NTNU
Norwegian University of Science and Technology
Department of Marine Technology

MASTER THESIS

Title: Underreporting of maritime accidents to vessel accident databases	**Delivered:** 14.06.2010
	Availability: Unavailable until 01.01.2011
Student: Martin Hassel	**Number of pages:** 41 pages

Abstract:

Underreporting of maritime accidents is a problem not only for authorities trying to improve maritime safety through legislation, but also to risk management companies and other entities using maritime casualty statistics in risk and accident analysis.

This study has collected and compared casualty data from 01.01.2005 to 31.12.2009, from IHS Fairplay and maritime authorities from: Norway, Sweden, Denmark, the United Kingdom, the United States of America, Canada, Greece and the Netherlands. Comparison of data to find common records has been done manually, while estimation of the true number of occurred accidents has been done by using conditional probability given positive dependency between data sources, several variations of the capture-recapture method (Lincoln-Petersen, Chapman and Chao), calculation of the best case scenario, assuming perfect reporting, and finally scaling up a subset of casualty information, from Cefor's Nordic Marine Insurance Statistics Database.

Estimated upper limit reporting performance for the selected flag states range from 7% (Greece) to 74% (Canada), while the estimated coverage of IHS Fairplay ranges from 4% (USA) to 62% (Canada). This study has found that on average, across the board, the amount of accidents that go unreported make up roughly 50% of all occurred accidents. Even in a best case scenario, only a few flag states come close to perfect reporting (Sweden, 94%). Based on the author's assessment of the data quality from the flag state authorities, most emphasis should be placed on the results for Norway (41% of accidents are not reported), The United Kingdom (56% of accidents are not reported) and Canada (10% of accidents are not reported).

This study not only proves, by the fact that some casualty data are exclusively in only one database, that underreporting exists, but also that it is a major problem. The best flag state in this study, Canada, is missing roughly a quarter of all accidents occurring in their area of responsibility. Norway and The United Kingdom covers just over a third of all occurred accidents in their area of responsibility, despite these flag state authorities' high level of perceived competency and quality.

The large amount of underreporting uncovered by this study, indicates that all users of statistical data should assume a certain amount of underreporting, and adjust their analyses accordingly. Whether they use a correction factor, safety margin or rely more heavily on expert judgement, must be decided on a case by case basis, as inaccurate basic data most often will have a significant impact on the end result.

Keywords:

- Underreporting
- Marine Databases
- Risk Management

Advisor:

Bjørn Egil Asbjørnslett

NTNU

Det skapende universitet

Master Thesis

Underreporting of maritime accidents to vessel accident databases

• Department of Marine Technology • Norwegian University of Science and Technology •

Preface

This master thesis is the product of work carried out during the spring semester of 2010, and is the product of a comprehensive study of selected marine databases from several flag state authorities, and is written in its entirety by the author, Martin Hassel.

I have several years of naval experience from previous service in the Royal Norwegian Navy (RNoN), as a navigator, commanding officer and instructor. This paper represents the final work, of my Master of Science degree in Marine Engineering, at the Norwegian University of Science and Technology.

I wish to thank the following people for their contribution and help on this project:

My supervisor Prof. Dr. Bjørn Egil Asbjørnslett, who has coached me throughout the project, and provided valuable support and guidance.

Ms Astrid Seltmann, analyst/actuary at Cefor (Nordic Association of Marine Insurers), has provided me with a unique opportunity, allowing me to compare data from NoMIS (Nordic Marine Insurance Statistics) with other sources.

Prof. Dr. Stein Haugen and Mr Thomas Eriksen from Safetec Nordic have provided valuable insight into the fundamentals of risk management work, and how basic data are used in risk assessments.

Prof. Dr. Dankmar Böhning was kind enough to provide feedback on the application, validity and accuracy of the statistical methods used in this paper.

All government officials who have been kind enough to respond to my data requests, representing the flag state authorities discussed in this paper have been of great support. They have politely helped me get the data I needed to perform this research, and endured my meticulous attention to details.

Trondheim, 14th June 2010.

Martin Hassel

Table of Content

Appendix list

List of Illustrations

List of Tables

Glossary and Abbreviations

AIS	-	Automatic Identification System
Cefor	-	Nordic Association of Marine Insurers (previously Central Union of Marine Underwriters)
DIS	-	Danish International Ship Register
DMA	-	Danish Maritime Authority
DS	-	Danish Ship Register
EMCIP	-	European Marine Casualty Information Platform
ESRA	-	European Safety and Reliability Association
ESREL	-	European Conference on Safety and Reliability
FOIA	-	Freedom Of Information Act
FSA	-	Formal Safety Assessment
GISIS	-	Global Integrated Shipping Information System
GT	-	Gross Tonnes
HCG	-	Hellenic Coast Guard
HSE	-	Health - Safety - Environment
H&M	-	Hull & Machinery
IHS	-	Information Handling Services
IMO	-	International Maritime Organization
IUMI	-	International Union of Marine Insurance
IVW	-	Inspectie Verkeer en Waterstaat (Dutch Transport and Water Management Inspectorate)
JTSB	-	Japanese Transport Safety Board
LISCR	-	Liberian International Ship & Corporate Registry
LMIU	-	Lloyd's Maritime Information Unit
MAIB	-	Marine Accident Investigation Branch (UK)
MISLE	-	Marine Information for Safety and Law Enforcement
MSC	-	Maritime Safety Committee
NIS	-	Norwegian International Ship Register
NMD	-	Norwegian Maritime Directorate
NoMIS	-	Nordic Marine Insurance Statistics
NOR	-	Norwegian Ship Register
NTNU	-	Norwegian University of Science and Technology
NTSB	-	National Transportation Safety Board (US)
P&I	-	Protection & Indemnity
STA	-	Swedish Transport Agency
SQL	-	Structured Query Language
TSB-(C)	-	Transportation Safety Board (Canada)
UNCTAD	-	United Nations Conference on Trade and Development
USCG	-	United States Coast Guard

Introduction

My master thesis is a continuation of work started in my project thesis, which was converted into an academic article and submitted to the European Safety and Reliability Conference (ESREL), after completion. It will be presented at ESREL 2010, in September. Due to the positive response, and the project thesis' acceptance to ESREL, the master thesis has been written as an academic article from its conception, with future journal publication in mind. It is written according to the IMRAD principle, and adheres to the format and standards of academic articles published by Elsevier Science.

I believe an adaptation of this paper could be submitted to one of the following scientific journals (all published by Elsevier Science): "Accident Analysis & Prevention", "Marine Policy", "Ocean Engineering", "Reliability Engineering & System Safety" (Level 2), "Safety Science" (Level 2) or "Transportation Research - Part A, D, E or F" (All level 2). Another option is the "Journal of Marine Science and Technology" (Level 2), published by Springer.

The main content of this master thesis is written in two-column format, with the style and appearance of a published academic article. Some additional raw materials have been provided at the end, to allow further work to be undertaken by other students at NTNU, building on the research described in this thesis.

The work presented in this paper has been performed as the Master Thesis of the author's Master of Science degree at the Norwegian University of Science and Technology, Department of Marine Technology. The opinions expressed in this article are solely those of the author, and does not represent the views of the university, its academic staff or collaborative partners.

UNDERREPORTING OF MARITIME ACCIDENTS TO VESSEL ACCIDENT DATABASES

Martin Hassel[*]

Norwegian University of Science and Technology, Institute of Marine Technology, Trondheim, Norway.

Abstract

Underreporting of maritime accidents is a problem not only for authorities trying to improve maritime safety through legislation, but also to risk management companies and other entities using maritime casualty statistics in risk and accident analysis.

