo’, and ‘ship’, have a large number of values, and therefore extra caution should be exercised in the transformation process.
- Based on the proposed model, analytical tool software could be designed to facilitate the analysis process.

The results have practical and theoretical implications for decision makers and academics. The model has the ability to explain and predict complex accident phenomena. It can facilitate accident or risk analysts to be more efficient and effective and, at the same time, produce detailed, valid, and reliable results. Finally, the application of the model may yield benefits in terms of saved time and resources while maintaining a high level of human safety as well as environmental and property protection.

Appendix A. A detailed structure of constructs and variables applied to the SOS database.

Number		Constructs and variables			Metric (M)/non-metric (NM)
		Name	Type	Measure	
	$\mathbf{X_1}$	**Time**			
1	x_{11}	Year	Date	Nominal	NM
2	x_{12}	Month	Date	Nominal	NM
3	x_{13}	Day	Date	Nominal	NM
4	x_{14}	Day of the week	Date	Nominal	NM
	$\mathbf{X_2}$	**Location**			
5	x_{21}	Port	String	Nominal	NM
6	x_{22}	Port of departure	String	Nominal	NM
7	x_{23}	Port of arrival	String	Nominal	NM
8	x_{24}	Map no	String	Nominal	NM
9	x_{25}	Map year	String	Nominal	NM
10	x_{26}	Map land	String	Nominal	NM
11	x_{27}	Latitude	String	Nominal	NM
12	x_{28}	Longitude	String	Nominal	NM
13	x_{29}	International geographical area	String	Nominal	NM
14	x_{210}	National geographical area	String	Nominal	NM
15	x_{211}	Fairway type	String	Nominal	NM
16	x_{212}	Traffic area	String	Nominal	NM
	$\mathbf{X_3}$	**Ship’s properties**			
17	x_{31}	IMO number	String	Nominal	NM
18	x_{32}	Call sign	String	Nominal	NM
19	x_{33}	Ship’s name	String	Nominal	NM
20	x_{34}	Ship type (1)	String	Nominal	NM
21	x_{35}	Ship type (2)	String	Nominal	NM
22	x_{36}	Class society	String	Nominal	NM
23	x_{37}	Nationality	String	Nominal	NM
24	x_{38}	Year built	String	Nominal	NM
25	x_{39}	Reconstruction year	String	Nominal	NM
26	x_{310}	Brutto Register Ton (BRT)	Numeric	Scale	M
27	x_{311}	Dead Weight (DWT)	Numeric	Scale	M
28	x_{312}	Length	Numeric	Scale	M
29	x_{313}	Material	String	Nominal	NM
30	x_{314}	Catamaran	String	Nominal	NM
31	x_{315}	High Speed Ship	String	Nominal	NM
32	x_{316}	Machinery type	String	Nominal	NM
33	x_{317}	Bridge manning	String	Nominal	NM
	$\mathbf{X_4}$	**Ship’s activity**			
34	x_{41}	Ship’s activity	String	Nominal	NM
35	x_{42}	Activity on board	String	Nominal	NM
36	x_{43}	Steering method	String	Nominal	NM
	$\mathbf{X_5}$	**Cargo**			
37	x_{51}	Cargo description	String	Nominal	NM
38	x_{52}	Dangerous goods	String	Nominal	NM
	$\mathbf{X_6}$	**Environmental conditions**			
39	x_{61}	Light	String	Nominal	NM
40	x_{62}	Visibility	String	Ordinal	NM
41	x_{63}	Sea	Numeric	Scale	M
42	x_{64}	Wind speed	String	Ordinal	NM

