

Assessing Grounding Frequency using Ship Traffic and Waterway Complexity

Arsham Mazaheri, Jakub Montewka, Pentti Kotilainen,
Otto-Ville Edvard Sormunen and Pentti Kujala

*(Aalto University, School of Engineering, Department of Applied Mechanics,
Marine Technology, P.O. Box 12200, FI-00076 AALTO, Espoo, Finland)*
(E-mail: arsham.mazaheri@aalto.fi)

Ship traffic is the factor that presents in almost all of the existing grounding risk models. It is considered to be one of the main factors affecting the expected frequency of ship groundings. This is mostly accepted by experts as common sense. However, there is no research available on the actual dependency between ship traffic and grounding accidents. In this paper, we conduct a study aimed at determining the statistical dependency between the density and distribution of traffic, the number and frequency of grounding accidents and the dependency between the complexity of waterways and an actual accident. For this purpose we utilise statistical analysis of maritime traffic, obtained from Automatic Identification System (AIS) data and grounding accidents, enhanced with the expert elicitation techniques delivering the waterway complexity index. The sea area under investigation is the Gulf of Finland. The results show statistical dependency between frequency of grounding and waterway complexity as well as the traffic distribution. However, the study does not reveal any significant dependency between grounding and traffic density.

KEYWORDS

1. Ship Grounding. 2. Waterway Complexity. 3. Ship Traffic. 4. Grounding Frequency.

Submitted: 18 October 2013. Accepted: 17 July 2014. First published online: 7 August 2014.

1. INTRODUCTION. There are many factors mentioned in the literature as affecting the frequency of ship Grounding Accidents (GA) (Jebson and Papakonstantinou, 1997; Fujii et al., 1974; Samuelides et al., 2009, Briggs et al., 2003; Lin, 1999; Quy et al., 2006; Martins and Maturana, 2010; Brown and Haugene, 1998; Praetorius, 2012; Akhtar and Utne, 2014; Amrozowicz et al., 1997; Pedersen, 1995; Fowler and Sörgård, 2000; Kite-Powell et al., 1999). Most of them are based on experts' opinions. Moreover, their effects on the expected frequency of GAs have rarely been examined (Mazaheri et al., 2013). In theory, it is possible to list hundreds of factors that may affect the frequency of GA (Jebson and Papakonstantinou, 1997), (Figure 1); nevertheless, if any model assessing the grounding frequency is supposed to

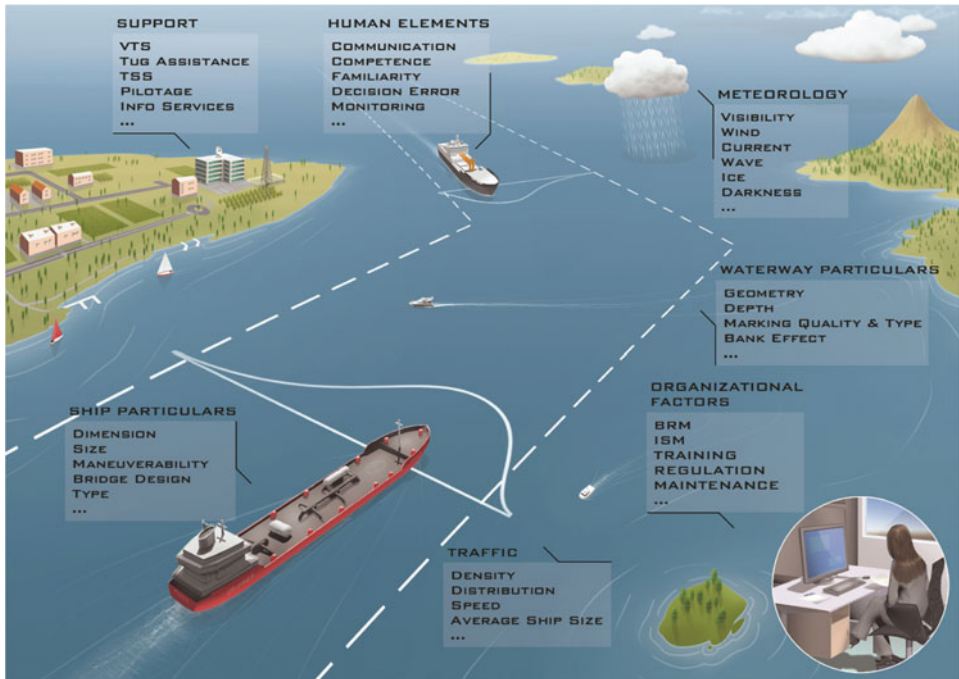


Figure 1. Theoretical factors that can affect the expected frequency of ship grounding accident.

be used for decision making purposes, the relationship of these factors with GA should be analysed more accurately.

In [Table 1](#) we list all the factors which are believed to affect GAs, which have been studied in the literature so far. However, the effect of ship traffic on GAs, has not been given attention yet, despite its presence in almost all of the existing grounding risk models, see [Mazaheri et al. \(2013\)](#); [Fowler and Sjørgård \(2000\)](#); [Otto et al. \(2002\)](#) and [Montewka et al. \(2011\)](#). This situation is mostly accepted as common sense by experts; e.g. [Pedersen \(1995\)](#) or simply as a result of generalisation from ship-ship collision assumptions; e.g. [Fujii et al., \(1974\)](#). One of the reasons behind the belief of the dependency of ship traffic and GA is that people assume when traffic is denser, the likelihood that the ships have to alter their courses to avoid collision and eventually ending up grounded is higher ([Jebsen and Papakonstantinou, 1997](#)). A contrary belief though exists as the less dense traffic areas may be the more frequent grounding areas, since pilots and navigators may be less experienced in navigating in those areas and thus more likely to end up grounded. However, there is no study available for supporting or rejecting either of these beliefs. Moreover, it is possible that researchers are convinced that navigating a ship is more difficult when traffic is denser; however the question is whether difficulty due to traffic necessarily increases the likelihood of GA.

One problem that may rise when beliefs without solid foundation become commonly accepted is that the models developed for analysing the phenomenon and assessing the frequency of the accident may not represent the reality, ([Goerlandt and Kujala, 2014](#); [Sormunen et al., 2014](#); [Goerlandt et al., 2014](#); [Hänninen et al., 2014a](#)). Moreover if the model is used for risk management and decision making it might

Table 1. Studied factors in the literature that affect the expected frequency of ship grounding accident.

