Reliability Engineering and System Safety 176 (2018) 174–186

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Maritime navigation accidents and risk indicators: An exploratory statistical analysis using AIS data and accident reports

Rolf J. Bye[a,*], Asbjørn L. Aalberg[b]

[a] NTNU – Norwegian University of Science and Technology, Norway
[b] Safetec Nordic, Norway

ARTICLE INFO

Keyword:
Navigation accidents
Risk indicators
Correspondence analysis
Logistic regression

ABSTRACT

This paper presents the results of statistical analyses of maritime accidents data and AIS data from Norwegian waters, to identify conditions that are associated with navigation related accidents (groundings and collisions) and could be used as risk indicators. Vessels involved in accidents reported in the accident database of the Norwegian Maritime Directorate (NMA), have been traced in historical AIS records, and data related to each ship have been transformed into variables. These variables are related to the behavior of the ship in front of the accident (e.g. nautical miles sailed, hours in operations, number of port calls etc.), technical and organizational conditions (ship categories, flag state, age, gross tonnage, Paris MoU ratings etc.) and the area where the accident occurred (number of vessels in the area, port calls in the area, nautical miles in the area etc.). Both the AIS data and the data from the NMA accident database have first been analyzed using correspondence analysis (categorical variables), *F*-tests (continuous variables), and then combined in a multivariate logistic regression model with "navigation accidents" and "other accidents" as dependent variables. The model is a strong predictor for whether the accident is navigation-related or not. Specifically, some vessel types, less vessel length, poor visibility condition, and a flag of convenience increased this probability.

1. Introduction

Mazaheri et al. [1] have argued that most models for ship-grounding accident does not fully utilize available evidence, but tend to predominately rely on expert judgment and accident statistics. They have called for a more evidence based model development, using different data sources in order to develop more real-life scenarios. This study is an exploratory statistical analysis of accident-related data, conducted to support an empirically based development process of Bayesian Belief Network (BBN) risk models for Norwegian waters [2].

Statistical analysis of data associated with navigation accidents (groundings and collisions) have been carried out in order to support the conceptualization of relevant nodes in the BBN models for groundings and collisions, and to eventually develop data that could be used when quantifying the models. The objective was to identify conditions that may be used as risk indicators, i.e. measurable variables that can be used to describe risk, and/or states of the nodes in the model [3,4].

The development of the BBN models was not entirely data driven, such as in the work of e.g. Mullai and Paulson [5], but based on an iterative process of comparing results from data analysis with major accident theories (see e.g. [6,7]), previous analysis of navigation accident and developed risk models (se e.g. [8,9,10] for model reviews). The combination of exploratory data analysis, theory, and expert judgment represents a development process that is more in line with the recommendations of Mazaheri et al. [1].

The BBN models that was developed through the project was supposed to only predict accident probabilities without including consequences. This goal was also reflected in this present analysis, where the aim was to identify factors that might influence the probability of accidents.

The identification of possible nodes in the models, and relevant historical data to be analyzed, was guided by the accident scenario model of Rasmussen [11], which conceptualize the relation between "sharp end" conditions (staff and work), "blunt end" conditions (government, regulators, company, management), and environmental "stressors" (e.g. financial pressure, technology, educational level etc.).

Explorative analyses used in the model development included this present work, a qualitative inductive analysis of a sample of accident investigation related to navigation accidents [12], analyses of accident data by the use of multiple logistic regression [13], comparison of normalized accident frequencies by the use of exposure data calculated from AIS data [14]. The development process of the models has been described in Haugen et al. [15].

* Corresponding author.
E-mail addresses: rolf.johan.bye@safetec.no, rjb@safetec.no (R.J. Bye).

https://doi.org/10.1016/j.ress.2018.03.033
Received 8 October 2017; Received in revised form 11 February 2018; Accepted 29 March 2018
Available online 21 April 2018

Data sources for this analysis have been the accident database of the Norwegian Maritime Authority (NMA) and processed AIS (Automated Identification System) data. AIS data were used to construct indicators that could reflect activity of the ship involved in the accident, such as e.g. speed, number of port calls, hours since last port departure, course alteration, and so forth, general ship data (ship category, age, flag), and vessel traffic in the area where the accident occurred.

The use of both reported accident data and AIS data has partly been done in order to improve the reliability (and the validity) of the data. Research have shown that underreporting and missing data is quite widespread within the maritime industry [16,17,18]. Due to this, e.g. Hänninen [19] calls for the use of additional data sources, such as e.g. results from Port State Controls

The use of AIS data in this present analysis has also been motivated by previous explorative accident analysis based on the use of AIS data [20–24], and an ambition to utilize the possibilities of using AIS data to construct relevant and reliable accident data variables that cannot be obtained by established accident report records.

The main objective for developing BBN models was to establish risk models for Norwegian waters, meaning the whole exclusive economic zone. Due to this, only accident and AIS data from these areas were obtained.

There could be several characteristics with the Norwegian sea areas that may reduce the generic validity and transferability of our findings to other sea areas. Characteristics that might represent area specific conditions could be e.g. the compositions of vessel types in use, fairways within dens coastal archipelago, the extent of transit traffic from Russian artic areas, liberal cabotage regulations, and use of open register in coasting.