This study has collected and compared casualty data from 01.01.2005 to 31.12.2009, from IHS Fairplay and maritime authorities from Norway, Sweden, Denmark, the United Kingdom, the United States of America, Canada, Greece and the Netherlands. Comparison of data to find common records has been done manually, while estimation of the true number of occurred accidents has been done by using conditional probability given positive dependency between data sources, several variations of the capture-recapture method (Lincoln-Petersen, Chapman and Chao), calculation of the best case scenario, assuming perfect reporting, and finally scaling up a subset of casualty information from Cefor's Nordic Marine Insurance Statistics Database.

Estimated upper limit reporting performance for the selected flag states range from 7% (Greece) to 74% (Canada), while the estimated coverage of IHS Fairplay ranges from 4% (USA) to 62% (Canada). This study has found that on average, across the board, the amount of accidents that go unreported make up roughly 50% of all occurred accidents. Even in a best case scenario, only a few flag states come close to perfect reporting (Sweden, 94%). Based on the author's assessment of the data quality from the flag state authorities, most emphasis should be placed on the results for Norway (41% of accidents are not reported), The United Kingdom (56% of accidents are not reported) and Canada (10% of accidents are not reported).

This study not only proves, by the fact that some casualty data are exclusively in only one database, that underreporting exists, but also that it is a major problem. The best flag state in this study, Canada, is missing roughly a quarter of all accidents occurring in their area of responsibility. Norway and The United Kingdom covers just over a third of all occurred accidents in their area of responsibility, despite these flag state authorities' high level of perceived competency and quality.

The large amount of underreporting uncovered by this study, indicates that all users of statistical data should assume a certain amount of underreporting, and adjust their analyses accordingly. Whether they use a correction factor, safety margin or rely more heavily on expert judgement, must be decided on a case by case basis, as inaccurate basic data most often will have a significant impact on the end result.

Keywords: Maritime accidents, risk analysis, underreporting, accident statistics, capture-recapture.

---

[*] Corresponding author. Telephone: (+47) 92011745, E-mail: martin.hassel@gmail.com

1. Introduction

It is commonly believed that a certain amount of maritime accidents go unreported, and this is confirmed by firsthand reports from ship operators, as described by Devanney (2009). This may be due to a variety of causes, ranging from oversight to deliberate withholding of information. Local reporting procedures are sometimes not known to the crew or ship owner, or company policy may be adverse to incident reporting, fearing for the company's reputation or ship delays.

Underreporting of maritime accidents is not only a problem for the authorities who are trying to improve maritime safety through legislation, but also for the risk management companies and other entities using maritime casualty statistics in risk and accident analysis. Frequency of casualties from statistical data provide an overall view of the levels of safety involved in the shipping industry. Quantification of the real safety levels for different ship types and flags, as well as the main failure modes are made possible by good statistical data.

Several commercial actors sell maritime transportation data to a wide range of businesses within the maritime sector. Some provide input to in-house databases or data management tools, while others sell data packets on request, or subscription services to online resources. The common denominator for all these services is that they give the impression to the customer that they are presenting a perfect solution, or accurate data in response to a request. In reality, all databases on maritime casualties are riddled with inaccuracies or missing data. A previous study performed by Hassel and Hole (2009) showed that the level of relevant and complete data in databases such as Lloyd's Register Fairplay Sea-Web (Now IHS Fairplay Sea-Web), the Norwegian Maritime Directorate's accident database and the IMO's GISIS casualty database are far from optimal, and in some cases surprisingly inadequate.

Most researchers and commercial businesses simply accept available statistics as facts, and do not concern themselves with possible inaccuracies or incomplete data sets (Soares and Teixeira, 2001).

Risk analyses are often based on statistical data of past events, relying on recorded data of ship accidents. Improvements in maritime safety are dependent on accurate and comprehensive statistical data. Even studies comparing the risks in the maritime transportation sector with other modes of transportation are vulnerable to inaccurate data and underreporting. The lack of precise information leads to the inability to allocate appropriate resources, the inability to initiate and prioritize targeted interventions, and the inability to evaluate the effectiveness of those interventions.

In order to carry out qualitative or quantitative safety analysis, it is essential to obtain reliability data. Admittedly, a qualitative risk analysis requires less detailed statistical failure data, compared to quantitative risk assessment. However, a certain amount of reliability data is considered necessary in either case, in order to determine the probability of occurrence and the extent of the consequences of unwanted events.

Data availability may also determine the choice of risk analyses methods (qualitative or quantitative) most suitable for incorporation in a Formal Safety Assessment (FSA). Wang and Foinikis (2001) state that accident statistics on a generic vessel type can be obtained from the following sources:

1. Field experience (historical data) including:
   - 1.1. Data collection programs by government agencies.
   - 1.2. Data collection programs by classification societies.
   - 1.3. Data collection programs by insurance companies and P&I Clubs.
   - 1.4. Statistics maintained by private shipping companies.
2. Agreed judgmental estimates of experts.

Wang and Foinikis further recommends great attention to the data resources, as the various databases have variable basis for data analysis. This is due to the fact that different bodies look at safety issues from different perspectives, facilitating own interests. Regarding the problem of data inaccuracy, it generally adds uncertainty to risk assessments and reduces confidence in the results. This can only be overcome by expert judgement.

Simply choosing the right data source for a safety study can be a difficult task, as there is no formal ranking or content quality benchmarking when it comes to maritime accident databases. In comparison, there are many national and international well-established databases for road, rail and aviation accidents. Land and air safety has long held the focus of public attention, and seen continuous improvements in the last decades. Statistics of land and air accidents are believed to have less underreporting, and a generally higher level of quality and accuracy. This has been achieved through numerous studies into data quality and underreporting, whilst the maritime sector has seen very few.

Several studies on railway safety have focused on precise reporting and the quality of input data (Johnson, 2002; Ryan et al., 2010).

Similarly, the aviation industry has studied the importance of proper reporting procedures, and the effects of inaccurate statistical data and accident information (Galea et al., 2006; Rose, 2006).

The majority of studies into accident data quality has been performed for road transportation, and underreporting has consistently been uncovered between police and hospital records with regards to accident fatalities and injuries, over a wide period of time and geographical locations (Alsop and Langley, 2001; Amoros et al., 2006; Harris, 1990; Sciortino et al., 2005; Yamamoto, 2008).

Nielsen and Roberts (1999) uncover widespread underreporting of human injuries and deaths, in their research on fatalities among seafarers, but the underreporting of maritime (ship) accidents has not been carefully studied. However, some research has been performed in recent years. According to findings by Thomas and Skjong (2009), only about 30% of occurred fire and explosion accidents in chemical tankers are reported to Lloyd's Register Fairplay (LRF) or the Norwegian Maritime Directorate (NMD). Another study by Psarros et al. (2010) investigates underreporting of maritime accidents for Norwegian flagged tankers between 1997-2007, by comparing data from the LRF and NMD databases. Its findings are that the upper bound reporting performance is 41% for the NMD and 30% for LRF. The upper bound reporting performance represents the maximum possible database coverage.

These preliminary findings encourage further research of a wider set of data sources and flag states. It also invites researchers to attempt to quantify the reporting performance, uncover the causality of underreporting and find benchmarks for maritime casualty databases and other suppliers of maritime data. The results from this field of research will lead to a better understanding of how government agencies may improve legislation, find better incentives for how ship crews and operators can be motivated to report more diligently, and how casualty databases may improve their causality documentation.