Category	Affecting Factor	Reference
Meteorological Conditions	Current	Briggs et al. (2003)
	Darkness	Lin (1999); Fujii et al. (1974)
	Tide	Lin (1999)
	Visibility	Kite-Powell et al. (1999); Jebsen and Papakonstantinou (1997)
	Wave	Briggs et al. (2003)
	Wind	Kite-Powell et al. (1999); Jebsen and Papakonstantinou (1997); Briggs et al. (2003)
Waterway Particulars	Depth	Quy et al. (2006); Kite-Powell et al. (1999); Jebsen and Papakonstantinou (1997)
Vessel Particulars	Geometry	Briggs et al. (2003); Lin (1999)
	Age	Samuelides et al. (2009)
	Size	Kite-Powell et al. (1999); Jebsen and Papakonstantinou (1997); Samuelides et al. (2009)
Traffic	Type	Kite-Powell et al. (1999); Jebsen et al. (1997)
Human & Organizational Factors	Speed	Quy et al. (2006)
	Human Elements	Lin (1999); Martins and Maturana (2010); Amrozowicz et al. (1997)
	Fatigue	Akhtar and Utne (2014)
	Operator Skill	Kite-Powell et al. (1999); Jebsen and Papakonstantinou (1997)
	ISM	Lin (1999); Jebsen and Papakonstantinou (1997)
	Organizational & Human Elements	Amrozowicz (1997); Brown et al. (1998); Lin (1999); Martins and Maturana (2010); Jebsen and Papakonstantinou (1997)
Support	VTS	Praetorius (2012)

result in ineffective or even wrong risk control options (see for example Montewka et al., 2014, 2014a; Hänninen et al., 2014).

Therefore, the aim of this article is twofold: firstly we study the statistical dependency between ship traffic and the frequency of GA; secondly we assess the relation between the waterway complexity and the frequency of GA.

For this purpose, GAs are subdivided into numbers and frequency of accident. Frequency of GA is used here as the number of groundings over the number of ship transits. Subdividing ship traffic into density and distribution, we test whether there is any statistical dependency between the number of GAs and density and distribution of ship traffic. Subsequently, we discuss the strength and weakness of the possible dependency. The dependency between the complexity of the waterways and the frequency of GAs is investigated in the same way.

The list given in Table 1 shows that waterway complexity, as the difficulty of a waterway in which to safely navigate a ship, has not yet been studied with regard to GA. Contrary to ship traffic, waterway complexity has not been used in any available risk models for GA so far, despite the commonly accepted belief that GA is a location-dependent accident. However, since waterway complexity is inherent in the design and geographical location of a waterway, it can be seen as a valid factor for assessing the frequency of GA; thus it is worth further analysis.

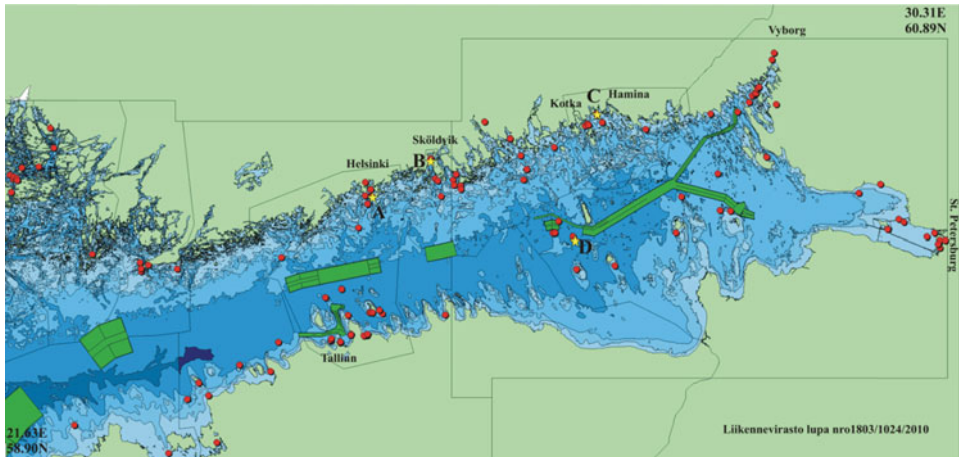


Figure 2. Grounding accidents happening in the GoF (1989–2010) shown in red dots. Green areas show Traffic Separation Schemes.

In this paper, we focus on the Gulf of Finland (GoF), due to its complex-to-navigate archipelagos and intense shipping volume. Besides, the area is one of the shallowest bodies of water in the world (Kuronen et al., 2008), which has resulted in grounding being the most frequent accident type in the area (Kujala et al., 2009).

The remainder of this paper is organised as follows: the data used for the analysis is presented in Section 2. The applied methodology and the implemented algorithms for data analysis are provided in Section 3; results of the data analysis are given in Section 4; followed by the discussion of the results and their limitations in Section 5. Finally, the paper is concluded in Section 6.

2. ACCIDENT AND TRAFFIC DATA. Two different sources of data are used in this study. First is Helsinki Commission (HELCOM) statistics regarding ship accidents that occurred in the Baltic Sea between the years 1989 and 2010; second is HELCOM Automatic Identification System (AIS) data of marine traffic in GoF for the year 2010.

2.1. HELCOM statistics. The accident data consist of inputs such as the date and time, geographical coordinates, and type of accident. From among all the available fields, the only input values that have been utilised in this research are the type and the location of the accident.

The dataset was first filtered regarding the type of the accident, and all the accidents that were recorded as “grounding” were chosen. Secondly, the dataset was filtered to limit the data to those accidents that have happened in the GoF, limiting the study area to 21–63° E and 30–31° E longitude, and to 58–90° N and 60–89° N latitude. Consequently, 123 records of GAs which occurred in this area between the years 1989–2010 were detected (Figure 2). The dataset was then filtered to remove those accidents that have erroneously reported as not having occurred near shoals or shallow water areas in the GoF or deep inland areas. As the result, 112 cases were retained and are used in Section 4 for the statistical analysis. The characteristics of the retained cases are shown in Figure 3. It can be seen from Figure 3 that 49 accidents

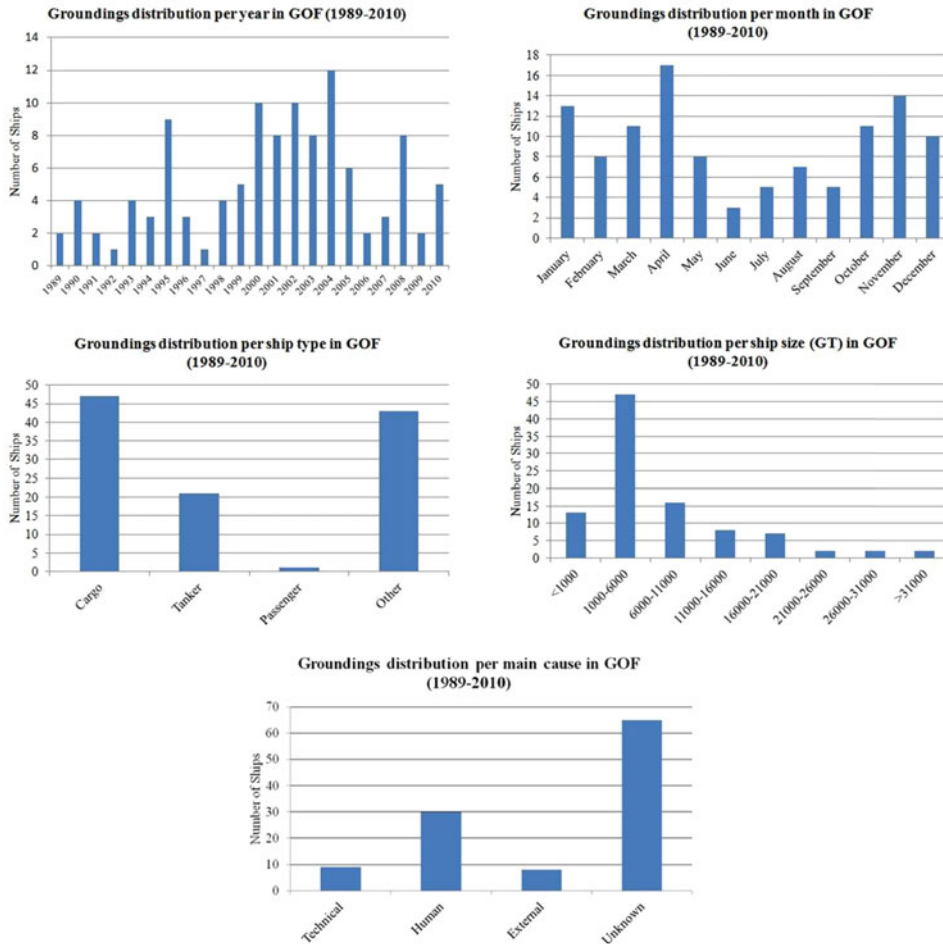


Figure 3. Characteristics of 112 groundings accidents in GoF between years 1989 and 2010 based on HELCOM statistics.