Number		Constructs and variables			Metric (M)/non-metric (NM)
		Name	Type	Measure	
43	x_{65}	Wind direction	String	Nominal	NM
44	x_{66}	Precipitation	String	Nominal	NM
	$\mathbf{X_7}$	**Cause**			
45	x_{71}	Cause description	String	Nominal	NM
	$\mathbf{X_8}$	**Other**			
46	x_{81}	Pilot on board	String	Nominal	NM
47	x_{82}	Pilot obligation	String	Nominal	NM
48	x_{83}	Ice breaker assistance	String	Nominal	NM
	$\mathbf{X_9}$	**Exposure**			
49	x_{91}	Number of crew	Numeric	Scale	M
50	x_{92}	Number of visitors on board	String	Nominal	M
51	x_{93}	Total number of persons on board	Numeric	Scale	M
	$\mathbf{Y_1}$	**Event**			
52	y_{11}	Case number	String	Nominal	NM
53	y_{12}	Event type	String	Nominal	NM
54	y_{13}	Event grading	String	Ordinal	NM
	$\mathbf{Y_2}$	**Consequence**			
	$\mathbf{Y_{21}}$	**Human consequence – aggregated**			
55	Y_{211}	Fatality – total	Numeric	Scale	M
63	y_{2111}	Fatality – total on board	Numeric	Scale	M
58	y_{21111}	Fatality – total crew	Numeric	Scale	M
56	y_{21111}	Fatality – Swedish crew	Numeric	Scale	M
57	y_{211112}	Fatality – foreign crew	Numeric	Scale	M
59	y_{21112}	Fatality – passenger	Numeric	Scale	M
60	y_{211121}	Fatality – passenger (natural)	Numeric	Scale	M
61	y_{21113}	Fatality – pilot	Numeric	Scale	M
62	y_{21114}	Fatality – other on board	Numeric	Scale	M
64	y_{2112}	Fatality – other not on board	Numeric	Scale	M
65	Y_{212}	Injury – total	Numeric	Scale	M
66	y_{2121}	Injury – Swedish crew	Numeric	Scale	M
67	y_{2122}	Injury – foreign crew	Numeric	Scale	M
68	y_{2123}	Injury – passenger	Numeric	Scale	M
69	y_{2124}	Injury – pilot	Numeric	Scale	M
70	y_{2125}	Injury – other on board	Numeric	Scale	M
71	y_{2126}	Injury – other not on board	Numeric	Scale	M
72	Y_{213}	Disappeared – total	Numeric	Scale	M
73	y_{2131}	Disappeared – Swedish crew	Numeric	Scale	M
74	y_{2132}	Disappeared – foreign crew	Numeric	Scale	M
75	y_{2133}	Disappeared – passenger	Numeric	Scale	M
76	y_{2134}	Disappeared –others on board	Numeric	Scale	M
77	y_{2135}	Disappeared – pilot	Numeric	Scale	M
	$\mathbf{Y_{22}}$	**Property damage – ship**			
78	y_{221}	Type of damage	String	Nominal	NM
79	y_{222}	Location of damage (port/starboard)	String	Nominal	NM
80	y_{223}	Location of damage (height above bottom)	String	Nominal	NM
81	y_{224}	Location of damage (length)	String	Nominal	NM
82	y_{225}	Damage (largest impression)	String	Nominal	NM
83	y_{226}	Damage (length)	String	Nominal	NM
	$\mathbf{Y_{23}}$	**Marine environment consequences**			
84	y_{231}	Amount of oil spill	Numeric	Scale	N
85	y_{232}	Amount of other environmental pollutant spill	Numeric	Scale	N
86	y_{233}	Type of oil spill	String	Nominal	NM
87	y_{234}	Type of other environmental pollutant spill	String	Nominal	NM

References

Brigitte, R., Carsten, D.G., 1997. Accidents and Risk Control. System Analysis Department, Risö National.

Bryk, A.S., Raudenbush, S., 1991. Hierachical Linear Models for Social and Behavioural Research: Applications and Data Analysis Methods. Sage Publications, Newbury Park, CA.

Carley, K., Palmquist, M., 1992. Extracting, representing and analyzing mental models. Social Forces 70 (3), 601–636.

Carol, S., Vilchez, J.A., Casal, J., 2001. A new approach to the estimation of the probable number of fatalities in accidental explosions. Safety Science 39, 205–217.

Celik, M., Lavasani, S.M., Wang, J., 2010. A risk-based modelling approach to enhance shipping accident investigation. Safety Science 48, 18–27.

Christou, M.D., 1999. Analysis and control of major accidents from the intermediate temporary storage of dangerous substances in marshalling yards and port areas. Journal of Loss Prevention in the Process Industries 12 (1), 109–119.

Denzin, N.K., Lincoln, Y.S., 1994. Handbook of Qualitative Research: Competing Paradigms in Qualitative Research. Sage Publication.

EMSA (European Maritime Safety Agency) Notice of the Swedish Maritime Administration, Maritime Investigation Unit, No. 1 2008. Summary of reported marine casualties, near accidents and accidents to persons – Swedish merchant and fishing vessels 2007. Available at: http://emsa.eu.int/ (accessed on December 2009).

