A framework for evidence-based risk modeling of ship grounding

Arsham Mazaheri
A framework for evidence-based risk modeling of ship grounding

Arsham Mazaheri

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Engineering, at a public examination held at the lecture hall 216 at Otakaari 4 (K1 Building) on 11th August 2017 at 12:00.

Aalto University
School of Engineering
Mechanical Engineering
Research Group on Maritime Risk and Safety
Supervising professor
Professor Pentti Kujala, Aalto University, Finland

Thesis advisor
Associate Professor Jakub Montewka, Aalto University, Finland

Preliminary examiners
Professor Margareta Lützhöft, Australian Maritime College (UTAS), Australia

Professor Ingrid Bouwer Utne, Norwegian University of Science and Technology (NTNU), Norway

Opponent
Associate Professor Marcelo Ramos Martins, University of São Paulo (USP), Brazil
Abstract

Most of the risk models for ship-grounding accidents do not fully utilize available evidence, since in general they are merely based on accident statistics and expert opinions. The major issue with models of such kind is their limitation in supporting the process of risk management with respect to grounding accidents; because they do not necessarily reflect the reality to the extent required.

This thesis proposes an evidence-based framework for building evidence-based risk models for probabilistic assessment of ship grounding. In order to build the evidence required for creating the evidence-based risk model, traffic characteristics as primary source of data are extracted from AIS data and accident statistics of the Gulf of Finland. Additionally, using expert knowledge of the local pilots in the Gulf of Finland, a location dependent and semi-quantitative index as Waterway Complexity Index is defined to assess the dependency of ship grounding and navigational difficulty of a waterway to handle a ship. Moreover, ship grounding incident and accident reports from Finnish, Swedish, and British maritime authorities are utilized as primary and secondary sources of data respectively to build the required evidence for constructing the risk model. In this regard, two frameworks are introduced in this thesis to review and extract the embedded information from the reports. A new version of Human Factors Analysis and Classification System (HFACS) is introduced as a framework to review the grounding accident reports; and a new positive taxonomy as Safety Factors, which are based on high level positive functions that are prerequisite for safe transport operations, is introduced to review the grounding incident reports.

Utilizing the proposed framework for evidence-based risk modeling as well as the built evidence from primary and secondary sources of data and expert knowledge, a Bayesian Network risk model is developed in this thesis for assessing the probability of ship grounding accidents. The uncertainties associated with the elements of the model are clearly communicated to the end user adopting a concept of strength of knowledge, and by introducing knowledge strength map for the built model. Therefore, it is argued in this thesis that the developed model is more suitable for risk management purposes, and the model can be used to suggest proper risk-control-measures to mitigate the risk of ship grounding accident. The developed model in this thesis suggests the high-level critical parameters that need proper control measures are complexity of waterways, traffic situations encountered, and off-coursed ships. The critical area that calls for more investigation is the onboard presence of a sea-pilot.

Keywords Evidence-Based Risk Modeling, Ship Grounding, Knowledge Strength, Background Knowledge, Bayesian Network

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ISSN-L</td>
<td>1799-4934</td>
<td>ISSN (printed)</td>
<td>1799-4934</td>
</tr>
<tr>
<td>Location of publisher</td>
<td>Helsinki</td>
<td>Location of printing</td>
<td>Helsinki</td>
</tr>
</tbody>
</table>
This dissertation is dedicated to my beloved parents, Katayoun and Kambiz, whom I am indebted my life and existence to the sacrifices they have made for me and to the endless love they have given to me. The thesis is also dedicated to my precious wife, Paria, who gave me the second life; to my lovely sister, Azadeh, for being my emotional support whenever I needed the most; and to my in-laws, Mohammad, Shahla, Shahab, Pouya, Leila, and Elina whom my life has become much more beautiful and meaningful by having them in my heart and by my side.
Preface

This dissertation is based on the work performed at Helsinki University of Technology in 2009 and afterwards at Aalto University - School of Engineering and at Kotka Maritime Research Center (MERIKOTKA) during 2010-2015. The financial support for this work is mainly provided by the European Union, European Regional Development Fund, the Central Baltic INTERREG IV A Programme 2007-2013, the City of Kotka, and Kotka-Hamina Regional Development Company (Cursor Oy) within the projects SAFGOF and MIMIC; and by Finnish Transport Safety Agency (TraFi) within smaller periodic research projects. The financial support is acknowledged and greatly appreciated. I would also like to thank Merenkulun Säätiö (Maritime Foundation Scholarship System) for their financial support for me to attend scientific conferences to present my works during these years as well as to finalize this compendium. Baltic Marine Environment Protection Commission (HELCOM), Finnpilot Pilotage Oy., and ForeSea are also thanked for providing the required traffic and accident/incident data for conducting the research.

I wish to thank my supervisor, Professor Pentti Kujala, for giving me the opportunity to work in his group. My ultimate gratitude towards my instructor, Associate Professor Jakub Montewka, who was more like a mentor to me during this path with all the encouragement and valuable discussions during the process. Special thanks to all the secretaries during all these years, Leila Silonsaari, Seija Latvala, Pirkko Suominen, and Marisa Lundström as well as Saeed Elhazaz, the janitor of the ship laboratory at Aalto, whose support and help with daily problems are recognized. I would also like to thank my colleagues and friends in Espoo, Helsinki, Kotka, Turku, as well as in Estonia and Sweden on all these years for the valuable discussions and occasional fun that we had. I am also grateful to my co-authors for their collaborations.

I should not, and for sure will not, forget my close friends in Finland: Mohammad A., Ali N., Javad H., Ali A., Nasim D., Pegah K., Nashmin E., Nima R., Ghazaleh A., Abolfazl K., Maral J., and Marizieh J., whose companionship was preciously calming when I was frustrated with the daily work. Let say if I have any remaining “mental health!” I owe that to you pals.

Finally, I would like to thank my family for always being there for me, no matter what; and for giving me the courage to take the final steps in this journey and finishing this compendium.

Stockholm

September, 2016

Arsham Mazaheri
## Contents

Dedication ............................................................................................................................................... 1  
Preface ..................................................................................................................................................... 2  
List of Publications ................................................................................................................................. 5  
Author’s Contribution ................................................................................................................................. 6  
Original Features ..................................................................................................................................... 7  
Special Terms .......................................................................................................................................... 9  
Acronyms .............................................................................................................................................. 10  

1. Introduction ................................................................................................................................... 11  
   1.1 Motivation and background: evidence-based risk modeling for ship grounding accidents . 11  
   1.2 Objectives and scope of the work .......................................................................................... 13  
   1.3 Delimitations ......................................................................................................................... 14  

2. A framework to model the risk of ship grounding from a risk management perspective .......... 16  
   2.1 Background knowledge ........................................................................................................... 18  
   2.2 An evidence-based risk modeling framework ....................................................................... 18  

3. Evidence building for ship grounding risk modeling ................................................................. 20  
   3.1 Accident and incident reports ................................................................................................ 20  
      3.1.1 Framework to review accident reports .......................................................................... 21  
      3.1.2 Framework to review incident reports .......................................................................... 23  
   3.2 Traffic and accident data ....................................................................................................... 24  
   3.3 Expert opinion ....................................................................................................................... 29  

4. Evidence-based risk model for ship grounding ........................................................................ 31  
   4.1 Model construction ................................................................................................................ 31  
   4.2 Model parametrization .......................................................................................................... 32  
   4.3 Model validation ................................................................................................................... 33  
      4.3.1 Nomological Validity .................................................................................................... 34  
      4.3.2 Construct Validity ......................................................................................................... 36  
      4.3.3 Criterion Validity .......................................................................................................... 37
List of Publications

This doctoral dissertation consists of an overview and of the following publications, which are referred to in the text by their Roman numerals.


Author’s Contribution

Publication I: “Modeling the risk of ship grounding - A literature review from a risk management perspective”

The author prepared the framework and performed the literature review, and was the main contributor to the manuscript. Montewka contributed to the manuscript and gave valuable comments and suggestions. Kujala provided valuable comments and suggestions.

Publication II: “Assessing grounding frequency using ship traffic and waterway complexity”

The author presented the idea, designed the algorithms, carried out the analyses and evaluations for the case studies, and was the main contributor to the manuscript. Montewka contributed to the manuscript and gave valuable comments and suggestions. Kotilainen performed the interviews and analyzed them. Sormunen and Kujala provided valuable comments and suggestions.

Publication III: “Usability of Accident and Incident Reports for Evidence-Based Risk Modeling - A case study on ship grounding reports”

The author presented the idea and developed the methodology for reviewing accident and incident reports, reviewed the accident and incident reports, carried out the analyses and evaluations, and was the main contributor to the manuscript. Montewka contributed to the manuscript and gave valuable comments and suggestions. Nisula contributed to the methodology development for reviewing incident reports, and contributed to the manuscript and incident reports reviewing. Kujala provided valuable comments and suggestions.

Publication IV: “Towards an evidence-based probabilistic risk model for ship-grounding accidents”

The author presented the idea and developed the framework, constructed the model, carried out the analyses and evaluations, and was the main contributor to the manuscript. Montewka contributed to the manuscript and gave valuable comments and suggestions. Kujala provided valuable comments and suggestions.
Original Features

This thesis aims to contribute to the evidence-based risk modeling of ship grounding accident for the purpose of managing the associated risk. The contribution is through the development of a Bayesian Network risk model and through communicating the knowledge strength of the model’s elements. The following features of this thesis are believed to be original:

1. Developing a methodological framework suitable for evidence-based risk modeling based on the recommendations given by the Formal Safety Assessment issued by the International Maritime Organization (PI, PIII, PIV).

2. Conducting a critical analysis, from a risk management perspective and by using the developed framework, on the risk models available in the literature for ship grounding accident risk analysis; and finding evidence-based risk modeling approach is extremely rare used practice in ship grounding accident risk analysis (PI).

3. Developing a framework for reviewing structured ship grounding accident reports based on the complex-linear method of Human Factors Analysis and Classification System; developing a framework based on high level positive functions and for reviewing semi- and unstructured ship grounding incident reports to extract the contributing factors and the existing safety barriers (PIII).

4. Performing a systematic approach to study accident and incident reports, with the aim of extracting the imbedded knowledge in the reports useful for evidence-based risk modeling (PIII).

5. Assessing the usefulness of ship grounding accident and incident reports on evidence-based risk modeling, and discovering that the voluntary incident reports in their current format cannot be used in evidence-based risk modeling (PIII).