The study of underreporting requires a set of well defined parameters of what constitutes underreporting. This paper looks at eight different flag states, each with their own set of guidelines and regulations of what to register. The most basic method to define underreporting is if one database has knowledge of accidents which are not known to another database when the accidents are within the scope and area of responsibility of both databases. Another way to look at underreporting, is if a database only requires the registration of a certain segment of accidents, or do not conform to international standards or norms. A third perspective is what a risk management professional would like to see registered, regardless of current standards or requirements. This paper has defined a set of accident parameters that all involved databases should register: The vessel in question must be a commercial vessel larger than 100 GT having an IMO identification number, the accident must happen while the ship is out of berth and the accident type must be as described in section 3.4.

In Section 2, a description of the selected data sources used in this study is provided. Section 3 describes the methodology used to estimate the amount of underreporting, while the results are given in Section 4. The results are analysed in Section 5, followed by further discussion and conclusions in Section 6.

2. Data

This study utilizes general casualty records from IHS Fairplay Sea-Web™ (Lloyd's Register of Ships Online), the Norwegian Maritime Directorate (NMD), the Swedish Transport Agency (STA), the Danish Maritime Authority (DMA), the Marine Accident Investigation Branch (MAIB), the United States Coast Guard (USCG), the Transportation Safety Board of Canada (TSB-C), the Hellenic Coast Guard (HCG), the Dutch Transport and Water Management Inspectorate (IVW) and Cefor's Nordic Marine Insurance Statistics Database (NoMIS), covering the period from 01.01.2005 to 31.12.2009. The registries of Antigua & Barbuda, Bahamas, Malta, New Zealand, Panama, Russia and St.Vincent & the Grenadines were also supposed to be part of this study, but the various flag state authorities did not reply to any of the requests for information from the author, and was thus not included. A short description of the selected databases / data sources are given in the following subsections.

2.1. IHS Fairplay - Sea-Web™

IHS Fairplay (formerly Lloyd's Register Fairplay) provides comprehensive details on the current world merchant fleet of 100 GT and above, accessible through their online subscription service, Sea-Web™. Accident data is available through a Casualty Module add-on. It is a very large and comprehensive database, spanning virtually every merchant ship in existence, even ships that have been scrapped, with about 110.000 reported marine incidents in total (As of March 2010).

Sea-Web could in popular terms be called the "Encyclopedia Britannica" of marine databases. It covers the entire world fleet, current and past, and is an excellent maritime tool. It is the oldest, largest and most comprehensive of all vessel databases. It has very good coverage of the connections between accidents, vessel and owner history and ship particulars, but does not have a satisfactory detail level of accidents. A previous study by Hassel and Hole (2009) uncovers inadequacies in key areas such as accident causes and consequences. That being said, it has all the necessary data fields relevant to this study, and is better structured than many flag state authorities' records.

2.1.1. New ownership

The old Lloyd's conglomerate is one of the oldest data banks available, and serves as the basis for most studies of global statistics. Currently, the two major sources of casualty information, namely IHS Fairplay Sea-Web and Seasearcher LMIU are no longer owned by the Lloyd's Register group, but continue to exist with apparent identical functionality under new ownership. However, to most people, it is hard to no longer associate the trademark name Lloyd's to these data sources. Seasearcher has thus kept the name Lloyd's MIU as a sub title, but will be rebranding the service in the near future, under the new name Lloyd's List Intelligence. Lloyd's List produces hundreds of supplements and special reports every year targeted at specific markets in the maritime sector and related industries, and is owned by Informa plc. Fairplay has been a wholly owned division of IHS Inc. since June 2009.

2.2. The Norwegian Maritime Directorate

The NMD's database covers accidents on all vessels registered under Norwegian flag (NOR/NIS registry) worldwide, and all foreign flagged vessels sailing in Norwegian waters. Access to the database requires access to the NMD's intranet, but a small excerpt is available on the NMD's website.

The database is well structured, suited for risk management work, and includes workplace accidents as well as near misses for both personnel and ships. The database registers a wide range of data fields relevant to maritime accidents, such as accident type, location of accident, date of accident, vessel specifics and flag, to name a few.

2.3. The Swedish Transport Agency

The STA's Maritime Department records accidents for all ships sailing in Swedish waters, and all Swedish flagged vessels worldwide. They also analyse selected accidents and near-misses. In general the structure and scope of the STA is very similar to the other flag state authorities, in particular the NMD, DMA and MAIB.

2.4. The Danish Maritime Authority

The DMA records accidents for all ships registered in the Danish Register of Shipping or the Danish International Register of Shipping (DIS), as well as all accident involving foreign ships in Danish waters. They also register accidents involving special Danish interests, regardless of location or vessel registration.

2.5. The Marine Accident Investigation Branch

The MAIB examines and investigates all types of marine accidents to or onboard UK vessels worldwide, and other vessels in UK territorial waters.

The MAIB is primarily concerned with investigating accidents to determine circumstances and causes, aiming to improving the safety of life at sea and the avoidance of future accidents. Consequently, as all reported accidents are logged and registered, the MAIB maintains a computerised database of reportable marine accidents which have occurred since 1991. Besides providing an accessible source of information, the database can be analysed to identify accident trends.

Accident statistics are set out in the MAIB's annual report, publicly available on their website.

2.6. The United States Coast Guard

The Coast Guard's Office of Marine Safety and Environmental Protection (USCG-MSEP) maintains the Marine Information for Safety and Law Enforcement (MISLE) information system. One module of the MISLE, the Marine Casualty and Pollution Database, contains data describing all safety-related investigations involving commercial vessels operating in United States' territorial waters or U.S.-registered commercial vessels operating elsewhere in the world.

Investigations are initiated for events exceeding specific thresholds of bodily harm, pollution, material damage or other specifications.

Computerised data, entered by Coast Guard staff, are reviewed by front-line supervisors, and then transmitted to the USCG-MSEP for inclusion in the MISLE safety module. Casualty and pollution data are available, in one form or another, from 1975 to the present. The MISLE was implemented in December 2001. Previously, data describing investigations of accidents or incidents involving commercial vessels were stored in the Marine Investigations Module (MINMOD) of the Marine Safety Information System (MSIS). Commercial vessel casualties, injuries, and deaths prior to 1992 can be found in the CASMAIN database. In the year 2000, approximately 17,000 casualty and 8,000 pollution events were added to the MISLE (then MSIS). The Coast Guard does not currently publish annual summaries of commercial marine casualties or vessel losses, but all recorded data can be made available through a Freedom Of Information Act (FOIA) request (NTSB, 2002).

2.7. The Transportation Safety Board of Canada

The TSB-C is an independent agency, striving to advance transportation safety through the investigation of occurrences in the marine, pipeline, rail and air modes of transportation. The Head Office is located in Gatineau, Quebec; however, most investigation staff are located in various regional and field offices across Canada, where they are better able to respond quickly to transportation occurrences anywhere in the country. The TSB-C was somewhat slow to return data, but compensated with a remarkable level of service once communication was established. Their are of responsibility is similar to that of the NMD, STA and MAIB.

2.8. The Hellenic Coast Guard

The Marine Casualty department is a division of the Hellenic Coast Guard, and co-located with the national search and rescue center in Piraeus.

Using the provided Greek data, it is important to note that the prerequisites for a formal investigation by the Hellenic Coast Guard are different than in the International Maritime Organization's (IMO) Code for Casualty Investigation, and consists of any of the following criteria:

- Total or partial loss of a ship/floating structure.
- Ship/floating structure is taken over by insurers.
- Permanent or temporary abandonment of ship by the crew.
- Cargo loss or failure (more than 25%).
- Prolonged loss of ship command due to serious failure.
- Loss of life or serious injury to a crew member or passenger.

As Greece has a different definition of marine accidents and incidents, there is an inherent level of underreporting in the Greek data, according to IMO definitions (Hellenic Coast Guard, 2010; Tzannatos and Kokotos, 2009).