Facchini A. and Brockhoff L.H. (1992) Comparison of road accidents during transport of dangerous goods in France, Italy and Germany. European Commission, DGXIII/D-2, Paper presented: Transport of Dangerous Goods by Road: Safety and Environment Protection, Karlstad (SE), June 2–4, 1992.

Glaser, B.G., Strauss, A.L., 1967. The Discovery of Grounded Theory: Strategy for Qualitative Research. Aldine de Gruyter, New York.

Grabowski, M., Merrick, J.R.W., Harrald, J.R., Mazzuchi, Th.A., René van, J., 2000. Dorp risk modeling in distributed, large-scale systems. IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans 30 (6), 2000.

Grabowski, M., You, Z.H., Song, H., Wang, H., Merrick, J.R.W., 2010. Sailing on friday: developing the link between safety culture and performance in safety-critical systems. IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans 40 (2).

Haastrup, P., Brockhoff, L.H., 1991. Reliability of accident case histories concerning hazardous chemicals – An analysis of uncertainty and quality aspects. Commission of the European Communities, Joint Research Centre Ispra, t.p. 321,1-21020 Ispra, Italy. Journal of Hazardous Materials 27 (1991), 339–350.

Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.B., 1998. Multivariate Data Analysis, Fifth Ed. Prentice Hall, Upper Saddle River, New Jersey.

Harrald, J.R., Mazzuchia, T.A., Spahna, J., Van Dorpa, R., Merrick, J., Shrestha, S., Grabowski, M., 1998. Using system simulation to model the impact of human error in a maritime system. Safety Science 30, 235–247.

Hashemi, R.R., Le Blanc, L.A., Rucks, C.T., Shearry, A., 1995. A neural network for transportation safety modelling. Expert Systems with Applications 9 (3), 247–256.

Hollnagel, E., 2002. Understanding accidents – from root causes to performance variability. In: IEEE 7 Human Factors Meeting Scottsdale Arizona.

Hollnagel, E., 2008. Risk + barriers = safety? Safety Science 46, 221–229.

Hollnagel, E., Woods, D.D., Leveson, N., 2006. Resilience Engineering: Concepts and Precepts. Ashgate Publishing Company, Aldershot, Hampshire, England, Laboratory, DK-4000 Roskilde, Denmark, 1997.

Huang, Y.H., Ljung, M., Sandin, J., Hollnagel, E., 2004. Accident models for modern road traffic: changing times creates new demands. International Conference on Systems, Man and Cybernetics.

Hubert, M., Blalock, J.R., 1979. Social Statistics, Second Ed. McGraw-Hill International Editions.

IMO (International Maritime Organisation), 1996. Code of the Investigation of Marine Casualties and Incidents, FSI 5/10/2.

IMO (International Maritime Organisation), 2002. Guidelines for Formal Safety (FSA) for Use in the IMO Rule-Making Process. The MSC/Circ. 1023, MEPC/Circ. 392.

Jens, Rasmussen, 1997. Risk management in a dynamic society: a modelling problem. Safety Science 27 (2/3), 183–213.

Kaplan, S, Haimes, Y.Y., Garrick, B.J., 2001. Fitting hierarchical holographic modelling into the theory of scenario structuring and a resulting refinement to the quantitative definition of risk. Journal of Risk Analysis 21 (5), 807–819.

Kawka, N., Kirchsteiger, C., 1999. Technical note on the contribution of sociotechnical factors to accidents notified to MARS. Journal of Loss Prevention in the Process Industries 12, 53–57.

Konovessis, D., Vassalos, D., 2008. Risk evaluation for RoPax vessels. Proceedings IMechE Part M: Journal of Engineering for the Maritime Environment 222, 13–26.

Konstantinos, G., Ernestini, B.G., 2002. Assessing the risk of pollution from ship accidents. Emerald (MCB), Disaster Prevention and Management An International Journal 11 (2), 109–114.

Krippendorf, K., 1980. Content Analysis: An Introduction to its Methodology. Sage Publications, Beverly Hills, CA.

Laracy, J.R., 2006. A systems theoretic accident model applied to bio-defense. Defense & Security Analyst 22 (3), 301–310.

Lars Harms-Ringdahl, 2004. Relationships between accident investigations, risk analysis, and safety management. Journal of Hazardous Materials 111, 13–19, 2004.