6. Developing algorithms to extract traffic density and traffic distribution from AIS data; conducting a detailed analysis on the effects of traffic density and traffic distribution on ship grounding accident frequency, and discovering that evidence suggests traffic density and frequency of ship grounding accident are statistically independent (PII).

7. Developing a location-dependent semi-quantitative index as Waterway Complexity Index to rank fairways with regard to the difficulty of ship handling; and conducting a detail analysis on the effects of waterway complexity on ship grounding accident frequency, and discovering that evidence suggests the waterway complexity and the frequency of ship grounding accident are statistically dependent (PII).
8. Constructing an evidence-based Bayesian Network risk model, suitable for risk management purposes, for assessing the probabilistic risk of ship grounding accidents; and developing a visualization method to communicate the strength of the knowledge of each element of the model to the end users of the model (PIV)
### Special Terms

| **Accuracy** | is the degree to which the result of a specification, measurement, or calculation adequately conforms to the correct value |
| **Evidence** | is the available body of facts or information, mostly in the form of hard data like document, recording, and collected data by various sensors on board the ship indicating whether a belief or proposition is true or valid. Testimonies of multiple witnesses that are based on real-life examples are also considered as evidence |
| **Knowledge** | is considered the same as Evidence. In general, evidence is to justify knowledge. Knowledge of a subject could be based on evidence and thus can be justified by evidence for larger group of subjects, or it can only be justified by the subject to accepting them as its own belief. As the result, one's belief is different than one's knowledge; which means knowledge and only knowledge constitutes evidence. Therefore, one's evidence is the same as one's knowledge (Williamson, 2000) |
| **Maritime** | refers to objects or activities related to the sea |
| **Primary Data** | is considered as raw data and with no (secondary) interpretations involved. Examples of such data are the data that is retrieved from the Voyage Data Recorders (VDR) onboard the ships, or the testimony of the witnesses of an event shortly after the event |
| **Reliability** | is the degree to which a specification, measurement, or calculation adequately produces the same result every time it is repeated |
| **Risk Analysis/Assessment** | is a process that identifies/quantifies potential problems/threats/mishaps that could undermine or jeopardize the safe journey of a ship |
| **Risk Management** | is a process that uses Risk Analysis to foresee and evaluate potential risks and identifies procedures to avoid or minimize their impact on maritime transportation |
| **Secondary Data** | is considered as processed and collected data based on primary data, and normally some level of (secondary) judgmental interpretation and analysis are involved in this source of information |
| **Validity** | is the degree to which a specification, measurement, or calculation adequately describes the specific concept one attempts to describe |
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIS</td>
<td>Automatic Identification System</td>
</tr>
<tr>
<td>BN</td>
<td>Bayesian Network, Bayes Network, Belief Network, Bayesian Belief Network (BBN)</td>
</tr>
<tr>
<td>BRM</td>
<td>Bridge Resource Management</td>
</tr>
<tr>
<td>BST</td>
<td>Behavior Sensitivity Test</td>
</tr>
<tr>
<td>CPT</td>
<td>Conditional Probability Table</td>
</tr>
<tr>
<td>DAG</td>
<td>Directed Acyclic Graph</td>
</tr>
<tr>
<td>EBT</td>
<td>Evidence Based Training</td>
</tr>
<tr>
<td>FSA</td>
<td>Formal Safety Assessment</td>
</tr>
<tr>
<td>GoF</td>
<td>Gulf of Finland</td>
</tr>
<tr>
<td>HFACS</td>
<td>Human Factors Analysis and Classification System</td>
</tr>
<tr>
<td>IMO</td>
<td>International Maritime Organization</td>
</tr>
<tr>
<td>PSSA</td>
<td>Particularly Sensitive Sea Area</td>
</tr>
<tr>
<td>QFT</td>
<td>Qualitative Features Test</td>
</tr>
<tr>
<td>SF</td>
<td>Safety Factor</td>
</tr>
<tr>
<td>TraFi</td>
<td>Finnish Transport Safety Agency</td>
</tr>
<tr>
<td>VDR</td>
<td>Vessel Data Recorder</td>
</tr>
<tr>
<td>VTS</td>
<td>Vessel Traffic Service</td>
</tr>
<tr>
<td>WCI</td>
<td>Waterway Complexity Index</td>
</tr>
</tbody>
</table>
1. Introduction

1.1 Motivation and background: evidence-based risk modeling for ship grounding accidents

Maritime transportation is the main mode of transportation for the global trade that accounts for about 80% of the world trade. This translates to 9.84 billion tons of trading cargos in 2015 globally (UNCTAD, 2015). The importance of shipping increases for the countries like Finland, who their import and export’s trades are greatly dependent on sea transportation (FinnishCustoms, 2014). Therefore, it is fair to assume major maritime accidents are threats not only to the safety of maritime transportation, but also to the global trade; because the consequences of such accidents could be severe and expensive (Helle et al., 2011, Montewka et al., 2014a, Montewka et al., 2013).

Ship grounding accident is one of the main types of maritime accident that involves the impact of a ship on the seabed or waterway side. It results the damage of the submerged part of the ship’s hull and in particular the bottom structure; potentially leading to water ingress and compromise of the ship's structural integrity and stability. Grounding still accounts for about one-third of commercial ship accidents all over the world and has the second rank in frequency in global perspective, after ship-ship collision (Samuelides et al., 2009). However, grounding is the major maritime accident in the Gulf of Finland (Kujala et al., 2009), which is part of the Baltic Sea Area, where is designated as Particularly Sensitive Sea Area (PSSA) by IMO (2005). A PSSA is an area where it needs special protection through action because of its vulnerable ecological, socio-economic, or scientific attributes where international shipping activities could damage such attributes (IMO, 2005).

Risk analysis and risk management in maritime domain aim at assessing the risk of international shipping activities and then implement controlling measures to decrease that risk. Depends on which definition of risk is taken (see Goerlandt and Montewka, 2015b) different elements may come into consideration for the risk assessment and thereafter defining the controlling measures. However, for the assessed risk to be beneficial for the decision makers and for risk management purposes, one element that must be taken into consideration from the first step is the strength of the knowledge behind each input of the model as the involved uncertainty; (see Rosqvist, 2010, Rosqvist and Tuominen, 2004, Aven and Heide, 2009). Without that element, the decision makers have almost no mean to assess the reliability of the model at hand. Therefore any decision that is made based on the results of the model may have unknown assumptions due to lack of available information, which increases the level of uncertainty in the made decision. Thus, taking into account evidence and the attached uncertainty is considered a fundamental characteristic of risk analysis application (Goerlandt and Montewka, 2015b).

Currently very few available approaches in ship grounding risk modeling take into account the uncertainty and strength of knowledge (Mazaheri et al. 2014); especially regarding the evidence-based modeling (Goerlandt and Montewka, 2015b), which is almost nonexistence (see Table 1). This is despite...
the fact that the knowledge- and evidence-based modeling\textsuperscript{1} is greatly encouraged in the realm of maritime risk assessment and risk modeling; see for example IMO (2002), Kristiansen (2010), IMO (2012), Aven (2013a), Mazaheri et al. (2014), Montewka et al. (2014b), and Goerlandt and Montewka (2015b). Evidence-based risk modeling of ship grounding in compare with the common way of modeling of ship grounding which is based on extensive use of expert knowledge as soft data (see Table 1), is referring to the use of hard data in hand to the extent of availability. Evidence-based modeling is not against of using expert knowledge, rather encourages to use expert knowledge that is based on hard data and scenarios. For instance, accident and incident reports of ship groundings as two sources of evidence (see Section 3) are more or less structured based on the knowledge of experts, while the experts in that context are exposed to hard data and real scenarios beforehand. However, in common way of risk modeling of ship grounding, experts mostly hypothesize scenarios, which are sometime purely based on their own intuition. This, in compare with evidence-based scenarios, may end up to have scenarios that have no or limited number of examples in reality. This can be interpreted that the analyzed scenarios in that situation are very unlikely, which in risk management perspective is not valid to be considered as feasible scenarios.

An extensive critical review on the existing literature on ship grounding risk modeling (Mazaheri et al., 2014) shows that from among all the available risk models for ship grounding there is only one model that considers actual evidence in the modeling process (i.e. Kristiansen, 2010); the rest of the available models are mostly utilizing expert opinions for that purpose. Moreover, the uncertainty discussion (Merrick and van Dorp 2006; Merrick et al. 2005; Chen and Zhang 2002) has brought into the attention by only four of the reviewed models, in which Kristiansen (2010) is not one of them (see Table 1). Although involved uncertainty in the model construction may be eased using the method adopted by Kristiansen (2010) (i.e., evidence-based model construction), from a risk management point of view, the remaining uncertainty should be exposed to the decision makers for further precautions.

Typically, the largest source of uncertainty is human elements in the forms of expert judgment and human failures quantification (Amrozowicz, 1996). Although there are some methods that can somehow decrease the effects of human elements uncertainty (Pyy 2000; Rosqvist 2003), they are very often not utilized and the uncertainty of this type is not addressed nor discussed in the reviewed literature. Historically, the possible ways to visualize the involved uncertainties were the probability of frequencies, risk curves, the imprecise probability, fuzzy logic, and the theories of possibility and evidence (see Kaplan 1997; Aven and Zio 2011; Aven 2011; Goerlandt and Reniers 2016). Recently, using color codes in combination with other methods is introduced and discussed by Mazaheri et al. (2016) and Goerlandt and Reniers (2016). The new proposals are practically more promising when it comes to the communication between model makers and users of the model (i.e. decision makers) regarding the strength of the knowledge of different elements of the model.