2.9. The Dutch Transport and Water Management Inspectorate

The "Inspectie Verkeer en Waterstaat" (IVW) is the Dutch maritime authority, and responsible for performing preliminary investigations of marine accidents and injuries to seafarers on Dutch seagoing vessels, to ensure the continuos safety at sea for all Dutch ships. The IVW covers marine accidents such as: collisions, labour related accidents on board, groundings, fire, capsizing, et cetera.

2.10. Cefor's Nordic Marine Insurance Statistics Database

The Nordic Association of Marine Insurers (Cefor) is the trade association for marine insurance companies in the Nordic region and has currently 15 members. To attain a larger data set of insurance claims, Cefor members support the NoMIS database, located at and administered by Cefor's administration,in order to improve evaluation of market trends and enhance the validity of statistical analyses made on the combined data. Cefor's joint database project (NoMIS) began in 1985, and was converted into an electronic and modern database around 1999. Now the members submit data electronically to NoMIS monthly, and may request combined data at any time. The reporting contains only about 10 data fields (mainly regarding the specific insurance claim/incident), but Cefor connects the NoMIS (accidents/claims) data with vessel details from Seasearcher LMIU and IHS Fairplay. Cefor publishes statistical data showing market trends relating to insurance claims and insurance premiums annually on their website. The NoMIS database is not a publicly accessible database, and cannot be used for individual maritime accident analysis, as no identifying information can be released. Only the Cefor administration has access to the full database, while members and other data recipients only receive anonymized information, making it purely usable for statistical purposes. It covers about 21% of the world fleet above 300 GT in terms of number of vessels and 40% in terms of tonnage, with a total of 40,613 claims (data points) registered in the period from 1985 to 2008. One of Cefor's objectives is to *"make available to its members appropriate data-based statistics to support the activities of the individual members and the* *general objectives of the association"* (Cefor, 2010).

2.11. The Japanese Transport Safety Board

The JTSB was established when the Japanese Marine Accident Inquiry Agency (JMAIA) combined the previously separate departments for air, land and sea, on 1st October 2008. The integration of the three departments was performed to enhance and combine the investigative capabilities, in an effort to improve accident analysis and prevention.

The JTSB was somewhat slow to respond, and the returned data were missing important details, such as IMO number for each casualty entry. Through correspondence with the JTSB it was established that the JTSB did not use IMO numbers in domestic casualty registers. The received data could not be processed with sufficient accuracy without proper identification of each ship, and the timeframe for this paper did not allow for further resources and time to be spent trying to find a solution to this problem. Japanese data have subsequently been omitted from this paper.

2.12. Liberian International Ship & Corporate Registry

Liberia was the biggest flag state in the world, both by deadweight and number of ships, as of 1st January 2009 (UNCTAD, 2009). The Liberian registry (LISCR) is administered by a U.S. owned and operated company that provides the day-to-day management of the registry. LISCR was contacted at the beginning of this study, and did return some data. However, the data were generic annual self assessment forms, and not a ship specific list of casualties. Further correspondence did not return the desired results, and no usable data were made available, preventing Liberia to be part of this study.

3. Methodology

The probability that a randomly selected accident is registered in a given database, is defined as the number of registered accidents in that particular database, divided by the actual number of occurred accidents.

The amount of (under)reporting must thus be ascertained for each individual database, as it is the difference between the number of registered accidents and the actual number of occurred accidents. For each database, we are only interested in finding the true level of accidents of the same type as those registered in Sea-Web. Only accidents within the scope of both databases in question are relevant to determine the true level of reporting, and Sea-Web is nearly always the database with the least amount of registered accidents. This is due to the fact that Sea-Web only registers vessel accidents, while most flag state authorities also register workplace accidents, near misses or incidents while at berth. This paper will only investigate accidents involving a commercial vessel larger than 100 GT having an IMO identification number. The accident must also occur while the ship is out of berth and the accident type must be as described in section 3.4. The following sections describe how the collection of data was performed.

3.1. IHS Fairplay - Sea-Web™

With a normal user account to Sea-Web (including the Casualty Module), it is possible to run simple SQL (Structured Query Language) queries from the website. For this study, the following query was performed:

SELECT ALL (ships/casualties) WHERE GT≥100, AND CASUALTY DATE is between 01.01.2005 and 31.12.2009 AND FLAG="selected flag states".

Sea-Web does not allow the SQL operator "OR" to be used, so a separate query was performed for each flag state covered by this study. It should be noted that during the course of this study the author uncovered a faulty/missing functionality in Sea-

Web, as the query for "FLAG=Norway" only returned NOR (Norwegian Ship Register) vessels, omitting vessels registered in NIS (Norwegian International Ship register). Similarly the query "FLAG=Denmark" only returned ships from the Danish Register of Shipping. This was pointed out to IHS Fairplay Sea-Web on 2nd March 2010, and the proper functionality was implemented after a few days. It is now possible to run queries like: "FLAG=Norway" or "FLAG=Norway (NIS)", and similarly "FLAG=Denmark" or "FLAG=Denmark (DIS)". IHS Fairplay seem to have made a comprehensive quality check and update to their casualty search attributes as a response to being made aware of the earlier shortcoming. Other countries with multiple ship registries are now also available as individual search options.

With accident locations only given in Marsden Grid References or zones as defined by Lloyd's World Casualty Statistics, it is not possible to make a query for all ships with a certain flag and/or vessels subject to an accident in the same flag state's national waters. This means it is not possible to achieve a full comparison between Sea-Web and flag state data, but simply casualties of a specific flag. This is unfortunate, as several flag states have pointed out that they have less control over foreign shipping in their national waters, subsequently expecting underreporting to be a bigger issue among this segment (IHS Fairplay, 2010; NMD, 2010).

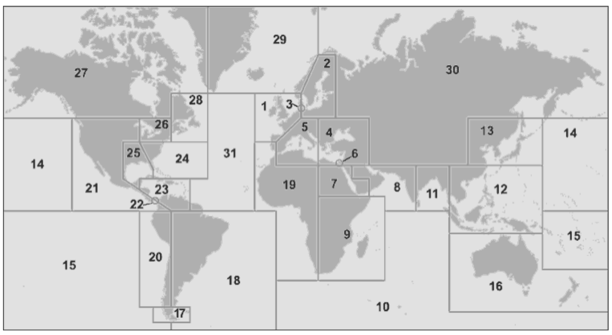


Figure 1.
Sea-Web sea zones.

The use of sea zones to categorize position will be further discussed later. This is a point of potential improvement to both database search ability and underreporting estimation. Sea-Web's current system is shown in Figure 1.

3.2. Flag State Authorities

Government officials from the relevant departments were able to run the aforementioned query in their own databases, and provide flag state specific data. The returned data were then filtered further, to remove duplicate entries and entries outside the scope of this study. To ensure that no true records were overlooked, the initial data request was somewhat larger than strictly necessary.

Duplicate entries were a major issue with some of the flag states, as they made separate entries for each step of an accident scenario (for example, the STA would make one entry for an engine failure, another for the subsequent grounding, and a third for injuries to personnel due to the grounding).

Greece presented another problem, as they had suffered data loss, and due to limited manning and resources, did not have good records of accident statistics for ships sailing under Greek flag in the time period to be examined.

3.3. Cefor's Nordic Marine Insurance Statistics Database

Direct access to the NoMIS database is not possible in the context of the legal insurance framework, but publicly available statistics are regularly posted on the organization's website. Even Cefor members are only privy to anonymized statistical data. Ideally, unlimited access to the database could significantly improve the estimation models, and provide a whole new data source. This was not possible, but by requesting only anonymized data and with the association's approval, Cefor's chief analyst and actuary was able to provide valuable data Including the number of vessels currently in the database, sorted by flag, and the number of claims received each year, also sorted by flag, for the selected flags covered by this study (Cefor, 2010).