have occurred in winter (January till April) and 63 cases in summer. This will be used later in Section 4.3 for seasonal analysis of the data.

2.2. *HELCOM AIS-database.* To extract the properties of marine traffic in the GoF, we utilise AIS data, recorded on average every third minute. The input values of the AIS records, utilised here, are timestamps and geographical locations of the ships. Before using, the AIS data needs to be cleaned, filtered, and sorted. For that, some thresholds are defined for the fields of an AIS record, and the AIS records are filtered based on the defined thresholds. For instance for the speed, the lower boundary of 2 knots (kt) is set in order to remove the stationary vessels, and upper boundary of 50 kt is set since no vessel can travel faster than 50 kt (van Dorp and Merrick, 2009). Thereafter, the AIS records of a single vessel are sorted into separate journeys if a time difference of more than 30 minutes is detected between two sequential AIS records of a single vessel (Goerlandt and Kujala, 2011). Afterwards, another threshold is defined as the maximum distance allowed between two sequential transmitted AIS records

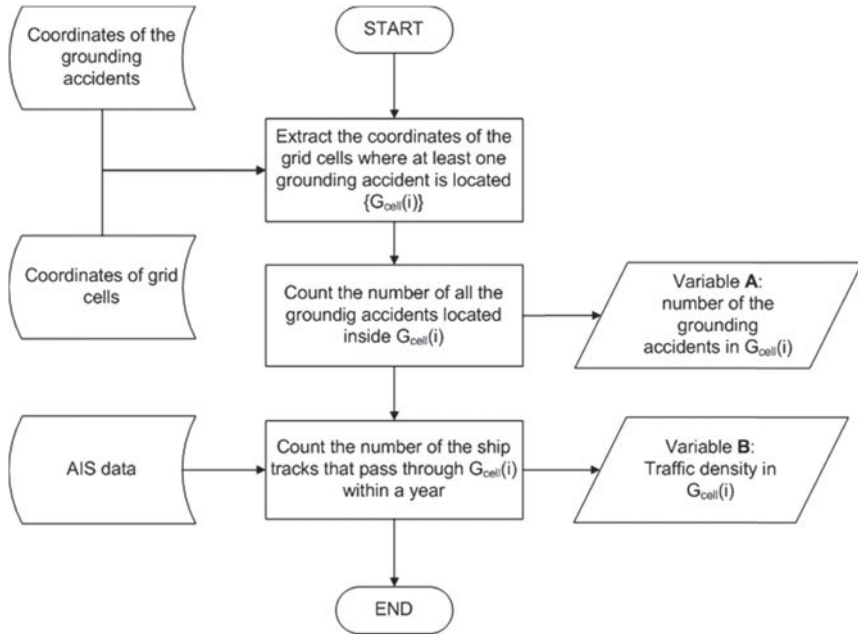


Figure 4. Algorithm to extract traffic density from AIS data.

from a single vessel, given the maximum speed that a vessel can travel. This was to remove the recording errors or location sensing errors that may appear during the data collection and transmission process. Thus, given the instant speed of the vessel and the time of the AIS records, any AIS record that has been transmitted within a larger distance of the defined threshold from either of its previous or following AIS records was detected as faulty and thus removed from the records (van Dorp and Merrick, 2009). The filtered AIS data is then used in Section 4 for the statistical analysis.

3. **METHODOLOGY.** The methodology has three phases: 1. Defining algorithms to extract the characteristics of ship traffic from the AIS data; 2. Quantifying the difficulty of navigating a ship in a waterway (i.e. waterway complexity); and 3. Implementing statistical analysis.

3.1. *Traffic Characteristics.* In the existing literature on grounding risk modeling, ship traffic is defined in two ways as density and distribution. Traffic density is defined as the number of ships per unit area within a time window (Mazaheri and Ylitalo, 2010). As we are dealing with large enough samples of the AIS data from year 2010, the time window for this study is defined as a year. The algorithm to estimate the traffic density based on the above definition is shown in Figure 4. In general, the GoF is divided into grid cells of size one by one nautical mile. Thereafter, the number of ship tracks over a year that passed through the cells with at least one GA is counted as traffic density. Although the records of the ships that are in stationary positions are removed from the AIS database beforehand, the speed of the traffic is taken into account by the way that traffic density is extracted. The number of GAs located inside

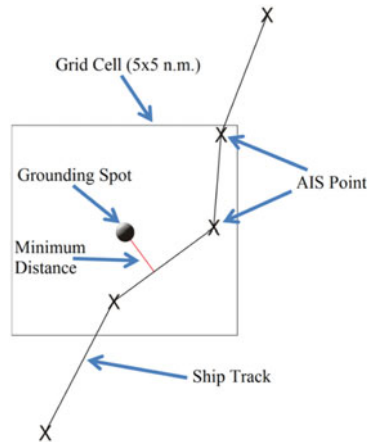


Figure 5. Minimum distance between a ship and location of a grounding accident.

each grid cell is also counted into another variable. These variables are used in Section 4.1 to find the statistical dependency between traffic density and number of GAs.

Traffic distribution is defined as the lateral distribution of ship tracks over a waterway. To extract the traffic distributions over a waterway with respect to the location of GAs, we have extracted the distribution of the minimum distances of ship tracks from GAs in the GoF (Figures 5 and 6). Thus, grid cells of size five by five nautical miles are created around the location of each GA in such a way that the location of the GA is located exactly in the middle of the grid. Then for each cell, the minimum distances of the ship tracks from the GA's location are estimated and taken as traffic distribution (see Figure 7). The number of GAs located inside each grid cell is also counted into another variable. These variables are used in Section 4.1 to find the statistical dependency between traffic distribution and number of GAs based on the first and second moments of the obtained distributions.

To find the effect of the grid sizes used on the results, an additional grid size of three by three nautical miles is used to repeat the studies of traffic density and distribution. Besides, to find the effect of winter season, the study on traffic distribution is repeated by separating the winter and summer traffics. In Sections 4.2 and 4.3 the results of these additional studies are presented respectively and discussed by comparing with the main analyses.