\textsuperscript{1} In general, evidence is to justify knowledge. Knowledge of a subject could be based on evidence and thus can be justified by evidence for larger group of subjects, or it can only be justified by the subject to accepting them as its own belief. As the result, one's belief is different than one's knowledge; which means knowledge and only knowledge constitutes evidence. Therefore, one's evidence is the same as one's knowledge (Williamson, 2000).
Table 1. Comparison of the existing risk models with noticeable contribution to the risk modeling of ship grounding accident, replicated from Publication-I (Mazaheri et al., 2014)

<table>
<thead>
<tr>
<th>Source</th>
<th>Year</th>
<th>Implemented method</th>
<th>Source of the used data</th>
<th>Risk Control Options</th>
<th>Uncertainty Discussion</th>
<th>Decision Making Potential</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Geometrical scenarios</td>
<td>Fault Tree Analysis</td>
<td>Bayesian Networks</td>
<td>Simulation</td>
<td>Historical data</td>
</tr>
<tr>
<td>Fujii et al.</td>
<td>1974</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Macduff</td>
<td>1974</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pedersen</td>
<td>1995</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Amrozowicz</td>
<td>1996</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fowler et al.</td>
<td>2000</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>DNV</td>
<td>2003</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>RAMBØLL</td>
<td>2006</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Eide et al.</td>
<td>2007</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>COWI</td>
<td>2008</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Uluscu et al.</td>
<td>2009</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>van Dorp and Merrick</td>
<td>2009</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Kristiansen</td>
<td>2010</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Montewka et al.</td>
<td>2011</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

* Based on real accident and incident cases
† H = High
M = Medium
L = Low

1.2 Objectives and scope of the work

The main objective of this thesis is to propose a methodological framework for evidence-based risk modeling of ship grounding by utilizing Bayesian Networks (BN). The framework suggests using evidential data in a hierarchical manner (i.e. bottom-up) to build up an evidence-based risk model. Implementing a hierarchical procedure helps the analysts and model developers to use the primary data as the initial building blocks for constructing a risk model, and then use the secondary data and expert judgments as the finishing blocks of the model. BN, as a modeling tool, is utilized in the proposed framework because it can overcome many inherent challenges that evidence-based risk modeling of ship grounding faces. BN can, for example, perform a two-way reasoning that is specifically useful when the
model is extended to a decision problem model, handle uncertainty and lack of evidential data, and model complex and dynamic systems.

The research that is conducted to achieve the main goal of the thesis can be divided into two steps, each with its own objectives which are published in four journal papers (see Figure 1). The objectives of the first step are: - to understand the state of the art on ship grounding risk modeling, and assess the usability of the current existing models for risk management and decision making purposes (Objective 1); and - to create a framework for evidence-based risk modeling, where the model can support decision making purposes (Objective 2). The objectives of the second step are: - to build required evidence for constructing an evidence-based ship grounding risk model based on the proposed framework (Objective 3); - to create and validate an evidence-based risk model for ship grounding based on the proposed framework and the built evidence (Objective 4); and - to develop an effective way for communicating the strength of the knowledge of the used evidence in the developed model (Objective 5).

The links between the above steps, objectives of the thesis, the journal articles (publications-I to IV), and the sections of this compendium are shown in Figure 1.

Figure 1. The links between the objective of the thesis, publications, and the structure of the compendium

Given the objectives of the work, various stakeholders can benefit from the study and its results. The main group that can benefit the most are risk assessors and managers, who need to identify the involved risk of a ship grounding accident in maritime transportation, and then decide on implementing different risk control measures to mitigate the risk. Additionally, other stakeholders involved in maritime transportation such as maritime legal and operational authorities, shipping companies as well as mariners and researchers can benefit from this study and its results.

1.3 Delimitations

The risk analysis and risk management of maritime transportation as a system are very broad topics, and their contents can be varied depends on the taken definition of risk and what actually intends to be addressed (see Goerlandt and Montewka, 2015b). This thesis focuses only on the risk of maritime transportation in the context of ship grounding accident. While the results may be more widely applicable also to other areas and problems such as other types of maritime accidents like ship-ship collision, this is not explicitly claimed or further addressed in this thesis.
From the taken definition of risk in this thesis as $R \sim \{S, L, C\} | BK$ (see Section 2 for a detailed explanation) the focus is only on the scenario part ($S$), i.e. series of mishaps that can cause an accident; and from that, the focus is on the strength of the background knowledge ($BK$) for the foreseen scenarios. Therefore, the likelihood ($L$) and consequence ($C$) part of the taken risk definition is not addressed in this work, nor any other aspects of scenario creation rather than the used knowledge to build up the scenarios. In other words and from the new framework of risk thinking (Aven, 2013b), in where likelihood, knowledge, and surprises are considered, this work only focuses on knowledge part, thus likelihood and surprises are not part of the research nor the discussion.

Moreover, as is discussed in Section 1.2, the main objective of this thesis is to propose a methodological framework for evidence-based risk modeling of ship grounding by utilizing Bayesian Network (BN). In that respect, it is acknowledged that other frameworks for the addressed problem may be feasible. Thus, no claims are made that the proposed framework is the only possible one. Besides, to show the scenarios captured from the data, BN as a modeling methodology is used. While BN has the features that makes it suitable for the purpose listed for this thesis (see Mazaheri et al., 2014), it is acknowledged that BN is only one modeling approach, among many, that is taken to tackle the problem in this work within the proposed framework.

At the end, it also worth to highlight that the data and information which have been used to construct a ship grounding risk model based on the proposed framework are mostly reflecting the situation of the geographical areas of their appearances. More specifically, the extracted knowledge from the reviewed accident and incident reports is reflecting only the maritime traffic situation of where the events have occurred, or characterizes the situation of the maritime authorities that the flags of the involved ships represent. Therefore, the constructed model can certainly be further improved, within the proposed framework, by utilizing the data and information from other geographical areas in order to reflect the situations of other maritime authorities.
2. A framework to model the risk of ship grounding from a risk management perspective

It is shown in Publication-I that many scholars have suggested risk models for ship grounding accident to predict the probability of an accident given certain criteria. However, a question is raised that how many of the proposed models allow risk estimation according to the guideline specified by IMO (2002)? In general, the guideline suggests risk models to be constructed for assessing the associated risk of an unwanted event in a way to provide information required for making decisions to mitigate the risk (IMO 2002). Thus, merely being aware of the risk, as a single number, is not crucial because it has no or limited use for decision makers. Publication-I argues that a suitable model for risk management purposes should reflect the knowledge on the analyzed system with satisfying accuracy (Aven 2013c), and should communicate that knowledge to the users of the model for knowledge-based decision making. Such model can efficiently help the decision makers in choosing the optimal controlling measures, which at the end will hopefully lessen the associated risk. Having mentioned that, the majorities of the available risk models for ship grounding risk analysis, which are presented and reviewed in Publication-I (see Table 1), are mostly focusing on giving risk figures rather than presenting the available background knowledge of the system (Goerlandt and Kujala, 2014, Goerlandt and Montewka, 2015a, Sormunen et al., 2015, Mazaheri et al., 2014). Lack of the background knowledge about the underlying causes of a system, or improper presentation of the available background knowledge leads to uncertainty in the used risk models (Aven and Zio, 2011). Therefore, evidence-based risk modeling that ensures a reasonably good balance between real accident scenarios and imaginary scenarios is encouraged; see for example IMO (2002), Kristiansen (2010), IMO (2012), Mazaheri et al. (2013), and Mazaheri et al. (2014).

In the context of risk analysis, presented in Chapter 6 of Formal Safety Assessment (FSA) guidelines (IMO, 2012), risk is defined as a product of the probability ($P$) and the consequences ($C$) of a given action:

$$ R = P \times C $$

Whereas, in the context of Chapter 7 (ibid), called “Risk control options” - aiming to determine the areas needing control, the risk is decomposed and the uncertainty aspect of two risk components is added as an important element of the decision process. Moreover, in the context of the recommendations, called “Presentation of FSA results”, discussion about the assumptions, limitations and uncertainties of the risk model is recommended (ibid). Therefore Publication-I argues that to make sure all these recommendations can be addressed at the last stage of the analysis, the initial definition of risk must allow for the knowledge-based scenario building, uncertainty analysis, and model validation; (see Rosqvist and Tuominen, 2004, Rosqvist, 2010, Aven and Heide, 2009, Goerlandt et al., 2016). Therefore, the most common definition of risk as the product of the probability of an accident and its consequences (Eq. 2.1) may lead to confusion, especially when comparing the risks associated with the following two situations: $A$) frequent events resulting in low consequences with $B$) rare events of high
consequence. Even though the products of $P$ and $C$ in both cases can be the same, these two situations differ substantially. The background knowledge on $A$ is most probably better than in the case of $B$, as $A$ occurs frequently and $B$ occurs rarely. This knowledge affects the amount of uncertainty associated with the descriptions of $A$ and $B$.

Therefore, describing risk as the combination presented in Eq. 2.1 and expressing it as a single number, leads to the situation where much of the relevant information needed for knowledge-based decision making is not properly reflected. Thereby, the wider concept of risk should be applied, allowing systematic reasoning of the causes that eases decision making on the proper mitigating measures to be in place.

Considering maritime transportation as a system (Figure 2), a well-founded approach to risk can be followed (Haimes, 2009, Aven, 2011), where the existing risk within the system can be defined as a complete set of triplets (Kaplan and Garrick, 1981):

$$R = \{S, L, C\}_C$$

Where, this triplet attempts to answer the following questions: what can go wrong in the system (Scenario - $S$), how likely is it that it goes wrong (Likelihood - $L$), and what are the consequences if the assumed scenario happens (Consequence - $C$)? However, describing the risk as a complete set of triplets is unattainable, simply because our knowledge on the system is never complete, thereby the system cannot be characterized exactly (Aven and Zio, 2011). Therefore, what we actually attempt to describe is an incomplete set of triplets, called “a set of answers” (Kaplan, 1997), which reflects the defined risk for the given system according to our best knowledge and anticipation. Nevertheless, certain triplets, yet existing, remain undiscovered and thus they cannot be captured. Publication-I argues that this incompleteness, which mostly results from lack of background knowledge on the given system, should be recognized and communicated to the decision makers.

![Figure 2. Maritime transportation as a complex socio-technical system presented in Publication-IV (Mazaheri et al., 2016)](image)

To define a set of outcomes, knowledge and proper understanding of the system or phenomena being analyzed is a prerequisite; this is referred to as background knowledge ($BK$). Risk perspective should account for the amount of available background knowledge, and so forth the description of risk perspective for the given system:
R = \{S, L, C|\text{BK}\} \tag{2.3}

In the above definition of risk \( S \) stands for a set of explanatory variables for a given scenario, where the variables and their relations can take different values due to the stochastic nature of the phenomena being analyzed or by applying different assumptions, which depends on the background knowledge (\( \text{BK} \)) of the analyzed process that is available to the risk assessors; \( L \) is a set of likelihoods corresponding to the set of consequences \( C \), for a given scenario and the given combination of assumptions governing the input variables.

This adopted way of risk definition by entering the background knowledge into the consideration enables the reliability check of the model together with the model validation (Aven and Heide, 2009), which is important from risk management perspective.