3.4. Further filtering of Flag State Data

The studied flag states have different thresholds and framework governing the scope of their databases, making it necessary to manually process the returned data to filter out any irrelevant entries. Only entries conforming to Lloyd's World Casualty Statistics' definition of the following accident categories were included in the final analysis:

(i) Foundered
(ii) Fire/Explosion
(iii) Collision
(iv) Wrecked/stranded
(v) Contact
(vi) Hull/Machinery

(World Casualty Statistics, 2008)

Additionally, all entries missing vital data, such as IMO number, type of accident, vessel size and type, or entries containing contradictory information were all rejected. Many entries lacked sufficient data to present a complete picture, and was eventually rejected. All accidents categorised as "workplace accidents", "personnel injuries", "pollution/spillage" or "near-miss" were also rejected. This filtering accounts for the sometimes large difference between the amount of entries/accidents received from flag state authorities and the values used in estimations, equations and tables.

3.5. Comparative analysis to find common entries

Once all the data from the flag states and Sea-Web was collected and properly filtered, a 1-to-1 comparison between Sea-Web data and individual flag state data was performed in order to find the amount of common entries. This had to be done manually, as the data were from several different sources and did not have identical formats or formats suitable for programmed comparison. Due to variable data quality, and irregularities between registration date and accident date, entries were deemed to match even if (accident) date differed by ±14 days, if all other data fields matched.

3.6. Estimation of true level of accidents

There are several ways to estimate the true amount of occurred accidents. This study will use four different methods: (1) conditional probability, (2) capture-recapture methods, (3) best case scenario and (4) upscaling of subset data.

Henceforth, the following notation is used:

- $P(FS_n)$ is the probability that an occurred accident is reported to the flag state (database).
- $P(IHSF_n)$ is the probability that an occurred accident is reported to IHS Fairplay.
- $P(FS_n \cap IHSF_n)$ is the probability that an occurred accident is reported to both the flag state and IHS Fairplay.
- $P(IHSF_n \mid FS_n)$ is the conditional probability that an occurred accident is reported to IHS Fairplay given that it has been reported to the flag state.
- FS_n is the number of recorded accidents in a specific flag state's accident database.
- $IHSF_n$ is the number of recorded accidents for a specific flag, in IHS Fairplay.
- C_n is the number of accidents reported to both the respective flag state and IHS Fairplay.
- a_n is the number of accidents only reported to the flag state authorities.
- b_n is the number of accidents only reported to IHS Fairplay.
- x_{mn} is the number of unreported accidents described in section 3.4, (according to method "m", applied to flag state "n").
- N_{mn} is the (estimated) total number of occurred accidents described in section 3.4, (according to method "m", applied to flag state "n").
- R_n is the accident ratio found in NoMIS
- TF_n is the total fleet size in number of ships
- ICC_n is the number of claims submitted to Cefor
- ICF_n is the number of ships covered by Cefor
- Z_n is the factor describing the relationship between $P(IHSF_n \mid FS_n)$ and $P(IHSF_n)$

- Index “m” indicates the method used to estimate the unknown variables “x” and “N”. Where “m” has one of the values shown in table 1.

Table 1. Description of index “m”

m	Method
1	Bayesian conditional probability estimation
2.1	Lincoln-Petersen capture-recapture estimation
2.2	Chapman capture-recapture estimation
2.3	Chao capture-recapture estimation
3	Best case scenario
4	Scaling up of Cefor accident ratio

- Index “n” indicates a specific flag state. Where “n” has one of the values shown in table 2.

Table 2. Description of index “n”

n	Flag state
1	Norway
2	Sweden
3	Denmark
4	United Kingdom
5	U.S.A.
6	Canada
7	Greece
8	The Netherlands

The other flag states mentioned in this paper did not provide (enough satisfactory) data to be included, thus stopping index “n” at numeric value 8.

3.6.1. Bayesian conditional probability

It is possible to estimate the true amount of accidents (N_{mn}) for a specific flag, by using available data on registered accidents from IHS Fairplay ($IHSF_n$) and the selected flag state (FS_n). This method is also used by Psarros et al. (2010) in their previously mentioned study on the underreporting of Norwegian tanker accidents between 1997 and 2007.

Estimated from available statistical data, the following thus applies:

$$P(FS_n) = \frac{FS_n}{N_{mn}} \tag{1a}$$

$$P(IHSF_n) = \frac{IHSF_n}{N_{mn}} \tag{1b}$$

$$P(FS_n \cap IHSF_n) = \frac{C_n}{N_{mn}} \tag{1c}$$

Using Bayes’ theorem of conditional probability, it is further given that:

$$P(FS_n \cap IHSF_n) = P(IHSF_n \mid FS_n) * P(FS_n) \tag{2}$$

It is a fair assumption to presume that the likelihood of an accident being registered with IHSF is larger if the same accident is already registered in the flag state database (and vice versa).

$$P(IHSF_n \mid FS_n) \geq P(IHSF_n) \tag{3}$$

This assumption is based on the growing coverage by media worldwide, and the likelihood that there may be some common sources of information in the maritime cluster. There also exists a certain level of interaction between flag state database officials, and IHSF. Another study, by Hole (2010) looks more closely at reporting chains, and the validity of this assumption.

Combining equations (2) and (3) produces the following statement:

$$P(FS_n \cap IHSF_n) = \\ P(IHSF_n \mid FS_n) * P(FS_n) \geq P(IHSF_n) * P(FS_n) \tag{4}$$

By inserting equations (1a)-(1c) into (4), the statement is simplified:

$$\frac{C_n}{N_{mn}} \geq \frac{FS_n}{N_{mn}} * \frac{IHSF_n}{N_{mn}} \tag{5}$$

Ultimately, a lower limit of the true number of occurred accidents are estimated by:

$$N_{1,n} \geq \frac{FS_n * IHSF_n}{C_n} \tag{6}$$

In turn, this combined with equations (1a) and (1b) produces an upper limit of reporting performance for each flag state, and IHS Fairplay.

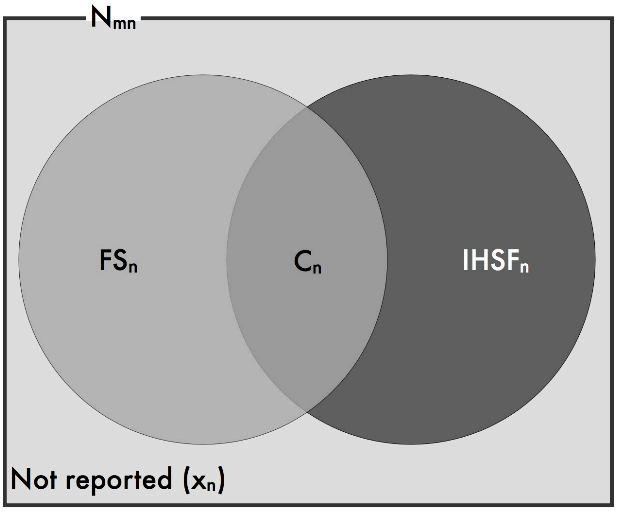


Figure 2.
Venn diagram modeling registered (FS_n, $IHSF_n$), common (C_n) and unknown (x_{mn}) accidents as subsets of the total number of occurred accidents (N_{mn}).