3.2. Waterway Complexity. Waterway complexity is defined as the level of difficulty in handling a ship safely through a waterway. For analysing the statistical dependency between GAs and waterway complexity, a location-dependent, semi-quantitative index, called a Waterway Complexity Index (WCI), is defined. WCI ranks fairways with regard to the difficulty of ship handling. Currently, there is no study available in the literature that shows the parameters that may affect the complexity of a waterway. Thus, WCI is here quantified based on the experience of the local pilots. Since navigating in a fairway is related to human activity, and qualitative methods are often successfully used for such studies (Robson, 2008) semi-structured interviews as a qualitative method is used here to acquire knowledge of the local pilots.

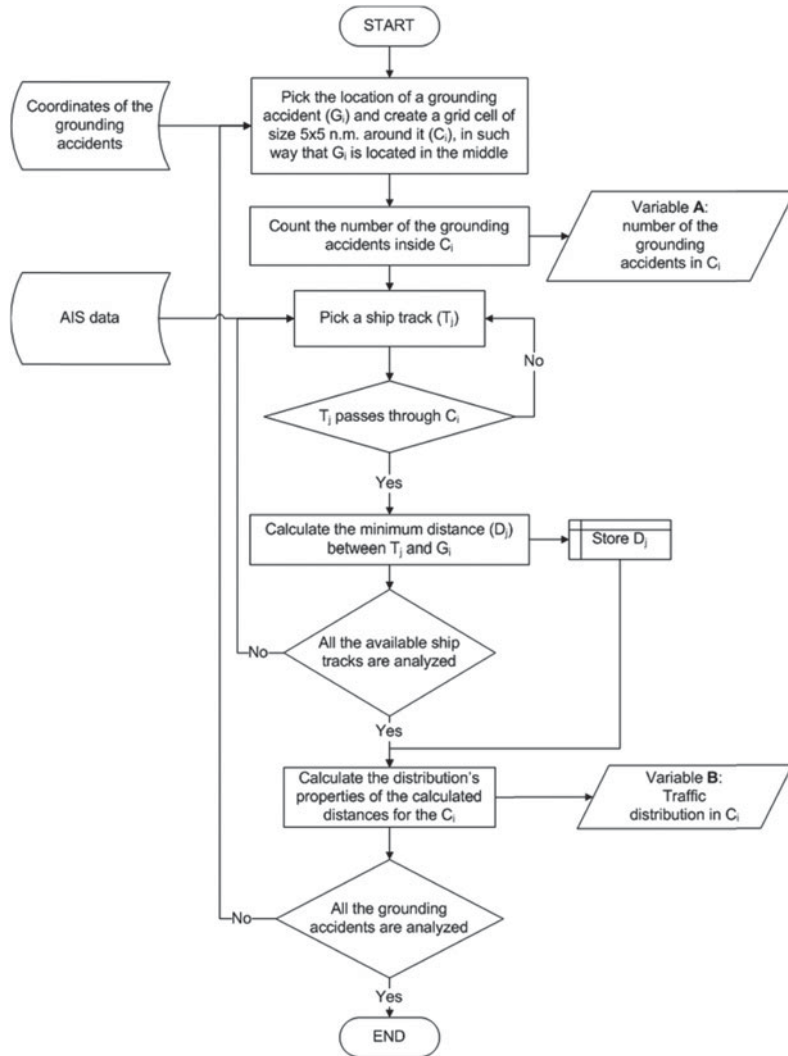


Figure 6. Algorithm to extract traffic distribution from AIS data.

WCI for this study is defined for five waterways (harbour entrance fairways) in the GoF, leading to the following harbours: Helsinki (West and South harbours), Sköldvik, Kotka, and Hamina (see Table 2 and Figure 2). The chosen fairways differ in their geometry and bottom topography. They are also intended to be used for various ship types and sizes. From the operational point of view, these fairways fall within two pilot stations as Helsinki – ships are piloted to Helsinki and Sköldvig – and Kotka – ships are piloted to Kotka, Hamina and Sköldvig. This overlapping in operation is good for the comparability of the gathered results for the studied fairways.

The focus of the study was mainly on fairway dimensions and geometry. The chosen perspective is based on ship navigators' experience. The experience varies between individuals and in regard to how frequently the fairway is used in general. The aim of the interviews was to make a numerical WCI for typical ships navigating

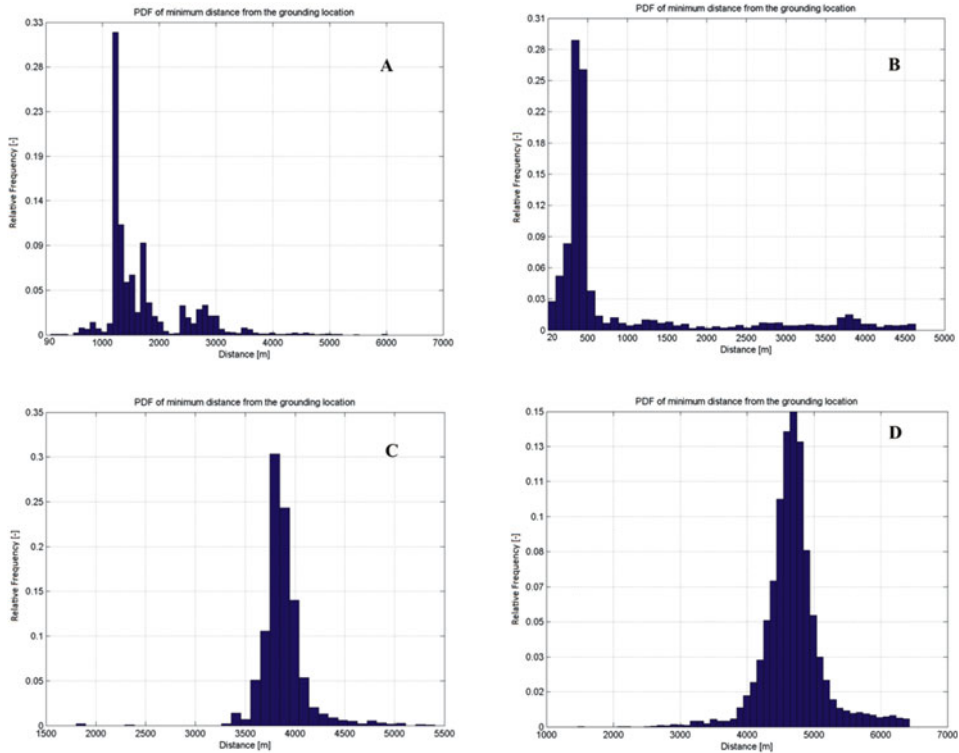


Figure 7. Extracted distributions of minimum distances between ship tracks and four grounding locations marked in Figure 2 with yellow stars.

Table 2. Waterway complexity graded from 4 (the most difficult) to 10 (the easiest) by the interviewed pilots. WCI is scaled based on the Average values and from 0 (the most difficult) to 1 (the easiest).