2.1 Background knowledge

A clear representation of background knowledge that is available about the given system is relevant for any model which is intended for practical use, see for example Aven (2013a). Since the lack of knowledge about the underlying phenomena governing the behavior of the analyzed system leads to uncertainty in the model parameters and on the hypotheses supporting the model structure (Aven and Zio, 2011), it is desirable for a risk framework to communicate the background knowledge level and the involved uncertainties (IMO, 2002, Aven, 2010, Rosqvist and Tuominen, 2004, Mazaheri et al, 2014, Goerlandt and Reniers, 2016). This is especially important in the case of risk modeling in the maritime transportation system, where background knowledge about the system being analyzed is limited and unequally distributed across the system. This, in turn, may introduce varying uncertainties depending on the elements of the system, and may result in a situation in which certain areas of the modeled system lack a sufficient level of background knowledge to satisfy the adopted formal definition of risk. Therefore, we repeatedly argue in Publications-I, III, and IV that it is desirable for a risk framework to communicate and preferably visualize the level of background knowledge (see also Goerlandt and Reniers, 2016) in order to determine whether the risk results are informative and can be used for decision-making, or should be used with great caution, or even the risk model shall not be used at all. In this regard, a solution for communicating the background knowledge is proposed in Section 5 of this thesis, which is also discussed in detail in Publication-IV.

2.2 An evidence-based risk modeling framework

It is discussed in Publication-III that a common way to divide the sources of data and information is to divide them into two groups as primary and secondary sources. The primary source is the one which is collected firsthand and directly from the source with no interference. These kinds of data or information can usually be collected from the installed recording devices that automatically collect and store the data from various available sensors, or from the witnesses who were directly involved in the observing event. Primary data are normally considered raw and with no (secondary) interpretations involved. Examples of such data are the data that is retrieved from the Voyage Data Recorders (VDR) onboard the ships, or the testimony of the witnesses of an event, normally shortly after the event. The secondary source, on the other hand, is the one that is prepared and collected based on the primary data, and normally some level of (secondary) judgmental interpretation and analysis are involved in those sources of information. Example of such data are official accident reports that are prepared by accident investigators after analyzing the data from various onboard sensors and also the testimony of the witnesses.
Based on the argument presented in Publication-I and III, the main idea of the evidence-based risk modeling framework as depicted in Figure 3 and presented in details in Publications-III and IV, is to start from the lowest hierarchy of the sources of data and information that is available, and use other sources on top of them to build-up the required knowledge for the risk modeling. This way, the secondary sources can normally be tested against the primary sources, in case of availability, in order to increase the available background knowledge and thus decrease the uncertainty. With this way of building-up a risk model, the level of background knowledge and the attached uncertainty can be recognized and thus communicated to the end users of the risk model. This practice, suggested in Publication-I, is valuable from a risk management perspective (IMO, 2012, Aven, 2013a, Montewka et al., 2014b), where the ultimate concern is risk mitigation.

Figure 3. A framework for evidence-based risk modeling, replicated from Publication-III & IV (Mazaheri et al., 2015b, Mazaheri et al., 2016)
3. Evidence building for ship grounding risk modeling

Based on the proposed framework for evidence-based risk modeling, it is suggested in Publication-I and II to use the factors that are supported by evidence in ship grounding risk modeling. However, the question is what source of evidence can be used to support the affecting factors and how the required support can be extracted from those sources of evidence?

3.1 Accident and incident reports

In Publication-III, we argue that one of the main sources of evidence that is available freely and can be used for evidence-based risk modeling is accident reports that are prepared by expert accident investigators (Schröder-Hinrichs et al., 2011). Since obtaining primary data about an accident that has happened in the past is nearly impossible, using accident reports as a secondary source of data is unavoidable (Mazaheri et al., 2013). However, there are some concerns regarding using only accident reports for modeling. One is that the accidents are scarce in frequency, thus the number of scenarios that can be analyzed is limited (Ladan and Hänninen, 2012). To overcome this imperfection, one of the suggested solutions is to utilize incident reports (Rothblum et al., 2002), as incidents occur much more frequently than accidents (Bole et al., 1987). Besides, since incidents are governed by the similar mechanism and underlying factors that cause accidents (Harrald et al., 1998) but they did not end in actual accidents, analyzing the incidents may likely give insights about the in-placed risk control options that stopped the incident to become an accident.

Publication-III suggests that utilizing accident and incident reports is beneficial for evidence-based risk modeling. This is because accident and incident reports can be useful for uncovering the factors that have contributed to occurrence of a mishap as well as for evaluating the level of importance of each factor. Besides, the way that contributing factors are linked together may be understood from such reports.

Generally, accident and incident reports are in text format and the information first need to be extracted before one is able to utilize them. The extraction normally needs human efforts, thus the risk of human opinion subjectivity exists. There are some text-mining techniques that use machine-learning algorithms to eliminate the need for human efforts for extracting the information and thus cope with the human opinion subjectivity issue; see for example Tirunagari et al. (2012b), Tirunagari et al. (2012a), and Artana et al. (2005). However, still quite many challenges exist in this regard as the reports are written in different natural languages with their own abbreviations and no standard template, and also they often contain misspelling (Hänninen et al., 2013). Additionally, since most of the available data sets whether in categorical- or text-format are prepared by humans at some stages, they contain the views of their creators and thus some level of subjectivity anyway (ibid). Therefore, being aware of such possible
subjectivity, still utilizing human to review the reports and extract the embedded information is beneficial.

For this thesis and as sources of evidence to be used, 115 grounding accident reports from Finnish and British maritime authorities as well as 163 incident reports prepared and collected by Swedish and Finnish maritime organizations have been reviewed. From among 163 reviewed incident reports, 73 reports were collected voluntarily and by a feed-by-user database, which in Publication-III their quality to be used as a reliable source of evidence for the presented risk model in Publication-IV were questioned.

3.1.1 Framework to review accident reports

For this study accident reports are categorized as a secondary source of data, in which the reports are prepared from the primary data that the investigator obtained first-hand by interviewing the operators and analyzing the evidence, normally short time after an accident (Mazaheri et al., 2013). In maritime safety analysis, the official accident reports that are prepared by the accident investigation boards usually present valuable information regarding why and how an accident happens.

Since the accident reports were prepared in a systematic way by expert accident investigators, in order to uniformly extract the information from all the reports, a framework is needed for reviewing the reports. There are a handful of tools and frameworks available for accident and incident analyzing and reporting (Johnson, 2003), which are mostly based on linear or non-linear accident theories that deal with complex socio-technical systems. Since a ship and her interactions within maritime traffic is also a complex socio-technical system (Hollnagel, 2004), a redefined version of a well-established complex-linear method as Human Factors Analysis and Classification System (HFACS) framework is utilized in this thesis (Publication-III) to review accident reports.

HFACS, which is based on the linear accident theory of Reason Swiss Cheese (Reason, 1990), was initially developed to study the contribution of human elements in military aviation accidents (Shappell and Wiegmann, 1997, Shappell and Wiegmann, 2000). The framework was further developed to also cover other causal factors than human factors, namely environmental factors like machinery failures and meteorological conditions (Wiegmann et al., 2005). The success of the method in detecting the contributing latent and active failures in the accident analysis made the method popular in the field of accident analysis, which is vastly used in analysis of civil aviation accidents (Shappell et al., 2007) as well as the accidents in other domains like railroad (Reinach and Viale, 2006) and maritime (Chen and Chou, 2012, Chen et al., 2013). Reinach and Viale (2006) have further developed the method by adding an additional level of factors, namely “external factors”, to the initial four levels in order to cover the latent failures that come from outside a particular domain. The same practice is followed by recent studies that used HFACS; see for example Schröder-Hinrichs et al. (2011), Chauvin et al. (2013), and Chen et al. (2013).

Since every single accident is unique from its own perspective, frameworks like HFACS try to assign the unique causes of an accident into more global factors to give better understanding of the phenomena by cumulating the causes into frequent factors. In this regard, having a specific framework with more specialized global factors for each domain and purposes seems beneficial. Different versions of HFACS that are recently introduced in the maritime domain like HFACS-MA for general maritime accidents (Chen et al., 2013), HFACS-Coll for collision accidents (Chauvin et al., 2013), and HFACS-MSS for machinery space accidents (Schröder-Hinrichs et al., 2011) support this belief. Therefore, HFACS is revised in Publication-III and a specific version suitable for grounding accident analysis (HFACS-Ground; see Figure 4) is developed for this thesis by implementing factors that are more related to grounding accidents. HFACS-Ground is also built as a five-level framework and has many similarities
with HFACS-Coll and HFACS-MA. However, in addition to the factors that cover traffic control and piloting services as affecting factors on grounding accident, “infrastructure” is added as a latent failure subcategory to the “environmental factors” in order to cover the waterway complexity related issues that are proposed and discussed in Section 3.3 of this thesis as well as in Publication-II.

HFACS-Ground is thus used in Publication-III for reviewing 115 accident reports in order to extract the factors required to structure the risk model in Publication-IV. Additionally, in the reviewing process of the accident reports, a directed acyclic graph (DAG) is structured for every reviewed accident based on the provided factual information in the report to demonstrate the relation of the contributing parameters in the accident. Therefore, practically for every reviewed accident, in addition to the extracted information using the HFACS-Ground, a DAG is also structured. These DAGs are later used in the iterative process of constructing the final model explained in Section 4.1 of this thesis as well as in Publication-IV.

Figure 4. HFACS-Ground developed and presented in Publication-III (Mazaheri et al., 2015b)
3.1.2 Framework to review incident reports

For this study incident reports are categorized as primary sources of data because the reports are normally prepared by the mariners or other operators that were directly involved in the event; thus can present the information without any secondary interpretation. This categorization, however, may be challenged. Incident reports, although reported by the same person directly involved in the event, could be subjective as it may be influenced by the reporter’s own opinion and interpretation. Additionally, incident reports’ quality could be affected due to organizational issues such as blame culture or other incentives for reporting an incident. As is discussed in Publication-III this may apply to both voluntary and mandatory reports. Hence, may it be questioned, whether incident reports can really be considered as primary sources of data? Considering that the presence of possible judgmental interpretation in the incident reports are less likely to be secondary judgmental, rather the reporter’s own opinion and interpretation, and also considering presented building blocks of background knowledge and their hierarchy shown in Figure 3, incident reports can acceptably be placed below accident reports and be considered as primary sources of data for this study. The difference between various primary sources like traffic data and incident reports can later be managed by considering the uncertainty level and strength of the background knowledge for the used sources, as is discussed in Section 5 and more extensively in Publication-IV.