3.6.2. Capture-recapture method

This method originates from epidemiology and estimations of animal populations sizes. However, it is also a valid method of data comparison, and has amongst others been used by Aptel et al. (1999) and Trépanier et al. (2009) in a similar context, and further elaborated by Brittain and Böhning (2009). As described by Desenclos and Hubert (1994) and Hook and Regal (1995), the method requires four assumptions to be used correctly: (1) the data sources must be independent, (2) the population must be closed, (3) there can not be any false matched records and (4) no true matches can be missed. The first assumption contradicts the presumption made in the previous method, but one could just as easily argue the data sets are independent, as there is no formal link between IHS Fairplay and most flag state authorities. It can thus be argued that to assume independence between data sets is not entirely wrong. However, several versions of the capture-recapture method exists, with different approaches to estimating the unknown population. The alternatives will be discussed and evaluated based on accuracy and their sensitivity to the assumption of independent data sources. The population is closed as no real recapture takes place. The "capture/recapture" is done instantaneously. This fulfills the second assumption. The third and fourth assumptions are subject to the accuracy of the manual comparison performed by the author, who have performed several iterations to ensure good precision.

Table 3. The two source model showing registered (a_n, b_n), common (C_n), unknown (x_{mn}) and total (FS_n, $IHSF_n$) accidents

Accident registered		In Flag State		
		Yes	No	
In IHS Fairplay Sea-Web	No	a_n	x_{mn}	
	Yes	C_n	b_n	$IHSF_n$
		FS_n		

($IHSF_n = C_n + b_n$ while $FS_n = C_n + a_n$)

The capture-recapture method for two sources, shown in table 3, has several ways to estimate the unknown quantity, and a study by Brittain and Böhning (2009) describes the most common alternatives. For instance the Lincoln-Petersen estimation, which is very similar to that of equation (6) from the last section. The Lincoln-Petersen estimation states that:

$$x_{2.1,n} = \frac{a_n * b_n}{C_n} \quad (7)$$

$$N_{2.1,n} = \frac{FS_n * IHSF_n}{C_n} \quad (8)$$

However, in the case of no matching records, this estimation cannot compute, and another estimation technique is needed.

The Chapman estimator is given by the following equations:

$$x_{2.2,n} = \frac{a_n * b_n}{C_n + 1} \quad (9)$$

$$N_{2.2,n} = \left[\frac{(FS_n+1)(IHSF_n+1)}{C_n+1}\right] - 1 \quad (10)$$

$$Var(N_{2.2,n}) = \frac{(FS_n + 1)(IHSF_n + 1) * a_n * b_n}{(C_n + 1)^2(C_n + 2)} \quad (11)$$

$$95\%CI = N_{2.2,n} \pm 1,96\sqrt{Var(N_{2.2,n})} \quad (12)$$

The Chapman estimator is a commonly used method in epidemiology and statistics, but is vulnerable to dependency between data sources.

Using a 95% confidence interval (CI), it is possible to say that the true value is within the returned interval range, with 95% certainty. A small interval means the method is more accurate than one with a large confidence interval.

A third estimation technique is Chao's lower bound estimate, using the following equations:

$$x_{2.3,n} = \frac{(a_n + b_n)^2}{4C_n} \quad (13)$$

$$N_{2.3,n} = a_n + b_n + C_n + \frac{(a_n + b_n)^2}{4C_n} \quad (14)$$

$$Var(N_{2.3,n}) = x_{2.3,n}\left(\frac{a_n+b_n}{2C_n} + 1\right)^2 \quad (15)$$

$$95\%CI = N_{2.3,n} \pm 1,96\sqrt{Var(N_{2.3,n})} \quad (16)$$

It is fairly easy to see that the Lincoln-Petersen and Chapman estimators will yield results similar to the method using conditional probability. Three converging methods would normally indicate good accuracy, but according to Brittain and Böhning (2009), Chao's lower limit estimation performs better than Chapman's estimator, as it has a generally lower relative bias, despite Chapman's lower relative variance. Brittain and Böhning further suggests that Chao's estimator has a better confidence interval coverage than the more commonly used Chapman estimator, especially when the assumption of independence between sources is in doubt, which is the case in our situation.

3.6.3. Best case scenario

Assuming a perfect world where all occurred accidents are reported to at least one database, the number of occurred accidents may be found by adding the number of registered accidents in both IHSF and the flag state database, and subtracting the common entries.

$$N_{3,n} = (FS_n + IHSF_n) - C_n \quad (17)$$

Assuming perfect reporting, this model returns a specific value, although it would be more correct to use the following modification:

$$N_{3,n} \geq (FS_n + IHSF_n) - C_n \quad (18)$$

With the modification made in equation (18), the assumption of perfect reporting is no longer required, as the equation now provides an absolute lower limit of the total population, based on proven records.

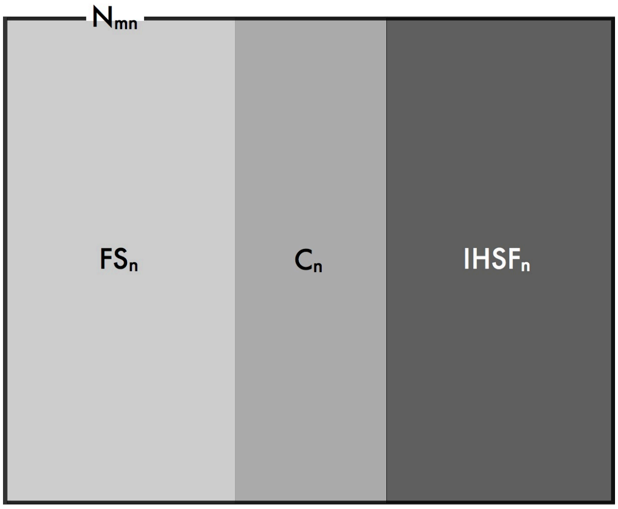


Figure 3.
Venn diagram modeling registered (FS_n, $IHSF_n$) and common (C_n) accidents as subsets of the total number of occurred accidents ($N_{3,n}$).

Calculating the upper limit reporting performance for flag states and IHS Fairplay, returns the best case scenario (absolute upper limit).

3.6.4. Comparative data set

Assuming all occurred accidents are reported to the vessel's insurance company, it is possible to estimate the relationship between P(IHSF|FS) and P(IHSF) better than simply "equal to" or "greater than". This can be achieved by finding an accident ratio (R_n) for the selected flag states from the insurance companies' database, and scale up the result to match the actual fleet size for each flag state. The (insurance company) accident ratio can be found by dividing the number of reported accidents (ICC_n) for a specific flag state with the total number of ships (ICF_n) from the same flag state, covered by the insurance companies' database.

$$R_n = \frac{ICC_n}{ICF_n} \tag{19}$$

Applying this accident ratio to the whole fleet of a particular flag state (TF_n) enables another estimation of the true number of occurred accidents:

$$N_{4,n} = R_n * TF_n \tag{20}$$

Assuming that the (insurance company) accident ratio is uniformly distributed, representative and scalable, gives $N_{4,n} \approx N_{mn}$. Using this in equations (1a)-(2) results in a factor (Z_n) describing the relationship between P(IHSF | FS) and P(IHSF):

$$P(IHSF_n \mid FS_n) = Z_n * P(IHSF_n) \tag{21}$$

Subsequently, the factor describing the relationship between P(IHSF | FS) and P(IHSF) is as follows:

$$Z_n = \frac{N_{4,n} * C_n}{IHSF_n * FS_n} \tag{22}$$

Using the value $N_{4,n}$ in equations (1a) and (1b) produces new approximations of reporting performance for each flag state, and IHS Fairplay.

4. Results

This chapter presents the refined data sets received from the various flag state authorities and IHS Fairplay Sea-Web. The data are presented in two-source tables showing how the combined amount of registered accidents are distributed between the flag state's database and IHS Fairplay Sea-Web. Data collected from IHS Fairplay Sea-Web are presented "as-is", with very little or no modification. This is due to the clear cut results returned from SQL-queries made directly in the database's web based graphical user interface (GUI). Whereas data returned from the various flag states had to undergo major refinement, quality control and filtering. Communication errors, language problems, database structure and other factors made it difficult to always get the information relevant to this study. As a result, the data requests put forward to the flag state authorities would usually ask for more data than were strictly necessary. This was done to ensure that no relevant data were missing, although it required some additional processing of the returned information.