Waterway	Pilot 1	Pilot 2	Pilot 3	Pilot 4	Pilot 5	Pilot 6	Pilot 7	Pilot 8	Pilot 9	Pilot 10	Pilot 11	Pilot 12	Average	WCI
Kotka	–	–	8	–	–	9	–	–	8	10	8	8	8.5	0.75
Sköldvik	8.5	8	9	10	7	9	9	8	–	–	–	–	8.6	0.76
Helsinki-West	7	9	8.5	9	8.5	–	–	8.5	–	–	–	–	8.4	0.74
Helsinki-South	6	8	8	8	5	–	–	8	–	–	–	–	7.2	0.53
Hamina	–	–	7	–	–	8	–	–	7.5	8.5	8	8	7.8	0.64

in the studied fairways. Twelve experienced maritime pilots were chosen for the interviews. Their pilot experience varied from seven to 25 years. Each maritime pilot has a set of fairways where he operates with various types of ships, and thus he was interviewed only based on those fairways (Table 2).

The interviews followed a six point routine:

1. Describing for the interviewee what is to be studied and why.
2. Discussing fairway quality in general. This includes the complexity in navigation regarding the turns, width of the fairway, and water depth.

3. Asking about quality-based observations of the interviewee about the fairways where he serves as a pilot. This includes the difficulties that the interviewee has met in the fairways.
4. Interviewee is asked to compare the fairways to rank them according to their difficulty of navigating.
5. Interviewee is asked to give a holistic grade of the difficulty of navigation in the waterway for typical ships navigating there today.
 - a. The grading was based firstly on the comparison to the other fairways where the interviewee is piloting, and secondly to the severity of the difficulties that are met in the fairway (see Table 2).
 - b. The scale of the grading was from 4 to 10, as this scale has been widely used earlier at schools in Finland, and Finnish pilots feel comfortable in using it.
 - c. Grade 4 should be given to a fairway which a pilot considers as being extremely difficult for navigation, and grade 10 to an ideal fairway.
6. Step 5 is repeated to cover all the studied fairways where the interviewee is piloting.

To formulate WCI the results of the grading were expressed on a scale between 0 (the most difficult) and 1 (the easiest), see Table 2. For this purpose, linear opinion pooling was used with equal weights for all the pilots. This method of combining multiple expert opinions into one is known for its simplicity and robustness (O'Hagan et al., 2006). The results gained during the interview process revealed that the pilots considered the following factors, related to a ship and a waterway, as important when assessing the complexity of a waterway:

- Ship draught and size in relation to the available space.
- The need for reduction of ship speed under certain circumstances, e.g. in the presence of two-way traffic.
- Width of the waterway, especially when two-way traffic is allowed.
- Number of turns and the magnitude of course alteration.
- The width of the waterway immediately following the turn.
- Fairway marking, especially in the areas which require increased attention of a pilot, e.g. due to significant course alteration.

Some of the factors related to ship and waterway geometry have been identified in the literature, as listed in Table 1. However to determine the complexity of a waterway, the joint effect of the factors needs to be recognised and the specific circumstances triggering such joint effect needs to be defined. Local experts, as shown in this study, can effectively do the latter.

3.3. *Statistical Analysis.* The existence of possible statistical dependency between ship traffic and GAs as well as WCI and GAs is studied with Pearson (r) correlation coefficient and Spearman (ρ) and Kendall (τ) rank correlation coefficients. The reliability of the resulting coefficients then tested with χ^2 – goodness-of-fit test assuming the null hypothesis (H_0) as “zero-correlation” against the alternative hypothesis (H_1) of “non-zero correlation” using significance level of 95%.

Although the correlation coefficients used can somehow reveal the possible statistical dependencies between the two variables, one should be aware about their limitations. For instance, the Pearson coefficient is defined assuming the data are

Table 3. Correlations between number of groundings and traffic, with grid sizes 1×1 nm for traffic density and 5×5 nm for traffic distribution. H_0 is Zero-correlation.

Traffic Property (D)	Coefficient	Correlation Value	P-Value	Accepted Hypothesis ($\alpha=0.05$)
Density	r	0.0045	0.9654	H_0
	ρ	0.1102	0.2850	H_0
	τ	0.0908	0.2857	H_0
Distribution (mu)	r	-0.0804	0.4385	H_0
	ρ	0.0130	0.9008	H_0
	τ	0.0093	0.9091	H_0
Distribution (std)	r	0.1415	0.1713	H_0
	ρ	0.1458	0.1586	H_0
	τ	0.1187	0.1324	H_0

normally distributed, or the Spearman coefficient is recommended when both variables are ordinal variables, or one is ordinal and the other is a continuous variable (Lehman, 2005). Therefore, since these assumptions are not perfectly matched with the limitation of our variables, a Mutual Information (MI) test is also utilised in this study for comparison, whenever it is plausible and informative (Steuer et al., 2002, Peng et al., 2005). MI shows how much knowing about one of the variables can decrease the uncertainty of the other variable. Thus, by using the MI test, the uncertainty coefficient (U) is calculated here that determines how large a proportion of the uncertainty about the number and frequency of GAs can be decreased by observing the other variables (Theil, 1970).

4. RESULTS

4.1. *Grounding vs. Traffic.* The results show that there is no statistically significant dependency between the two variables of traffic and number of GAs. Table 3 shows the correlation coefficients between the number of GAs and the traffic density with the grid cells of size 1×1 nm and the traffic distribution with the grid cells size of 5×5 nm. As is seen, the null hypothesis (i.e. zero-correlation) cannot be rejected for traffic density and distribution at 95% significance level.

Nevertheless, the MI test on the same data shows that even though knowing about traffic density will not considerably increase our knowledge about the number of GAs, information about the distribution of traffic can significantly increase our knowledge in this regard (Table 6). Therefore, the distribution of the traffic is more informative than the traffic density to assess the expected frequency of GAs.

4.1.1. *Effect of Different Grid Sizes.* The extracted traffic density and traffic distribution were based on grid sizes of 1×1 nm and 5×5 nm respectively. For finding whether the size of grids can affect the results, the properties of ship traffic were extracted again from the same database and were analysed the same way, but using the grid size of 3×3 nm for both traffic density and distribution.

As is seen in Table 4, similar results for correlation coefficients are obtained using different grid sizes. However, the results of MI test show noticeable changes. Table 6 shows that U for traffic density is increased while for traffic distribution it either remained the same or decreased. Recalling that the size of the grids for traffic density and traffic distribution was increased for the former and decreased for the latter, this shows that the size of the grid has direct relation with the amount of

Table 4. Correlations between number of groundings and traffic, with grid size 3×3 n.m. H_0 is Zero-correlation.

Traffic Property (D)	Coefficient	Correlation Value	P-Value	Accepted Hypothesis ($\alpha=0.05$)
Density	r	-0.0163	0.8719	H_0
	ρ	0.0267	0.7923	H_0
	τ	0.0209	0.7943	H_0
Distribution (mu)	r	0.0813	0.5165	H_0
	ρ	0.0997	0.4285	H_0
	τ	0.0768	0.4349	H_0
Distribution (std)	r	0.2017	0.1043	H_0
	ρ	0.1924	0.1217	H_0
	τ	0.1468	0.1339	H_0

Table 5. Correlations between traffic distribution and number of groundings for summer and winter traffic in the GoF. H_0 is Zero-correlation.