Currently, there are quite few available sources in the maritime realm that can be used for obtaining incident or near-miss data, of which not all are available for public use (Ladan and Hänninen, 2012). Besides, contrary to the accident reports that were prepared in a systematic analytical way by expert accident investigators, the currently available incident reports suffer from the lack of a systematic view of the event. As is discussed in Publication-III, this issue is even more severe for the reports that are prepared voluntarily in compare with the reports that are collected by an official organization as part of the routine daily duty of their officers. The voluntary incident reports can be prepared by people of different expertise that some may not even be maritime related; while the official maritime incident reports are normally requested by an organization involved in maritime services and as part of the routine for performing the daily tasks of the mariners. Therefore it is more likely that such reports be more structured and be prepared by people with a maritime background. It is shown in Publication-III that such reports are more reliable and thus have more potential to be used as background knowledge for evidence-based risk modeling in compare with the reports that are prepared voluntarily.

Nevertheless and irrespective of how the reports are prepared, incident reports are normally short and their qualities depend on the reporters’ skills (Hänninen et al., 2013). Incident reports, in general, have high potential to be biased, because the reports are normally prepared by the same person who was involved in the event. This has resulted for the reports to have different qualities with regard to the provided data and information. Therefore, using HFACS-Ground for analyzing incident reports is not practical, rather it might be misleading because normally active failures (see Publication-III) are the only causes that are reported by the incident reports; thus looking at the incidents from the same perspective as accidents, meaning from the failure of an action, could be misleading. In reality, rarely there is a chain of failed actions as such in incidents that can be analyzed; rather there is always one or more successful controlling option that has stopped an incident due to an active failure to become an accident. Therefore, for incidents, contrary to accidents, it makes more sense to find what has gone right (i.e. as planned) that stopped the transformation of an incident to an accident. In contrary, the chain of failed actions is completed in accidents; thus it makes more sense to investigate the accidents from the perspective on what has gone wrong. Overall, incidents are more informative in terms of the controlled options, while accidents are more informative in terms of chain of failures. Therefore in Publication-III, instead of HFACS-Ground, another approach and taxonomy as Safety Factor (SF) has been introduced and followed for reviewing the incident reports.
SFs are high level positive functions that are believed to be prerequisite for safe transport operations. SFs are initially drafted by Nisula (2014) and include the (airline) pilot competencies that were produced in the Evidence Based Training (EBT) project (IATA, 2013). The principles for creating SFs are:

1. A safety factor need to be a positive function and not failure condition or technical device
2. The set of safety factors should cover all high-level safety critical functions
3. Overlap among safety factors should be avoided

SFs that are defined based on the above principles are then refined and customized for maritime purposes within expert panel discussions at the Finnish Transport Safety Agency (TraFi) as part of an experimental project at TraFi. The customized SFs are phrased as positive functions like “Controllability of the ship”, “Availability of propulsion”, or “Awareness of ship position in relation to the correct safe route”. Such positive functions are desirable to detect the presented safety factors that acted as barriers and stopped the incident situation to become severe into an accident. However, for the purpose of analyzing the contributing factors in the incidents as well as comparing incident and accident events with each other, it was also necessary to have the negation of SFs like “Loss of controllability of the ship”, “Loss of propulsion”, or “Unawareness of ship’s position”. Therefore in Publication-III, 163 incident reports are reviewed to simultaneously find those SFs that were reported present in the event as barriers and also to find if the absence of any SFs contributed to the occurrence of the event. This means that not only the SFs are collected as safety barriers, but also the negations of SFs are collected as contributing factors in the incidents. The collected SFs from among 90 of the reviewed incidents (i.e. excluding the voluntary reports) are then used to support the factors required to structure the risk model in Section 4 of this thesis as well as in Publication-IV.

3.2 Traffic and accident data

In Publication-II, we argue that traffic data, which is extracted from AIS² data, is one of the primary sources of data that can be used for evidence building. The patterns and traffic characteristics that can be extracted from the AIS data, and in combination with the statistical data of the historical accidents as a secondary source of data, give the model makers the opportunity to understand and then model the behavior of traffic as well as individual ship types in various situations based on different factors like location and weather condition.

For this thesis and as sources of evidence to be used, the accident statistics of ship groundings in the Gulf of Finland between the years 1989 and 2010 (Figure 5 and Figure 6) as well as the AIS data of the Gulf of Finland for the year 2010 are utilized. The utilized accident data consists of inputs such as date, time, and geographical coordinates of accidents. The utilized AIS data, which are recorded on average every third minute, also consists of inputs like geographical position, speed, course over ground, heading, and type of the vessel. As is explained comprehensively in Publication-II, in order to utilize the data for evidence building, the data is first filtered and cleaned from the possible errors and then is sorted and structured into templates that can be easily used in the designed algorithms for analyzing traffic characteristics.

---

² AIS or Automatic Identification System is an automatic tracking system required by IMO (2003) to be installed on all ships of 300 gross tonnage and upwards engaged on international voyages, cargo ships of 500 gross tonnage and upwards not engaged on international voyages, and all passenger ships irrespective of size. AIS is capable of automatically providing information about the ship, like the ship's identity, type, position, course, speed, navigational status and other safety-related information to appropriately equipped shore stations like coastal authorities, other ships, and aircrafts.
For analyzing traffic characteristics, two algorithms are designed in Publication-II to extract the required patterns from the AIS data; one for extracting traffic density and one for extracting traffic distribution. In the existing literature on grounding risk modeling, ship traffic is defined in two ways as density and distribution (Mazaheri and Ylitalo, 2010). Traffic density is defined as the number of ships per unit area within a time window; and traffic distribution is defined as the lateral distribution of ship tracks over a waterway.

The algorithm for extracting traffic density is basically designed to count the number of ships, which their path tracks pass through a predefined area within a year (Figure 7). However, since the objective of the study was to assess the effect of traffic density on the grounding accident, the algorithm only calculates the traffic density for the areas that have previously been subject to a grounding accident. The algorithm is designed in this way to reduce the required calculation effort and thus to increase the performance.

The algorithm for extracting traffic distribution generates distributions of minimum distances of ship tracks from the spot where a grounding accident has previously occurred according to the utilized accident statistics (Figure 8). For the same reason as above (i.e. reducing calculation effort and increasing performance), the designed algorithm only takes the waterways that have previously been subject to grounding accidents, and generates the lateral distribution of ship tracks over that specific waterway (Figure 9).

Both algorithms are then coded and run in MATLAB® to extract the required traffic characteristics for the study. The existence of possible statistical dependency between ship traffic (i.e. density and distribution) and number of grounding is studied using three correlation coefficients (i.e. Pearson, Spearman, and Kendall) as well as Mutual Information (MI) test. The reliability of the resulted coefficients then tested with $\chi^2$ – goodness-of-fit test assuming the null hypothesis ($H_0$) as “zero-correlation” against the alternative hypothesis ($H_1$) of “non-zero correlation” using significant level of 95%. The results show no significant statistical dependency between traffic density and grounding frequency. Nevertheless, the MI test shows the traffic distribution can give us information regarding the frequency of ship grounding. The results are comprehensively explained and discussed in Publication-
II, and are also used to support the factors required to structure the risk model in Section 4 of this thesis as well as in Publication-IV.

Figure 6. Characteristics of groundings accidents in Gulf of Finland between years 1989 and 2010 based on HELCOM Statistics (Mazaheri et al. 2015a)
Figure 7. Developed algorithm in Publication-II to extract traffic density from AIS data, and find the possible relation between traffic density and number of grounding accidents (Mazaheri et al. 2015a)
Figure 8. Developed algorithm in Publication-II to extract traffic distribution from AIS data, and find the possible relation between traffic distribution and number of grounding accidents (Mazaheri et al. 2015a)
3.3 Expert opinion

Primary and secondary sources of data in most cases are scarce sources of data in compare with the complexity of the accident model and the including parameters that need to be investigated. Besides, some sources like accident and statistical data cannot be analyzed over a long period of time due to the change in the conditions such as regulations that affect the integrity of the collected data. In such situation, using expert opinion, aka expert knowledge, with different background and expertise can be seen as a reliable and rich source of information to be used for modeling (Hänninen et al., 2014). Additionally, in the areas that no primary or secondary data are available, expert knowledge is the only source of knowledge that can be used for building the required evidence. Therefore, it is suggested by the proposed framework (Figure 3), expert knowledge should be used on top of the primary and secondary sources of data to fill the gap of the knowledge in the built evidence.

It is argued in Publication-II that despite the commonly accepted belief that grounding accident is a location dependent accident, the difficulty of a waterway to safely navigate a ship (i.e. complexity of a waterway) has not been used in any available risk models for grounding accident so far, nor the dependency of frequency of grounding accidents and waterway complexity has been studied before (Mazaheri et al. 2015a). However, since waterway complexity is inherent to the design and geographical location of a waterway, it can be seen as a valid factor that affects frequency of grounding accidents, thus worth further analysis.
In this regard and in Publication-II, we argue that since currently no established reliable method exists to assess complexity of a waterway with regard to navigating a ship, expert knowledge can be used for that purpose. Moreover, complexity of a waterway in its conceptual meaning is a term that should naturally be felt and defined by human operators. Introducing a method or model in this regard could only ease the quantification of the human operators’ feelings regarding the complexity of a waterway based on different factors. Thus, using expert knowledge in this context is not only the sole available data source but also the most logical data source to be used. To acquire expert knowledge in this thesis, semi-structured interviews as a qualitative method is used in Publication-II. This method is considered acceptable because qualitative methods are often successfully used for studies related to human activities (Robson, 2008).

For this thesis and to assess the complexity of a waterway, a location dependent and semi-quantitative index, called *Waterway Complexity Index (WCI)*, is defined. WCI ranks fairways with regard to the difficulty of safely handling a ship through the fairway. Therefore, WCI in this thesis is quantified through the conducted semi-structured interviews and based on the experience of the local pilots.