Larsson, P., Dekker, S.W.A., Tingvall, C., 2010. The need for a systems theory approach to road safety. Safety Science 48, 1167–1174.

Le Blanc, L.A., Hashemi, R.R., Rucks, C.T., 2001. Pattern development for vessel accidents: a comparison of statistical and neural computing techniques. Expert Systems with Applications 20 (2), 163–171.

Leveson, N., 2004. A new accident model for engineering safer systems. Safety Science 42, 237–270.

Li, K.X., Wang, Y., Min, J., 2009. Quantitative analysis of materiality in marine insurance. Maritime Policy & Management 36 (5), 437–455.

Merrick, J.R.W., Singh, A., 2003. Propagation of uncertainty in a simulation-based maritime risk assessment model utilizing bayesian simulation techniques. In: Chick, S., Sanchez, P.J., Ferrin, D., Morrice, D.J. (Eds.), Proceedings of the 2003 Winter Simulation Conference.

Mullai, A., 2004. A risk analysis framework for maritime transport of packaged dangerous goods. In: Brindley, C. (Ed.), Supply Chain Risk. Ashgate Publishing Company, Aldershot, Hampshire, England, pp. 130–159.

Mullai A. A Risk Analysis Framework for Maritime Transport of Packaged Dangerous Goods: A Validating Demonstration. PhD thesis, Lund University, Lund Institute of Technology, Department of Industrial Management and Logistics, Lund, Sweden, 2007; Volume I & II, pp. 1–210, 1–330.

Mullai, A., Larsson, E., 2008. Hazardous materials incidents: some key results of a risk analysis. WMU Journal of Maritime Affairs 7 (1), 65–108.

Mullai, A., Paulsson, U., 2002. Oil Spills in Öresund – Hazardous Events Causes and Claims Report on the SUNDRISK Project. Lund University Centre for Risk Analysis and Management (LUCRAM), Department of Industrial Management and Engineering Logistics, Lund University, Lund, Sweden.

Mullai, A., Larsson, E., Norrman, A., 2009. A study of marine incidents databases in the Baltic sea region. In: Weintrit, A. (Ed.), Marine Navigation and Safety of Sea Transportation. Taylor & Francis Group, London, pp. 247–253.

Nikolaos, P., Ventikos, Harilaos, N., Psaraftis, 2004. Spill accident modeling: a critical survey of the event-decision network in the context of IMO's formal safety assessment. Journal of Hazardous Materials 107, 59–66.

OECD (Organisation for Economic Co-operation and Development), 2004. The Chemical Accident Risk Assessment Thesaurus (CARAT), Available at http://www1.oecd.org.

Özgecan, S., Ulusçu, 2009. Birnur Özbaş Tayfur Altıok, and İlhan (Risk analysis of the vessel traffic in the strait of Istanbul). Risk Analysis 29 (10).

Psarros, G., Skjong, R., Strandmyr Eide, M., 2010. Under-reporting of maritime accidents. Accident Analysis and Prevention 42, 619–625.

Reason, J., 1990. Human Error. Cambridge University Press, Cambridge.

Reason, J.T., 1997. Managing the Risks of Organizational Accidents. Ashgate, Aldershot, UK.

Sii, H.S., Ruxton, T., Wang, J., 2001. A fuzzy-logic-based approach to qualitative safety modelling for maritime systems. Reliability Engineering and System Safety 73 (1), 19–34.

SMA (Swedish Maritime Administration). Summary of marine casualties and near-accidents involving Swedish merchant and fishing vessels 2007. Report Series B 2008-10, 2008. Available at: http://insjo.test.interlan.se, (accessed on December 2009).

Trucco, P., Cagno, E., Ruggeri, F., Grande, O., 2008. A Bayesian belief network modelling of organisational factors in risk analysis: a case study in maritime transportation. Reliability Engineering and System Safety 93, 823–834.

Van Drop, J.R., Merrick, J.R.W., Harrald, J.R., Mazzuchi, Th.A., Grabowski, M., Risk, A., 2001. Management procedure for Washington state ferries. Risk Analysis 21 (1).

Vanem, E., Skjong, R., 2006. Designing for safety in passenger ships utilizing advanced evacuation analyses – a risk based approach. Safety Science 44, 111–135.

Yin, K.R., 1994. Case Study Research: Design and Methods, Second Ed. Saga Publication.