Traffic Property (D)	Coefficient	Correlation Value	P-Value	Accepted Hypothesis ($\alpha=0.05$)
Distribution (mu) <u>Summer</u>	r	-0.2686	0.0864	H_0
	ρ	-0.2250	0.0813	H_0
	τ	-0.1728	0.0872	H_0
Distribution (std) <u>Summer</u>	r	0.1512	0.2446	H_0
	ρ	0.1803	0.1643	H_0
	τ	0.1370	0.1755	H_0
Distribution (mu) <u>Winter</u>	r	-0.0290	0.8447	H_0
	ρ	0.0276	0.8523	H_0
	τ	0.0097	0.9387	H_0
Distribution (std) <u>Winter</u>	r	-0.1409	0.3396	H_0
	ρ	-0.1415	0.3372	H_0
	τ	-0.1115	0.3272	H_0

information that can be retrieved from the data. Thus, the critical matter in such studies is to define the optimum size or even shape for a grid in order to deliver meaningful information.

4.1.2. *Effect of Season.* The GoF is often frozen in winter time, which affects the traffic pattern in the area for this period. Therefore, we run the study again only for traffic distribution and number of GAs by separating the summer and winter traffic, using grid size of 5×5 nm. The GoF was frozen in the first four months of the year 2010. Thus, for this study we assigned the data of ship traffic and GA of January till April to winter traffic and the rest to summer traffic.

The results for the correlation coefficients remain the same, showing no statistically significant dependency between the studied variables (Table 5). The uncertainty coefficient (U) for the first moment of the traffic distribution for summer traffic does not show remarkable changes either, while it is increased for winter traffic (Table 6). This may be explained as quite often the ships in winter move in convoy following an icebreaker, or navigate inside an ice channel that is opened by an icebreaker. Therefore, they normally pass a shoal at the same distance, resulting in the winter traffic being more informative regarding the expected value of minimum distance from

Table 6. Uncertainty Coefficient (U) for ship traffic (D) and grounding (Gr).

Traffic Property (D)	Condition	$U(Gr D)$
Density	Annual traffic 1×1 n.m.	30.4%
	Annual traffic 3×3 n.m.	58.9%
Distribution (μ)	Annual traffic 5×5 n.m.	73.9%
	Annual traffic 3×3 n.m.	75.1%
	Summer traffic 5×5 n.m.	75.0%
	Winter traffic 5×5 n.m.	80.7%
Distribution (std)	Annual traffic 5×5 n.m.	60.2%
	Annual traffic 3×3 n.m.	51.2%
	Summer traffic 5×5 n.m.	66.6%
	Winter traffic 5×5 n.m.	52.9%

Table 7. Grounding frequency of the studied waterways– WCI scaled from 0 (the most difficult) to 1 (the easiest).

Fairway	WCI	No. of groundings (1989–2010)	Ship Transits in 2010	f_{gr}
Kotka	0.75	4	6226	2.92 E-5
Sköldvik	0.76	5	3632	6.26 E-5
Helsinki-West	0.74	1	4887	0.93 E-5
Helsinki-South	0.53	5	5119	4.44 E-5
Hamina	0.64	4	3473	5.24 E-5

a shoal. Table 6 also shows that U is increased for the second moment of the summer traffic while it is decreased for winter traffic. Knowing that the second moment of traffic distribution gives information about the traffic scattering, this may be explained as the ships in winter mostly choose not to sail where the ice situation is harsh, which is not constant and may change during a day. Thus, winter traffic is more scattered compared with summer traffic, which makes the data of winter traffic less informative about the scattering of the traffic.

4.2. *Grounding vs. WCI.* Using WCIs defined in Table 2 and the number of GAs occurring from 1989 till 2010 in the studied waterways (see Table 7), the correlation and rank coefficients show small positive rank and negative linear dependency between WCI and grounding frequency (ρ and $\tau=0.2$ and $r=-0.24$; see Figure 8). Grounding frequency is defined in Equation (1) as the number of GAs over the ship transits for the studied waterways. Ship transits are estimated based on the AIS data of 2010 for each waterway. Thus, grounding frequency is estimated as:

$$f_{gr} = N_{gr} / (22 \times N_{Transit}) \quad (1)$$

where f_{gr} is grounding frequency, N_{gr} is the number of GAs in the waterway over 22 years (1989 till 2010), and $N_{Transit}$ is the ship transit of the studied ports based on the AIS data of 2010.

Looking at Figure 8, one notices that data point of Sköldvik can be seen as an outlier since the waterway to Sköldvik is ranked as the easiest waterway to navigate among the studied waterways, while it has the highest frequency of GA (see Table 7). By omitting the data-point of Sköldvik from the dataset, the linear dependency

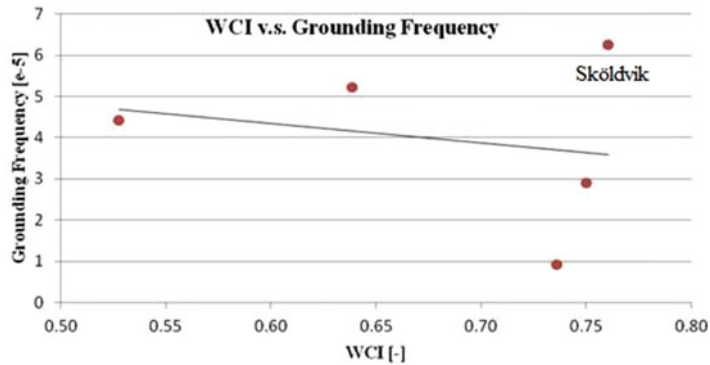


Figure 8. Weak linear dependency of -0.24 is detected between WCI and Grounding Frequency; WCI scaled from 0 (the most difficult) to 1 (the easiest).

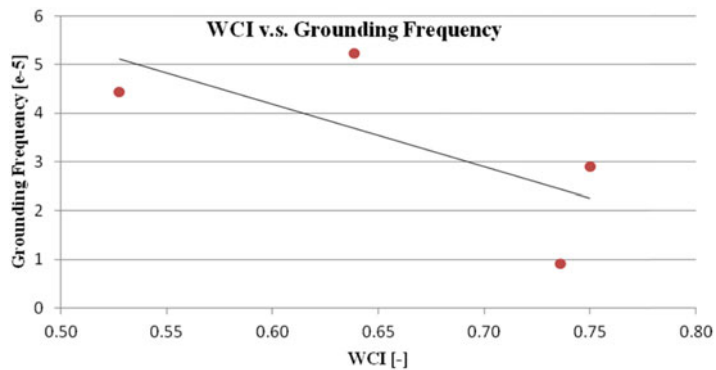


Figure 9. Strong linear dependency of -0.71 is detected between WCI and Grounding Frequency when data point related to Sköldvik is removed from the dataset; WCI scaled from 0 (the most difficult) to 1 (the easiest).

between WCI and the grounding frequency is significantly increased to $r = -0.71$ (Figure 9), and rank coefficients are also now showing the similar dependency (negative) between the grounding frequency and WCI ($\rho = -0.6$ and $\tau = -0.34$). Considering Sköldvik as an outlier can be justified by the fact that tankers, which are quite difficult to manoeuvre, especially when they are loaded, use this specific waterway. Separating the tanker traffic from the traffic of other ship types is often practised, see for example Montewka et al. (2010).