The focus of the study was mainly on fairway dimensions and geometry. The aim of the interviews was to make a numerical WCI for typical ships navigating in the studied fairways. 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 markings, especially in areas which require increased attention of a pilot, e.g. due to significant course alteration

The statistical analysis between WCI of the fairways and the grounding frequency shows rather strong negative linear correlation (Pearson correlation coefficient $r = -0.71$), which means the more complex a waterway the higher frequency of grounding in that waterway is. This result together with the quantified WCIs are used to support the factors related to location and waterway complexity in the constructed risk model in Section 4 as well as in Publication-IV.
4. Evidence-based risk model for ship grounding

Probabilistic causal modeling is known as one of the most suitable methods for modeling the risk of complex systems with high uncertainty like maritime transportation system (Ozbas 2013; Goerlandt and Montewka, 2015a; Hänninen et al., 2014; Montewka et al., 2014a; Fu et al., 2016). Therefore, a stepwise approach, which is originally proposed by Kragt (2009) and further modified by Akhtar and Utne (2014) to account for the specificity of maritime accidents, is followed in this thesis for constructing a probabilistic causal model for ship grounding based on the proposed framework. The approach has five steps as follows:

1. Define the system with its characteristics and boundaries (performed in Publication-I to IV)
2. Collect the required data (performed in Publication-II & III)
3. Develop the model’s structure (performed in Publication-III & IV)
4. Parametrize the model (performed in Publication-IV)
5. Validate the model (performed in Publication-IV)

So far steps 1 and 2 are established and explained in Sections 1 to 3. The remaining three steps, which are fully discussed in Publication-III & IV, are briefly explained in the following.

4.1 Model construction

In Bayesian Network (BN) modeling, model construction refers to choosing the relevant parameters for the model as nodes, as well as the way those nodes are interconnected, that is, the links between the parameters as edges. Model construction in this thesis was an iterative process. This means an initial model is constructed first based on the cumulated contributing factors from the accident and incident reports. Thereafter, the model is iteratively evolved towards the final model by comparing it with the individual DAGs, previously created during the reviewing process of the accidents.

In the first round of the iteration, the variables that are extracted as the contributing parameters from the reviewed accident/incident reports form the initial nodes of the model; and a link is made between every two parameters that shown statistical dependency (see in Publication-III & IV). Afterwards, the states of the nodes are defined using only two states; meaning the presence or absence of the contributing parameter. This is a simple and safe way to discretize the variables when they are supported by reports because each report only expresses whether a parameter was present at the time of the accident/incident or whether it was absent. Nevertheless, during the iteration process and during the parameterization phase (see Section 4.2), the states of the nodes are adjusted when more information is obtained. This way of discretization is performed purely based on the evidence at hand, and other sort of discretization that might be regarded as rather arbitrary is avoided at this study.

In the next rounds of the iteration, the resulting network is iteratively cross-checked with the individual DAGs of accident reports in order to define the direction of the added links as well as to retrieve the
parameters and connections that may have been lost during the combination process. Here, the combination process is the process in which the individual parameters are accumulated into the global parameters used in HFACS-Ground and SF. As the result, every node and link in the final model have acceptable level of support in the utilized evidence (see Publication-IV). The resulting model has 32 nodes (excluding the Grounding node) as contributing factors on the grounding accident; the nodes are interconnected with 49 edges, including the links to the Grounding node (Figure 10).

Figure 10. The evidence-based BN model for grounding accidents constructed in Publication-IV (Mazaheri et al, 2016)

4.2 Model parametrization

Model parametrization is referred to the step where the prior and posterior probabilities of the nodes are filled. In principle, there are two common practices to fill the conditional probability tables (CPTs) of a BN. These practices are not necessarily exclusive, but each has its own uncertainties involved.

One common method is to use techniques that allow elicitation of experts’ opinions on the likelihood of the occurrence of each scenario defined by the model, see for example DNV (2003), Hänninen et al. (2014), Goossens et al. (2008), Cooke and Goossens (2004), and O'Hagan et al. (2006). Another way is
to use the statistical data and try to find the statistics related to each scenario represented by the constructed model. However, there are several issues that need to be considered when one wishes to use the statistical data. First, there is the fact of under-reporting accidents. Nonetheless, based on studies regarding the under-reporting phenomena in the maritime realm (Hassel et al., 2011, Thomas and Skjong, 2009, Psarros et al., 2010), one can judge the related uncertainty and thus it can be compensated. Second, according to the Safety Investigation Authority of Finland, not all the reported accidents to the authorities are investigated (SIAF, 2014). Therefore, using the accident reports as the statistics for estimating the CPTs can under- or overestimate the probabilities, which translates to uncertainty. Unfortunately, the studies for under-investigated cases do not exist, thus the related uncertainty remains unknown. Third, each chain of events leading to an accident is different; the involved vessels, crew, external and internal conditions, or the situational context are different in each accident. This means that the exchangeability of events, which is a prerequisite for the frequency assessment, is not valid (Apostolakis, 2014).

As the results of the above discussion, the limitation of the available data on under-reporting and under-investigated, and also taking into account the lack of exchangeability, the experts’ judgment is utilized in this thesis (Publication-IV) as the main approach to fill the CPTs of the constructed model.

The CPTs of the nodes are filled using the probabilities presented in the existing literature that have utilized expert opinions for defining the conditional probabilities of different scenarios for a grounding accident (Hänninen et al., 2014, Haapasaari et al., 2014, DNV, 2003). Additionally, the results of the analysis in Publication-II on the statistics of the grounding accidents are used to fill the CPTs of the nodes. In this regard, if the conditional probability of a scenario in the constructed model is recognized in either of the mentioned sources, the probabilities are filled using the same values. However, in case that a scenario in the constructed model is not matched with any scenario presented in the mentioned sources, the probabilities are filled using scenarios similar to those detected in the dataset created in Publication-III.

4.3 Model validation

The last and probably the most important step of the taken approach for the model construction is validation. Validation of a model means finding whether the built model actually fulfills the purpose for which it was built, given that it confers the specification of the system that it is representing (Wentworth, 2012, Marwedel, 2011).

Model validation can be understood in a wider sense than as a comparison with observed data, by inspecting the model qua model. Such approaches to validation are widely used in social science research (Trochim and Donnelly, 2008), system dynamics modeling (Forrester and Senge, 1980), expert-based Bayesian Network modeling (Pitchforth and Mengersen, 2013) and recently in the maritime risk modeling (Zhang et al., 2015, Goerlandt et al., 2015c, Goerlandt and Montewka, 2014, Pitchforth et al., 2014, Fu et al., 2016). Therefore, a fairly similar holistic validity framework with six types of validity tests is adopted for this thesis (see Table 2). The framework is applied on the constructed model and the results are fully discussed in the following sections as well as in Publication-IV.

In general, the adapted model validation approach has two major phases. The first phase is to evaluate whether the model adequately operationalizes the construct it intends to measure, i.e. how well it concretizes the object of inquiry for the given purpose. This is evaluated in terms of Nomological and Construct validities. The latter is broken down into face and content validities. Face validity is a subjective, heuristic interpretation of whether the model is an appropriate operationalization of the
construct. Content validity is a more detailed comparison of the elements in the risk model in relation to what is believed to be relevant in the real system.

In the second phase, a number of specific tests can be performed on the model to evaluate whether the model adequately meets certain criteria. Here we refer to it as criterion validity. A behavior sensitivity test (BST) is used to assess to which elements of the model the results are sensitive. The parameter sensitivity of a model can be calculated and the results can be evaluated by domain experts. In a qualitative features test (QFT), the response of the model is evaluated for a number of test conditions in terms of a qualitative understanding on how the system is believed to respond under these conditions. In a concurrent validity test, the model elements are compared with the elements in another model for a similar purpose. This can also include a comparison with the output of such a model if the scope of the applications is the same.

One should bear in mind that the validity tests do not “prove” the model results are correct; they only indicate the extent to which the model is a plausible representation of the object of inquiry. The model should be plausible enough to serve as a basis for further reflections, leading to deliberative judgments in the risk analysis (Goerlandt and Montewka, 2015a, Zhang et al., 2015). Moreover, the validity of the model can only be established to the extent of the validity of the used evidence in the validation process; thus the model cannot be validated beyond the strength of the knowledge of the used evidence.

Table 2. Type of validity tests for BN models used in Publication-IV (Adapted from Pitchforth and Mengersen, 2013)

<table>
<thead>
<tr>
<th>Validity Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nomological validity</td>
<td>Confident that the model domain fits within a wider domain established by the literature.</td>
</tr>
<tr>
<td>Face validity</td>
<td>Approval from the domain experts that the model looks as expected.</td>
</tr>
<tr>
<td>Content validity</td>
<td>Confident that the available information is integrated into the model via the nodes, links, and discretization as well as parametrization of the nodes.</td>
</tr>
<tr>
<td>Concurrent validity</td>
<td>Confident that the whole network or sections of it behave identically to sections or another network from a similar domain.</td>
</tr>
<tr>
<td>Qualitative Features Test</td>
<td>Showing that the model structure, discretization, and parameters are similar to the nomologically proximal models.</td>
</tr>
<tr>
<td>Behavior Sensitivity Test</td>
<td>Approval that the behavior of the model as well as its sensitiveness to the constructed components is what is expected of the modeled system.</td>
</tr>
</tbody>
</table>

4.3.1 Nomological Validity

Nomological Validity discusses the position of the constructed model within similar existing models in the context of the studied phenomena, which here is ship-grounding accident. This type of validity needs extensive literature analysis to find and to understand the existing models in the realm under the study (Pitchforth et al. 2014). In this way, a nomological map is drawn to estimate the distance of the
constructed model from the existing models in four areas of uncertainty in BN modeling, as follows (Pitchforth et al. 2014):

1. Structure of the model, which includes the nodes that are included in the model as well as the way those nodes are interconnected (i.e. the edges)
2. Discretization, which is about the way the states of each node are defined.
3. Parametrization that defines the conditional probabilities of the nodes and their states
4. Behavior of the model, which describes the uncertainties related to the output of the model

The distance for each analyzed pair of models is indicative only, showing the position of the model among other solutions, for four analyzed areas (see Figure 11). To determine the distance, subjective judgment is used (Goerlandt and Montewka, 2015b). In the extensive literature analysis performed in Publication-I, we have detected three practical BN models currently existing for ship-grounding accidents. The models are listed in Table 3 as No. 1, 2 & 3 together with three recent BN models from other proximal realms. The nomological map for the models is shown in Figure 11. However, due to different modeling choices adopted and availability of the models, not all the pairs could have been compared in all uncertainty areas.

Observing distances that the constructed model in different areas of uncertainty has with the existing models in the maritime domain (Figure 11), specifically in the grounding accident domain, the nomological validation can be confidently considered as established for the constructed model.