The refined data and calculated estimations are presented in the following sub-sections, and illustrations may be found in appendices (a) and (b).

4.1. IHS Fairplay - Sea-Web™

The Sea-Web query for all casualties from 01.01.2005 to 31.12.2009, of all ships greater than 100 GT, for the selected flags, returned the results shown in table 4. This data were retrieved manually through the web-based interface available with a normal user account (Sea-Web, 2010).

Table 4.
Number of registered accidents in IHS Fairplay Sea-Web ($IHSF_n$), from 01.01.2005 to 31.12.2009

Flag state	$IHSF_n$
Norway (NOR/NIS)	529
Sweden	109
Denmark (DS/DIS)	189
United Kingdom	401
United States	632
Canada	608
Greece (only for 2009)	109
The Netherlands	304

Sea-Web registers the vessels' flag at the time of the casualty, avoiding missing records due to later flag changes. The same applies to all flag state databases in this study. All recorded vessel data is from the time of the accident. Later changes to flag or vessel are not reflected in the casualty statistics.

4.2. The Norwegian Maritime Directorate

The original data received contained all registered casualties since 01.01.2005, totaling 11873 entries. This included 9558 workplace accidents not relevant to this study, and other entries irrelevant to this comparative analysis. Removing all additional entries outside the scope of the Sea-Web query, and entries deemed irrelevant due to other reasons made the final comparable number of entries 596, which is the sum of the left content column in table 5 (NMD, 2010).

(Tables 5-12 are structured in the same manner as table 3, where the sum of the bottom row is the total amount of registered accidents in IHS Fairplay Sea-Web, while the sum of the left content column is the sum of the registered accidents in the flag state database. The number of common and exclusive entries are shown in tables 5-12.)

Table 5.
Registered accidents of Norwegian flagged vessels (NOR/NIS), from 01.01.2005 to 31.12.2009

Accident registered		In NMD	
		Yes	No
In Sea-Web	No	393	x_1
	Yes	203	326

4.3. The Swedish Transport Agency

The STA data were full of duplicate entries due to the reporting procedures of making separate entries for each major event in an accident scenario. The original data set contained 625 entries, but the relevant number of entries was reduced to 333 after data filtering. The distribution of the Swedish data are shown in table 6 (STA, 2010).

Table 6.
Registered accidents of Swedish flagged vessels, from 01.01.2005 to 31.12.2009

Accident registered		In STA	
		Yes	No
In Sea-Web	No	247	x_2
	Yes	86	17

4.4. The Danish Maritime Authority

The DMA database has a very similar structure and framework as the NMD, and the total number of casualty entries returned was 948, however only 220 conformed to the requirements set by this study. Table 7 has the numbers (DMA, 2010).

Table 7.
Registered accidents of Danish flagged vessels (DS/DIS), from 01.01.2005 to 31.12.2009

Accident registered		In DMA	
		Yes	No
In Sea-Web	No	175	x_3
	Yes	45	144

4.5. The Marine Accident Investigation Branch

The MAIB has a well structured database, and returned the most comprehensive data transcript of all the flag states covered by this study. Of the 5197 entries received, 1428 entries remained after data filtration. A lot of the discarded data were leisure crafts and non-IMO numbered fishing vessels. Table 8 shows the distribution of the collected data (MAIB, 2010).

Table 8.
Registered accidents of British flagged vessels, from 01.01.2005 to 31.12.2009

Accident registered		In MAIB	
		Yes	No
In Sea-Web	No	1 199	x_4
	Yes	229	172

4.6. The United States Coast Guard

The USCG returned a spreadsheet with formidable 24747 entries, of which only 2362 entries remained after irrelevant/deficient entries had been discarded. The initial data contained a lot of barges, vessels without IMO number and entries with void data fields. Interestingly, all accident types were collision/allision or grounding, no other categories were present. The USCG contact person was unable to explain why this was the case and did not know where to direct further inquiries regarding this issue. See table 9 for details (USCG, 2010).

Table 9.
Registered accidents of U.S.A. flagged vessels, from 01.01.2005 to 31.12.2009

Accident registered		In USCG	
		Yes	No
In Sea-Web	No	2 227	x_5
	Yes	135	497

4.7. The Transportation Safety Board of Canada

The TSB-C data contained 833 entries, of which 722 entries complied with the comparative requirements mentioned earlier. It should be noted that the TSB-C had the least amount of rejected entries, and a very well structured regulatory framework and data management system. See table 10 for details (TSB-C, 2010).

Table 10.
Registered accidents of Canadian flagged vessels, from 01.01.2005 to 31.12.2009

Accident registered		In TSB-C	
		Yes	No
In Sea-Web	No	268	x_6
	Yes	454	154

4.8. The Hellenic Coast Guard

Greek authorities were unable to provide satisfactory data, partly due to data loss and problems with manning and resources. Data for 2009 were given, but due to the difference between IMO standards and the national standards used by Greece, the data is of a significantly lesser quality compared to the other flag states. The relatively small selection is also more susceptible to inaccuracies. See table 11 for details (HCG, 2010).

Table 11.
Registered accidents of Greek flagged vessels, from 01.01.2009 to 31.12.2009

Accident registered		In HCG	
		Yes	No
In Sea-Web	No	12	x_7
	Yes	12	97

4.9. The Dutch Transport and Water Inspectorate

The IVW provided a list of 500 entries, where 342 entries remained after data filtration, the distribution is shown in table 12 (IVW, 2010).

Table 12.
Registered accidents of Dutch flagged vessels, from 01.01.2005 to 31.12.2009

Accident registered		In IVW	
		Yes	No
In Sea-Web	No	271	x_8
	Yes	71	233

4.10. Cefor's Marine Insurance Statistics Database

Data returned from Cefor is shown in table 13, and shows the total amount of registered claims and the 5-year average number of vessels for each flag state studied in this paper.

Table 13.
Number of registered vessels and claims in NoMIS, from 2005-2009

		Annual average
Norway	Vessels	1 170
	Claims/accidents	349
Sweden	Vessels	422
	Claims/accidents	185
Denmark	Vessels	181
	Claims/accidents	82
UK	Vessels	339
	Claims/accidents	110
USA	Vessels	156
	Claims/accidents	30
Canada	Vessels	41
	Claims/accidents	10
Greece	Vessels	406
	Claims/accidents	48
The Netherlands	Vessels	307
	Claims/accidents	116

Additional data from NoMIS can be found in table 16, as well as appendix (d).

The combined results from all the flag states, using the Lincoln-Petersen capture-recapture method is shown in table 14a. It shows how the refined data from the flag states and IHS Fairplay are used to find an estimated number of the true amount of occurred accidents. Using that value it is possible to calculate the flag states' reporting performance by dividing FS_n by $N_{2.1,n}$ or IHS Fairplay's reporting performance by dividing $IHSF_n$ by $N_{2.1,n}$.