Separating tanker traffic from the analysis, the obtained results reveal that the more complex a waterway is, the higher frequency of GA in that waterway. Nevertheless, due to the limited number of the studied waterways, the χ^2 – and MI tests cannot be performed meaningfully to further confirm the obtained results.

5. DISCUSSION. The results of this study show that although there is no significant statistical correlation between the number of GAs and traffic density and distribution (Table 3), it cannot be said that ship traffic and GA are statistically

independent. The MI test shows that although knowing about the traffic density may not significantly increase our knowledge about the expected frequency of GAs in an area, the information about the traffic distribution will do (Table 6). Thus, based on the results of this study and within its boundaries, it can be concluded that only spatial distribution of traffic in an area (not traffic volume) is a factor related to the frequency of GA. This finding can be supported by the recent accident report of HELCOM on the Baltic Sea area (HELCOM, 2011), showing that despite an increase in the traffic density of the area during recent years, the number of GAs has decreased.

The results also show statistical dependency between WCI and the frequency of GA, which means that the more complex a waterway is to navigate, the higher frequency of GA is expected to be seen in the area.

The implemented data, methods, and algorithms add uncertainty into the final results; therefore the limitations of the study are discussed below, along with our recommendations for further studies.

The data utilised for this study comes from statistics of the GoF and the results can only be considered valid for the studied area and for the utilised datasets. Therefore, the derived conclusions are only consistent with the resultant data and they should be later tested and verified using statistical data of different areas.

The AIS data utilised for this research represents maritime traffic of year 2010 in the area, while the historical accident data were from years 1989 till 2010. The recent traffic pattern does not necessarily represent the traffic pattern in the past. In fact, the ship traffic in the area has changed during the past decades, due to the opening of new ports in the area and the economic growth of the neighbouring nations (Kujala et al., 2009). Besides, traffic separation schemes in GoF (see Figure 2) have applied since 2001 as part of the joint Vessel Traffic Management and Information System (VTMIS) in the GoF, which has affected the ship traffic in the area (Sonninen et al., 2008). The same effect may have been caused by the implementation of Vessel Traffic Services (VTS) (MSC, 2003) and thereafter Gulf of Finland Reporting (GOFREP) (MSC, 2006) system in the area in 2003 and 2006 respectively. Therefore, the results of the study should be confirmed by performing similar studies and using traffic schemes of different years.

The frequency of GA was estimated assuming uniformity during 22 years and using the AIS data of the year 2010. This assumption brings uncertainty to the result. However, this uncertainty can be reduced if the AIS data of more years are used in order to estimate a better representative distribution of ship transits than uniform distribution.

The size of the data sample that is used for WCI study was small, which was due to the limited available resources. Since WCI can only be estimated via qualitative methods at this stage, the number of studied waterways was limited, which may have increased the uncertainty of the results. Therefore, defining quantitative methods for estimating WCI can help the researchers to increase the size of the studied waterways, which then may decrease the involved uncertainties. In this regard, the next step with WCI is to split up a fairway into pieces regarding its characteristics, such as number and quality of turns, length of the legs, and width of the fairway and then grading these fragments. This can lead us to a quantitative method to easily estimate WCI for every waterway only based on its characteristics.

The results show that although the size of the grids for extracting the traffic properties may not have significant effect on the correlation tests: it certainly has

effects on the results of MI tests. This may indicate that the sizes and shape used for gridding are not optimum for such studies. Hence, some other ways of dividing the study area such as hybrid grid cell size, using smaller grid size in more dense areas and bigger one for the less dense areas, may perform better. Additionally, since the results of this study also confirm the statistical dependency between GA and WCI, dividing the study area based on the waterways may be more realistic for such studies. In this way, the waterway legs can serve as the “grids”, from where the traffic properties should be extracted. These are the methods that could be taken into account for future similar studies.

The results show that separating the traffic into winter and summer traffic affects the results of the MI tests on the first and second moments of traffic distribution. Given the result of the dependency between the traffic distribution and GA, it is concluded that the season is an affecting factor on GAs through the traffic distribution. Figure 3 also shows different trends for the number of GAs for the GoF between winter and summer seasons, which can be seen as a support for the above argument. Thus, we recommend future related studies separate the winter and summer traffic in the areas where the sea is covered with ice in winter.

6. CONCLUSION. The objective of this paper was to study the commonly accepted premise about the relationship between the characteristics of maritime traffic (density and distribution) and the frequency of Grounding Accidents (GA). This has been achieved by studying the statistical dependency between: 1. ship traffic and the frequency of GA; and 2. waterway complexity and the frequency of GA.

The study utilised AIS data to obtain traffic characteristics, accident statistics with respect to the grounding accidents in the Gulf of Finland, and experts’ knowledge elicitation techniques to deliver an index describing the complexity of waterways. Applying correlation tests and Mutual Information (MI) analysis we carried out statistical analysis on the data.

The obtained results reveal that statistical dependency exists between the frequency of grounding accident and traffic distribution and between the former and the complexity of the waterway. Using these results we can conclude that traffic distribution and waterway complexity are affecting factors on the expected frequency of grounding accident, and are valid variables to be used when modelling the risk of grounding accidents.

It is also shown that there is no statistically significant dependency between the traffic density and grounding accidents. Therefore, the results of this study recommend not using traffic density as an input variable to models assessing the frequency of grounding accidents, especially for the studied area.

Since this outcome is a counter to the currently accepted beliefs, it highlights that it is important to conduct more statistical research on recognising the most significant affecting factors on the expected frequency of grounding accident when this type of accident is being modelled.

ACKNOWLEDGEMENTS

HELCOM is acknowledged for providing us with the AIS and accident data related to the study area. The authors are also grateful to the sea pilots, who willingly participated in the interviews, for their valuable inputs to this research.

FINANCIAL SUPPORT

This study was conducted as part of “Minimizing risks of maritime oil transport by holistic safety strategies” (MIMIC) project. The project was partly funded by the European Union and the financing comes from the European Regional Development Fund, The Central Baltic INTERREG IV A Programme 2007–2013; the City of Kotka; Kotka-Hamina Regional Development Company (Cursor Oy); Centre for Economic Development, and Transport and the Environment of Southwest Finland (VARELY).