**Table 3.** Existing BN models for ship-grounding accidents discussed in Publication-I and comparable BN models in other domains presented in Publication-IV (Mazaheri et al, 2016)

<table>
<thead>
<tr>
<th>No.</th>
<th>Nodes</th>
<th>Edges</th>
<th>Source</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>33</td>
<td>49</td>
<td>The current model</td>
<td>Transportation/Maritime/Grounding</td>
</tr>
<tr>
<td>1</td>
<td>69</td>
<td>107</td>
<td>DNV 2003</td>
<td>Transportation/Maritime/Grounding</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>136</td>
<td>Hänninen et al. 2014</td>
<td>Transportation/Maritime/Grounding</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>48</td>
<td>Rambooll 2006</td>
<td>Transportation/Maritime/Grounding</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>27</td>
<td>Wang et al. 2013</td>
<td>Transportation/Maritime/Collision</td>
</tr>
<tr>
<td>5</td>
<td>62</td>
<td>85</td>
<td>Ancel et al. 2015</td>
<td>Transportation/Aviation</td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>30</td>
<td>O’na et al. 2011</td>
<td>Transportation/Road</td>
</tr>
</tbody>
</table>
4.3.2 Construct Validity

As is mentioned before, the construct validity of the model comprises face and content validities, which are described below:

**Face Validity**
Face validity is established through an iteration process, explained in Section 4.1 as well as in Publication-IV, when the preliminary established network is iteratively checked with the individual DAGs from the accident reports. In the iteration process, the added nodes and edges in the network are all iteratively checked and confirmed with the individual DAG of each accident. Any possible disagreements between DAGs of different accidents are assessed and confirmed by the researchers who are experts in the field of ship accident modeling. Therefore, since the utilized methodology involves domain experts in the model construction and seeks their approvals on the involved contributing parameters as well as the connected links, the face validity of the model is considered as established.

**Content Validity**
Content validity checks whether the parameters or relationships that are considered as important by the available literature are included in the model. As is explained in Section 3 as well as in Publication-II&III& IV, the contributing parameters in the model and their relations are directly extracted from the real grounding scenarios and statistics; thus they are considered as supported by evidence. Additionally, as is discussed in Section 4.2 and Publication-IV, the model is parametrized using the available expert
knowledge as well as all the available data in hand. Therefore, the content validity of the model can be considered as confidently established, and thus the construct validity.

### 4.3.3 Criterion Validity

The criterion validity in this study is broken down into three elements, namely concurrent validity, qualitative feature and behavior sensitivity tests.

**Concurrent validity**

Concurrent validity is about the behavior of the whole model or a section of it, in comparison with other models developed for similar purposes.

The nomological map in Figure 11 shows that the current model is fairly close to the identified proximal models in the literature within the realm of maritime transportation, especially the ones related to ship grounding accidents. It can also be seen in Figure 11 that the distance of the current model from proximal models is increasing when we move from the models in maritime transportation system towards the models in other modes of transportation. The only exception is Rambøll (2006) model, which although is from the maritime domain, the structure of the model stands by a distance from the current model. This, however, is expected, because although both models are describing the same phenomena, the cores of the models are different. The core of Rambøll (2006) model is based on imaginary geometrical scenarios that are rooted on the scenarios first defined by Pedersen (1995), while the core of the presented model in this study is based on the interactions between different parts of the system (see Figure 2) that themselves are supported by real life scenarios as evidence.

The accident models in the maritime domain normally consist of the sections that account for the factors related to 1- hazardous situation; 2- danger detection; and 3- accident avoidance by evasive actions. It is shown and extensively discussed in Publication-IV that the constructed model has ten general sections which all can be fitted within one or two of the groups of factors for similar models in the domain. Sections with similar functionalities also exist in DNV (2003) and Hänninen et al. (2014) models, as the most proximal models in the literature of the same domain. This basically means that all three models follow the same logic and thus should behave the same way. This observation, as is also discussed in Publication-IV, assures us that the concurrent validity of the model is acceptable.

**Qualitative Feature Test**

Qualitative features test (QFT) evaluates a model for a number of test conditions to check the model response with the actual understanding of the system behavior under these conditions.

In Publication-IV, we observed the probability of grounding (i.e. Grounding = “yes”) while setting each state of each variable in a turn. The difference in the probability of grounding for a state producing the largest probability and a state corresponding to the smallest probability is recorded for each variable (for detail analysis please refer to Publication-IV). The difference \( \Delta P_G \) informs us about the behavior of the model and helps to evaluate the most influential variables:

\[
\Delta P_G = \max P(Grounding = "yes" | X = x_i) - \min P(Grounding = "yes" | X = x_j)
\]

(4.3.3.1)

where \( x_i \) and \( x_j \) are the states of variable \( X \) that produce the largest and smallest grounding probabilities, respectively. This difference describes the maximum change that variable \( X \) could cause on the model output.

The results of this analysis, as presented in Figure 12 and discussed in Publication-IV, show that there are two variables producing the largest changes in the model output if they are set at either of their states.
These are “Being off Course” and “Loss of Control”, which seems to follow the expectation regarding the system under analyzed and other similar systems; see for example Hänninen and Kujala (2012), Hänninen et al. (2012), and Hänninen et al. (2014). The next most influential variable is “Waterway Complexity”, which has almost the same level of effect on the model as “Technical Failure”. The latter also follows the expectation regarding the system under the study; and not surprisingly the former is also aligned with the statistical results discussed on Section 3.3 on the relation between WCI and grounding frequency, thus satisfying the expectations.

![Figure 12. The top 15 variables producing the largest difference in grounding probability when at their worst state (producing the largest $P_G$) and best state (producing the smallest $P_G$) presented and discussed in Publication-IV (Mazaheri et al, 2016) ](image)

**Behavior Sensitivity Test**

Behavior Sensitivity Test (BST) confirms if the model correctly predicts the behavior of the system that is modeled. The test basically checks the sensitivity of the model to see if the model is sensitive to the parameters and scenarios that the system itself is also expected to be sensitive to.

As it is explained in Publication-IV, to perform one-way sensitivity analysis in the way that is described in Castillo et al. (1997) and Kjaerulff and Gaag (2013), the sensitivity analysis tool of GeNle™ software from the Decision Systems Laboratory of University of Pittsburgh is used. To perform the analysis, first the sensitivity function should be developed, defining the output probability of interest as a function of the parameter $Y$; (see for example Coupé and van der Gaag, 2002), van der Gaag et al., 2007)

38
\[ f(Y) = \left( \frac{c_1 Y + c_2}{c_3 Y + c_4} \right) \]  

(4.3.3.2)

where \( f(Y) \) is the output probability of interest given observations, and \( c_1 \ldots c_4 \) are the constants, which are identified based on the model. Parameter \( Y \) is calculated as follow:

\[ Y = p(X = x_i | \pi) \]  

(4.3.3.3)

where \( x_i \) is one state of the parameter \( X \), and \( \pi \) is a combination of states for \( X \)’s parent nodes. When the value of \( Y \) is being varied, the other states of the same parameter \( X \); \( p(X = x_j | \pi), j \neq i \); should also be varied in order to keep the probability of the sum of the states related to \( X \) equal to one (Hänninen and Kujala 2012).

The first derivative of the sensitivity function \( f(Y) \) describes the effect of minor changes in a variable on the output and is called the sensitivity value.

In publication-IV, we have analyzed the sensitivity to changes in the posterior probabilities for the following:

- a case where “Grounding” node is set as a target,
- a case where “Being off Course” and “Loss of Control” nodes are set as the target nodes. With respect to these two nodes, the model is at its most sensitive, as detected by QFT method.

When the target node is set as “Grounding”, there are no evident dominating variables in the model; see Figure 13. This supports the statement about the complexity of the topic analyzed. On the other hand, the uncertainties associated with the parameters of such complex model affect the model output to a lesser extent. For this reason, such insensitiveness of the model to the changes of the parameters is desirable, as argued by Pradhan et al. (1996) and Henrion et al. (2013).

However, when the target nodes are set to “Being off Course” and “Loss of Control”, it becomes evident that the model is sensitive to several nodes, including “Traffic distribution”, “Technical failure”, “Detection”, “Lack of training” or “Incapacitated”; see Figure 14. Such analysis allows for more detailed assessment of the grounding model and its most sensitive elements.

Evaluating the results of the conducted sensitivity analyses shows that the model sensitiveness to the different parameters of the model and to different combinations of parental states (i.e. scenarios) is in accordance with the expectations of the modeled system. As the result, the final validity test as BST is confidently established so as the criterion validity.
Figure 13. Sensitivity analysis of the model presented in Publication-IV by setting Grounding as the target node. The model’s output as probability of grounding is almost insensitive to all parameters of the model except the Grounding node (Mazaheri et al, 2016)
Figure 14. Sensitivity analysis of the model presented in Publication-IV by setting “Being off Course” and “Loss of Control” as target nodes (Mazaheri et al, 2016)
As is explained in Publication-I and also in Section 2 of this thesis, considering maritime transportation as a system, a well-defined approach by Haimes (2009) and Aven (2011) can be followed for defining the risk of the system as $R = (S, L, C|BK)$ (Mazaheri et al., 2014), where $S$ is the scenario for a mishap to occur, $L$ is the likelihood of that specific scenario to occur, $C$ is the consequence of that specific scenario if it occurs, and $BK$ is the amount of available background knowledge about the system. $BK$ is the representation of our incomplete knowledge about the system, and that the risk is formulated merely based on our best knowledge about the system. This incompleteness should always be visualized and communicated to the end users of risk models (Aven, 2013a, Montewka et al., 2014b, IMO, 2012, Mazaheri et al., 2014, Goerlandt and Reniers, 2016). This is essential when the risk model is used for decision making and the $BK$ level is weak, thus the uncertainty is high. Such situation may result in risk estimates that falls in the acceptable level, whereas the associated uncertainty may be larger than the margin between risk level and acceptance boundaries, finally resulting in a situation where the risk shall not be deemed acceptable (Aven, 2015, Montewka et al., 2014b).

The $BK$ in this thesis is understood as a mixture of knowledge, understanding, beliefs, and acceptance about the analyzed phenomena (Aven, 2013a, Montewka et al., 2014b). The adopted concept, which originates from the earlier work of Aven (2013b) and is applied by others in the maritime field (Goerlandt and Montewka, 2015a, ValdezBanda et al., 2015), seems the best-suited for our purpose. Assigned to the parameters of a model, there are three levels of uncertainty based on the amount and quality of the $BK$ that was available as evidence for a particular parameter at the time of the model construction (see in Table 4).