The scientific contribution of this present work is that it contributes to knowledge regarding navigation accident, by the findings in our analysis and by comparing these findings with previous work. Secondly, the work represents one of relatively few analyses which utilize AIS data in order to construct accident related variables. This present work might contribute to a further development of the development of AIS based variables.

The remainder of this paper is organized as follows: Section 2 presents a literature review. Section 3 is a description of the method, including a description of the variables. Section 4 is a presentation of the results of the analysis. In Section 5 the findings will be discussed, with reference to previous research. Sections 6 presents the final conclusions.

2. Previous empirical research on maritime accidents

More than a decade ago, Hetherington et al. [25] conducted a meta-study to identify *risk factors* associated with human actions. Their work was based on a review of 30 papers that based the discussion on empirical data from the maritime industry. They concluded that the main factors were related to fatigue, automation, situational awareness, communication, decision-making, teamwork and health, and health and stress.

In this paper, a reviewed 32 journal publications and conference papers between 2000 and 2017 has been conducted. This is papers that use samples of empirical data from ship accidents (more than one single accident) to identify risk factors associated with maritime accidents. The methods employed by these studies may be divide into three groups: studies that rely on (1) qualitative inductive analysis of accident reports and/or investigations, (2) explorative statistical analysis of incident related data (both qualitative and quantitative), and (3) a combination of the qualitative inductive analysis and explorative statistical analysis.

Examples of studies based entirely on inductive analysis include Macrae [26], Kum and Sahin [27], and Nilsen et al. [12]. Macrae [26] re-analyzed 30 accident reports related to grounding and collision. He found that there was a common pattern in grounding and collision accidents. He concluded that groundings are associated with “planning failure” and collision with “inadequate watch.” Kum and Sahin [27] found that accidents in arctic regions were associated with inadequate quality and extent of training. Nilsen et al. [12] identified five conditions that occurred frequently in relation to groundings and collisions, based on the analysis of 95 accident investigations worldwide. These conditions were denoted as “pilot-related incidents,” “inadequate use or non-use of navigational aids,” “inadequate navigational attention,” “no dedicated lookout,” and “involvement of fishing vessels in collisions.”

Goerlandt et al. [20] have analyzed winter navigation in the Northern Baltic Sea area by using visual data mining related to a sample of 45 accidents. This qualitative approach was based on data from different sources, including, for example, accident databases, AIS data records, and weather data. They found that groundings occurred in port areas and archipelagic waterways, and that collision occurred in open sea areas and archipelagic waterways. Further, they found that the accidents occurred mainly during moderate to excellent visibility conditions. Collisions occurred when the vessels had low speed compared to the average speed of vessels. They also found that the impact area in collisions could vary with the type of vessel operation. Still further, they found that smaller cargo vessels in archipelagic waterways dominate groundings.

Wróbel et al. [28] have conducted an inductive analysis of 100 maritime accidents, using the framework of Human Factor Analysis and Classification System (HFACS), to identify human contribution to different accident scenarios that might be reduced or increased due to use of unmanned vessels. Conditions increasing the likelihood of a grounding accidents which was assumed to be eliminated by computerization, was e.g. distraction of the officers and improper passage plan. Their analysis of 19 collisions showed that the most frequent associated conditions was insufficient “look –out”, low visibility conditions and proximity to a port.

Explorative statistical analysis of accident data has also been conducted, using different quantitative methods. Zhang et al. [29] have carried out statistical analysis of accidents data limited to the Tian Port area in order to develop a BBN model for accident consequences. Their descriptive statistics indicated e.g. that accidents were associated with the position/area where the accident occurred, and types of ships. Banda et al. [30] have used descriptive accident statistics, collated with expert judgment in their risk analysis of winter navigation in Finnish sea areas. In their analysis they combined accident data with historical data of ice conditions. Most of the accidents were associated with navigation.

As a part of their work on developing a model for predicting accident frequencies in different zones of the Aegean Sea, Ventikos et al. [31] have conducted a study of historical data. They found that most of the vessels involved in accidents was older than 25 years, and that 60% of the accident involved RO-RO and bulk carriers.

Li and Wonham [32] have used vessel data from Lloyd´s Register of Shipping to e.g. calculate accidental total loss rates for different registry countries. They found that the worst group in terms of total loss rate is the open registry countries, except for two registries. Yip [33] has used negative binominal regression models to analyses accident data from 2012 marine accident in Hong Kong port, to identify factors associated with personal injuries and fatalities. He constructed the following four variables: “port of registry”, “type of vessel”, “type of accident”, “type of waterway” and “underway” (phase of operation). He found that the potential for accidents with injuries increase with specific vessel types (e.g. passengers ship, sampam,[1] and pleasure vessels), and types of accidents (e.g. structural damage, fire/explosion, and capsizing). Fatal accident was associated with fishing vessels and “miscellaneous vessels”.

Eleftheria et al. [34] have compared ship accident frequencies (normalized by annually operating ships) of different categories of vessels, using a sample of 4,572 accidents word wide. They found that the most accident-prone vessels were general cargo ships, bulk carriers, car carri-

[1] A flat bottomed Chinese wooden boat, used as fishing boats or for transportation in coastal areas.