REFERENCES

- Akhtar, M.J. and Utne, I.B. (2014). Human Fatigue’s Effect on The Risk of Maritime Groundings – A Bayesian Network Modeling Approach. *Safety Science*, **62**, 427–440.
- Amrozowicz, M.D., Brown, A. and Golay, M. (1997). A Probabilistic Analysis of Tanker Groundings. *7th International Offshore and Polar Engineering Conference*. Honolulu, Hawaii.
- Briggs, M.J., Borgman, L.E. and Bratteland, E. (2003). Probability Assessment for Deep-Draft Navigation Channel Design. *Coastal Engineering*, **48**, 29–50.
- Brown, A. and Haugene, B. (1998). Assessing the Impact of Management and Organizational Factors on the Risk of Tanker Grounding. *8th International Offshore and Polar Engineering Conference*.
- Fowler, T.G. and Sørgård, E. (2000). Modelling Ship Transportation Risk. *Risk Analysis*, **20**, 225–244.
- Fujii, Y., Oshima, R., Yamanouchi, H. and Mizuki, N. (1974). Some Factors Affecting The Frequency Of Accidents In Marine Traffic: I- The Diameter Of Evasion For Crossing Encounters, II- The Probability Of Stranding, III- The Effect Of Darkness On The Probability Of Collision And Stranding. *The Journal of Navigation*, **27**, 239–247.
- Goerlandt, F. and Kujala, P. (2011). Traffic Simulation Based Ship Collision Probability Modeling. *Reliability Engineering and System Safety*, **96**, 91–107.
- Goerlandt, F. and Kujala, P. (2014). On the reliability and validity of ship–ship collision risk analysis in light of different perspectives on risk. *Safety Science*, **62**, 348–365.
- Goerlandt, F., Montewka, J., Kujala, P. (2014). Tools for an Extended Risk Assessment for Ropax Ship-Ship Collision. *Vulnerability, Uncertainty, and Risk*, 2292–2302, doi: 10.1061/9780784413609.230.
- Hänninen, M., Mazaheri, A., Kujala, P., Montewka, J., Laaksonen, P., Salmiovirta, M., Klang, M. (2014). Expert elicitation of a navigation service implementation effects on ship groundings and collision in the Gulf of Finland. *Proceedings of the Institution of Mechanical Engineers, Part O, Journal of Risk and Reliability*, **228**, 19–28.
- Hänninen, M., Valdez Banda, O., Kujala, P. (2014a). Bayesian network model of maritime safety management. *Expert Systems with Applications*, DOI: 10.1016/j.eswa.2014.06.029.
- HELCOM. (2011). Report On Shipping Accidents In The Baltic Sea Area During 2011. Helsinki Commission, Baltic Marine Environment, Protection Commission.
- Jebsen, J.J. and Papakonstantinou, V.C. (1997). Evaluation of The Physical Risk Of Ship Grounding; *Department of Ocean Engineering*. Massachusetts Institute of Technology.
- Kite-Powell, H.L., Jin, D., Jebsen, J., Papakonstantinou, V. and Patrikalakis, N. (1999). Investigation of Potential Risk Factors for Groundings of Commercial Vessels in U.S. Ports. *International Journal of Offshore and Polar Engineering*, **9**, 16–21.
- Kujala, P., Hänninen, M., Arola, T. and Ylitalo, J. (2009). Analysis of the Marine Traffic Safety in The Gulf of Finland. *Reliability Engineering and System Safety*, **94**, 1349–1357.
- Kuronen, J., Helminen, R., Lehto, A. and Tapaninen, U. (2008). Maritime Transportation in the Gulf of Finland in 2007 and in 2015. University of Turku.
- Lehman, A. (2005). *JMP for Basic Univariate and Multivariate Statistics : A Step-by-Step Guide*, Cary, NC, USA, SAS Press.
- Lin, S.-C. (1999). Physical Risk Analysis of Ship Grounding. Massachusetts Institute of Technology.
- Martins, M.R. and Maturana, M.C. (2010). Human Error Contribution in Collision and Grounding of Oil Tankers. *Risk Analysis*, **30**, 674–698.
- Mazaheri, A., Montewka, J. and Kujala, P. (2013). Modelling the Risk of Ship Grounding – A Literature Review from a Risk Management Perspective. Published Online In *Wmu Journal Of Maritime Affairs*; doi: 10.1007/S13437-013-0056-3.

- Mazaheri, A. and Ylitalo, J. (2010). Comments on Geometrical Modeling of Ship Grounding. *5th International Conference on Collision and Grounding of Ships (ICCGS)*. Espoo, Finland.
- Montewka, J., Hinz, T., Kujala, P. and Matusiak, J. (2010). Probability Modelling of Vessel Collisions. *Reliability Engineering and System Safety*, **95**, 573–589.
- Montewka, J., Krata, P., Goerlandt, F., Mazaheri, A. and Kujala, P. (2011). Marine Traffic Risk Modelling – An Innovative Approach and a Case Study. *Proceedings of The Institution of Mechanical Engineers, Part O, Journal of Risk and Reliability*, **225**, 307–322.
- Montewka, J., Goerlandt, F., Kujala, P. (2014). On a systematic perspective on risk for formal safety assessment (FSA). *Reliability Engineering & System Safety*, **127**, 77–85.
- Montewka, J., Ehlers, S., Goerlandt, F., Hinz, T., Tabri, K., Kujala, P. (2014a). A framework for risk assessment for maritime transportation systems—A case study for open sea collisions involving RoPax vessels. *Reliability Engineering & System Safety*, **124**, 142–157.
- MSC. (2003). Mandatory Ship Reporting Systems, SN/Circ.225. In IMO.
- MSC. (2006). Mandatory Ship Reporting Systems, SN.1/Circ.258. In IMO.
- O’hagan, A., Buck, C.E., Daneshkhan, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E. and Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts’ Probabilities*. New York, John Wiley & Sons.
- Otto, S., Pedersen, P.T., Samuelides, M. and Sames, P.C. (2002). Elements Of Risk Analysis For Collision And Grounding of a Ro-Ro Passenger Ferry. *Marine Structures*, **15**, 461–474.
- Pedersen, P.T. (1995). Collision and Grounding Mechanics. *Proceedings of WEMT ‘95*. Copenhagen, Denmark, The Danish Society of Naval Architecture and Marine Engineering.
- Peng, H., Long, F. and Ding, C. (2005). Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**, 1226–1238.
- Praetorius, G. (2012). Safety within the Vessel Traffic Service (VTS) Domain – Understanding the Role of The VTS for Safety within Maritime Traffic Management, Licentiate Thesis, *Department of Shipping and Marine Technology*. Gothenburg, Chalmers University of Technology.
- Quy, N.M., Vrijling, J.K., Gelder, P.H.A.J.M.V. and Groenveld, R. (2006). On the Assessment of Ship Grounding Risk in Restricted Channels. *The 8th International Conference on Marine Sciences and Technologies – Black Sea Conference*. Varna, Bulgaria.
- Robson, C. (2008). *How to do a Research Project*, Singapore, Blackwell Publishing.
- Samuelides, M.S., Ventikos, N.P. and Gemelos, I.C. (2009). Survey on Grounding Incidents: Statistical Analysis and Risk Assessment. *Ships and Offshore Structures*, **4**, 55–68.
- Sonninen, S., Nuutinen, M. and Rosqvist, T. (2008). Development Process of The Gulf of Finland Mandatory Ship Reporting System, 614, VTT.
- Sormunen, O.-V.E., Goerlandt, F., Häkkinen, J., Posti, A., Hänninen, M., Montewka, J., Stahlberg, K., Kujala, P. (2014). Uncertainty in maritime risk analysis: Extended case study on chemical tanker collisions. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment* doi: 10.1177/1475090213515640.
- Steuer, R., Kurths, J., Daub, C.O., Weise, J. and Selbig, J. (2002). The Mutual Information: Detecting and Evaluating Dependencies Between Variables. *Bioinformatics*, **18**, 231–240.
- Theil, H. (1970). On the estimation of relationships involving qualitative variables. *American Journal of Sociology*, **76**, 103–154.
- van Dorp, J.R. and Merrick, J.R.W. (2009). On a Risk Management Analysis of Oil Spill Risk Using Maritime Transportation System Simulation. *Annals of Operations Research*, 1–29.