At a simple glance and using the presented guideline in Table 4, Figure 15 visualizes the strength of the $BK$ for every parameter (i.e. node) and its parental relations (i.e. edges) of the current model based on the sources of evidence that are utilized to support the existence of the parameter or the link in the model (see also the appendix of Publication-IV). It also visualizes the strength of the $BK$ of the probabilities implemented into the model according to the sources where the probabilities are estimated from.

The colors of the border of the nodes in Figure 15 show how certain the model makers are regarding the existence of the parameter in the model and the role that the parameter plays in the modeled phenomena. The color of the center of the nodes shows the uncertainty of the implemented prior and posterior probabilities; and likewise, the colors of the arrows show the certainty of the model makers on the existence of the presented parental relation between the parameters. Therefore, based on the presented knowledge strength map of the model (Figure 15) as well as the sensitivity analysis of the model presented in Section 4.3.3, the outcome of the model can be assessed for the scenarios of interest that are studied as the risk of the system. This basically means if a certain scenario is being studied, the
model can only be used reliably to assess that scenario if the nodes and links that are involved in that scenario have an acceptable level of knowledge strength. Otherwise, the sensitivity of the model to the nodes involved in the scenario should also be considered. This means the results of the model should be very cautiously interpreted for assessing a scenario that the level of knowledge strength is not acceptable and the model is too sensitive to the nodes involved in the scenario. The level of acceptability should, however, be decided by the decision maker and according to its utility value (Grabowski et al., 2000)

Table 4. The guideline employed for categorizing the strength of the evidence in Publication-IV (Mazaheri et al., 2016) based on the early works of Aven (2013b)

<table>
<thead>
<tr>
<th>Level of knowledge strength</th>
<th>At least two of the following conditions are met</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH</td>
<td>High number of supporting cases as evidence (normally more than nine cases)</td>
</tr>
<tr>
<td></td>
<td>High reliability of the data/information sources used</td>
</tr>
<tr>
<td></td>
<td>High detected statistical dependencies</td>
</tr>
<tr>
<td></td>
<td>High accuracy of the method used for the data analysis</td>
</tr>
<tr>
<td></td>
<td>High acceptance of the data sources used among the domain experts</td>
</tr>
<tr>
<td>MEDIUM</td>
<td>Conditions between the characteristics of the high and low strength knowledge</td>
</tr>
<tr>
<td>LOW</td>
<td>Low number of supporting cases as evidence (normally fewer than four cases)</td>
</tr>
<tr>
<td></td>
<td>Low reliability of the data/information sources used</td>
</tr>
<tr>
<td></td>
<td>Low detected statistical dependencies</td>
</tr>
<tr>
<td></td>
<td>Low accuracy of the method used for the data analysis</td>
</tr>
<tr>
<td></td>
<td>Low acceptance of the data sources used among the domain</td>
</tr>
</tbody>
</table>
Figure 15. Knowledge strength map of the model shows the strength of knowledge of each node (as the color of the border of the node) and edge (as the color of the link) as well as the conditional probabilities of each node (as the color of the center of the node) based on the sources of the evidence presented in Publication-IV (Mazaheri et al., 2016)
6. Conclusion

This compendium has summarized the major points of the four published articles as a united entity to meet the five objectives that are defined for this thesis. Lack of an evidence-based approach to ship-grounding for building suitable risk models for risk management purposes in the way to satisfy the requirements mentioned in FSA guideline by IMO (2012) was shown in Publication-I. This is further supported in Publication-II by showing how many of the existing factors in the current models are actually supported by evidence. Thus, combination of Publication-I&II addresses Objective 1 of the thesis, which is to understand the state of the art on ship grounding risk modeling, and to assess the usability of the current existing models for risk management and decision making purposes.

Thereafter, and in response to the findings in Publication-I&II, an evidence-based framework to build probabilistic risk models suitable for risk management purposes is proposed in Publication-I&III (Figure 3). The proposed framework, which satisfies Objective 2 of the thesis, attempts to help the model makers to develop models that efficiently maximize the decision makers’ utility values by communicating the level of the background knowledge of the whole model. This approach is aligned with the doctrine of utilitarianism, as presented by Kaplan (1997), which helps the decision makers to decide whether they should accept the resulting risk of the model or not; because not all of the decision makers have the same priorities and not even the same degree of feeling about the risk, they have different utility values. According to Aven (2013b), the strength of knowledge can help the decision makers to decide whether to accept the risk or not, only if the calculated risk by the utilized risk model is acceptable for them and thus satisfies their utility values.

According to the proposed framework in Publication-I&III, the first step in evidence-based risk modeling is to build the required evidence. This is defined as Objective 3 of the thesis. The required evidence is built and presented in Publication II&III to meet that objective. Traffic characteristics of ships navigating in the Gulf of Finland are extracted in Publication-II using the AIS data and accident statistics of the same waterbody, and with the help of two algorithms which are developed for that purpose. Moreover, a location dependent index as Waterway Complexity Index is defined and quantified for five waterways in the Gulf of Finland and based on the expert opinion of the local pilots, to rank waterways with regard to the difficulty of safely navigating a ship. It was found and shown in Publication-II that ship grounding frequency is statistically dependent on traffic distribution and waterway complexity, and thus traffic distribution and waterway complexity can be reliably used for evidence-based risk modeling of ship grounding according to the proposed framework. It is, however, found that traffic density is not evidently reliable to be used for the same purpose. Additionally, the usability of two other sources of evidence as accident and incident reports on evidence-based risk modeling are also tested in Publication-III. To build the evidence and extract the required knowledge from the accident and incident reports, two reviewing frameworks are developed and suggested; one for accident reports as HFACS-Ground and one for incident reports as Safety Factor. The developed frameworks are then used to review 115 accident reports from Finnish and British authorities, and 163 incident reports from Finnish and Swedish authorities. What is found was that accident reports have
acceptable levels of reliability to be used for evidence-based risk modeling according to the proposed framework. The same was true for incident reports that are prepared and reported by maritime organizations as part of their daily routine. Nevertheless, it is found that the incident reports that are prepared voluntarily have lower reliability, in compare with other two sources of evidence, to be used for evidence-based risk modeling.

The built evidence are then used in Publication-IV to construct and validate an evidence-based Bayesian Network risk model for ship grounding in order to satisfy **Objective 4** of the thesis, which is building and validating an evidence-based risk model.

Following the proposed framework in Publication I&III and also according to what is highlighted in FSA guideline, it is necessary to expose the assumptions and uncertainties behind the model’s elements to the decision makers, who are the ultimate users of any risk model. Therefore, in order to fully and operationally satisfy the requirements regarding the uncertainty discussion, the proposed framework in Publication-I&III not only gives model makers the opportunity to use the least subjective evidence at hand to initiate the model structure and use the more subjective evidence as well as expert opinions on top of them, but also suggests to record and later clearly expose the strength of the background knowledge of the utilized evidence to the end users of the model. This concern is addressed under **Objective 5** of this thesis. In Publication IV, in addition to the risk model, the users of the model are equipped with a knowledge strength map of the model that clearly and efficiently communicates the uncertainty of the model’s elements to them in one glance. This concludes the thesis by meeting the last objective of the thesis.

It can be argued that the constructed model stands above the current state of the art of ship grounding risk models. It is shown in Publication-IV that the constructed model is structurally simpler with fewer nodes and edges, in compare with the existing comparable models. This makes it easier for the decision makers to understand the mechanism of the model. When this combines with the knowledge strength map of the model, it can fully and operationally satisfy the requirements intended by IMO (2012) because it will let the decision makers decide whether to trust the results of the model or not.

The model can be used to suggest areas that need to be controlled by proper risk mitigation measures or areas that need more study to be fully understood. In this regard and within the limitation of the utilized evidence in the model construction, the model suggests the critical parameters that need proper control measures in order to mitigate the risk of ship grounding are “complexity of waterways”, “traffic encounters”, and “a ship being off course”. The critical area that calls for more investigation and study, in addition to the above-mentioned parameters, is “the onboard presence of a sea-pilot”.

With this being said, what makes the model stands out from its predecessors is the extensive use of available evidence and clear communication of the involved uncertainty. This makes it possible to define the areas of the model that can be improved whenever more data sources are available. This can be done either by further supporting the current structure with more evidence to increase the strength of knowledge behind each parameter and its parental relation, or by modifying the current structure when more reliable and trustworthy sources of evidence suggest adding or eliminating a certain parameter or link. These are seen as the potential areas to be improved by future works. Moreover, since the focus of the thesis was mostly on scenario (S) creation of risk (according to the taken definition of risk in this thesis), the proposed framework can be tested by applying it on the other parts of risk definition as Likelihood (L) and Consequence (C) in the future works. Furthermore, the proposed framework can be tested further on studies related to other types of maritime accidents as well as other domains. Additionally, waterway complexity in this thesis is assessed and quantified merely by using expert opinions and only via qualitative methods. Therefore and as a suggestion for future works, defining
quantitative methods for assessing waterway complexity only based on the characteristics of the waterways, such as number and quality of turns, length of the legs, and width of the fairway, will likely increase the knowledge strength behind the waterway complexity as a factor to be used in evidence-based risk modeling.
References


SIAF (2014) Marine Investigation Reports. Safety Investigation Authority of Finland.


Errata

Publication I

- In Section 3.3 Probabilistic Models, “By a FTA, the grounding accident as the top event is broken down into its initiating factors, and by the Bayesian approach, the key factors that contribute to the accident are recognized together with the degree of belief associated with the states of each factor, and they are taken into account as a Bayesian Belief Network (BBN)” should read “By a FTA, the grounding accident as the top event is broken down into its initiating factors, and by the Bayesian Network approach, the key factors that contribute to the accident are recognized together with the degree of belief associated with the states of each factor, and they are taken into account as a Bayesian Belief Network (BBN).”

- In the caption of Figure 6, “Overview of the Bayesian grounding model for passenger ships” should read “Overview of the Bayesian Network grounding model for passenger ships”

Publication II

- In Section 1 Introduction, “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” should read “Moreover, it is possible that researchers are convinced that navigating a ship is more difficult when traffic is denser; however the question is whether navigational difficulty due to traffic necessarily increases the probability of GA”

Publication IV

- In the caption of Table A2, “The states producing the largest and smallest grounding probabilities, the corresponding probability values, and their differences for all the parameters of the network” should read “The states producing the largest and smallest grounding probabilities, the corresponding probability values, and their differences for all the parameters of the model”
A framework for evidence-based risk modeling of ship grounding

Arsham Mazaheri