Table 14a. Results using capture-recapture method (Lincoln-Petersen estimator)

Flag state	Number of registered accidents: Sea-Web	Number of registered accidents: Flag state	Common entries	Estimation of true number of accidents ($N_{2.1,n}$)	Estimated reporting performance: Sea-Web	Estimated reporting performance: Flag state
Norway (NOR/NIS)	529	596	203	1 553	34 %	38 %
Sweden	109	333	86	422	26 %	79 %
Denmark (DS/DIS)	189	220	45	924	20 %	24 %
United Kingdom	401	1 428	229	2 501	16 %	57 %
United States	632	2 362	135	11 058	6 %	21 %
Canada	608	722	454	967	63 %	75 %
Greece (2009 only)	109	24	12	218	50 %	11 %
The Netherlands	304	342	71	1 464	21 %	23 %

The Lincoln-Petersen estimation returns the same numerical values as the conditional probability method. However, for the conditional probability method $N_{1,n}$ is equal or greater to the values in table 14a, meaning the estimated reporting performance should be regarded as upper limits.

Table 14b shows the combined results from the Chapman capture-recapture method, and it is easy to see that the results vary very little from the Lincoln-Petersen method. The final reporting performance are given as a 95% confidence interval, and it is clear to see the interval envelops the findings from table 14a.

Table 14b. Results using capture-recapture method (Chapman estimator)

Flag state	Number of registered accidents: Sea-Web	Number of registered accidents: Flag state	Common entries	Estimation of true number of accidents ($N_{2.2,n}$)	Estimated 95% CI reporting performance: Sea-Web		Estimated 95% CI reporting performance: Flag state	
Norway (NOR/NIS)	529	596	203	1 550	31 %	37 %	35 %	42 %
Sweden	109	333	86	421	24 %	28 %	73 %	86 %
Denmark (DS/DIS)	189	220	45	912	17 %	27 %	20 %	31 %
United Kingdom	401	1 428	229	2 497	15 %	17 %	53 %	62 %
United States	632	2 362	135	10 997	5 %	7 %	19 %	25 %
Canada	608	722	454	967	61 %	65 %	73 %	77 %
Greece (2009 only)	109	24	12	211	39 %	79 %	8 %	17 %
The Netherlands	304	342	71	1 452	18 %	25 %	20 %	29 %

Both Lincoln-Petersen and Chapman estimators are prone to underestimating the total value N_{mn} if there are dependency between the sources. Another capture-recapture method has been developed by Prof. Dr. Anne Chao, which is less susceptible to dependency between sources. However, this method is unable to improve the results of the Lincoln-Petersen or Chapman estimators if the two sources have similar values ($a_n \approx b_n$).

It should also be noted that Chao's estimation method calculates $N_{2.3,n}$ from a_n and b_n, not FS_n and $IHSF_n$ as the other methods. The results from Chao's lower bound estimation method is found in table 14c. One can easily see that the 95% confidence interval is generally more narrow and with somewhat lower values compared to the previous methods, while $N_{2.3,n}$ is larger.

Table 14c. Results using capture-recapture method (Chao's lower bound estimate)

Flag state	Flag state only (a_n)	Sea-Web only (b_n)	Common entries (C_n)	Estimation of true number of accidents ($N_{2.3,n}$)	Estimated 95% CI reporting performance Sea-Web		Flag state	
Norway (NOR/NIS)	393	326	203	1 559	32 %	36 %	37 %	40 %
Sweden	247	23	86	568	18 %	21 %	55 %	63 %
Denmark (DS/DIS)	175	144	45	929	18 %	23 %	21 %	27 %
United Kingdom	1 199	172	229	3 652	10 %	12 %	37 %	41 %
United States	2 227	497	135	16 600	4 %	4 %	13 %	15 %
Canada	268	154	454	974	62 %	63 %	73 %	75 %
Greece (2009 only)	12	97	12	369	24 %	39 %	5 %	9 %
The Netherlands	233	271	71	1 469	21 %	26 %	19 %	23 %

Table 15 shows the combined results from the best case scenario method. It assumes perfect reporting, in the sense that every accident that occurs are either reported to the respective flag state, or to IHS Fairplay. Elementary algebra then gives the total amount of reported accidents ($N_{3,n}$). The reporting performance found by dividing FS_n by $N_{3,n}$ or $IHSF_n$ by $N_{3,n}$ may then be regarded as the maximum possible coverage of the databases. The existence of a certain amount of unknown accidents will simply reduce the upper limit reporting performance found by this method. It is possible to rephrase the results from this method into a 100% confidence interval, ranging from 0% to the upper limit reporting performance. This interval is too large to be useful, but it returns an absolute upper limit, based on de facto occurred and reported accidents.

Table 15. Results using best case scenario

Flag state	Number of registered accidents Sea-Web	Flag state	Common entries	Minimum amount of occurred accidents ($N_{3,n}$)	Upper limit of reporting performance Sea-Web	Flag state
Norway (NOR/NIS)	529	596	203	922	57 %	65 %
Sweden	109	333	86	356	31 %	94 %
Denmark (DS/DIS)	189	220	45	364	52 %	60 %
United Kingdom	401	1 428	229	1 600	25 %	89 %
United States	632	2 362	135	2 859	22 %	83 %
Canada	608	722	454	876	69 %	82 %
Greece (2009 only)	109	24	12	121	90 %	20 %
The Netherlands	304	342	71	575	53 %	59 %

Table 16 shows the data from Cefor, and calculations of the average annual accident ratio for each flag state. The values represent the average number of claims and vessels from the five year period covered in this study, resulting in the average annual accident ratio which is then multiplied with the total fleet size and time interval. This returns an estimated reporting performance, which represents an accident ratio identical to Cefor's. It should be noted that some values exceed 100%, and should thus be discarded. It simply means that Cefor's accident ratio is better than the already established best case scenario accident ratio, and thus unusable for this purpose. However, this is only the case for 25% of the total data set.

Table 16. Results from scaling up Cefor's accident ratio

Flag state	ICC_n	ICF_n	R_n	TF_n	$N_{4,n}$, per year	$N_{4,n}$, whole period	Estimated reporting performance Sea-Web	Estimated reporting performance Flag state
Norway (NOR/NIS)	1 169	338	29 %	1 784	516	2 579	21 %	23 %
Sweden	179	81	45 %	348	157	787	14 %	42 %
Denmark (DS/DIS)	420	184	44 %	789	346	1 728	11 %	13 %
United Kingdom	338	105	31 %	863	268	1 340	30 %	107 %
United States	156	29	19 %	1 726	321	1 604	39 %	147 %
Canada	41	10	24 %	371	90	452	134 %	160 %
Greece	407	45	11 %	2 913	322	322	34 %	7 %
The Netherlands	307	116	38 %	845	319	1 596	19 %	21 %

5. Analysis

Several studies on underreporting of road accidents find the level of underreporting increases when the injury severity level decreases, as shown by Elvik and Mysen (2007). It is natural to assume a similar pattern in the maritime sector. As discussed by Amoros et al. (2008) the assumption of independence, in this case between IHS Fairplay and flag states, may be faulty. It would be logical to assume a certain degree of positive dependence for serious maritime accidents, due to global media coverage, interagency communication and social transparency. If that is the case, the Lincoln-Petersen and Chapman capture-recapture methods are biased downwards and underestimate N_{mn}, meaning the projected estimations of N_{mn} should be interpreted as lower limits. With increasing dependencies between sources the Lincoln-Petersen and Chapman estimators will steadily underestimate the population size. A more thorough study of the link between data sources would thus be beneficial to the accuracy and choice of estimation method (Böhning, 2010).

Method (1) using conditional probability returns almost identical values as the Lincoln-Petersen and Chapman estimators from methods (2.1) and (2.2), but conditional probability returns the number of true occurred accidents as a lower limit. However, method (1) assumes positive dependency between sources, in which case the Lincoln-Petersen and Chapman estimators will underestimate the true value.

A better estimator is thus Chao's lower bound estimator, from method (2.3). However, when there is symmetry between the two data sources ($a_n \approx b_n$), Chao is unable to improve on Lincoln-Petersen and Chapman, returning similar results. This can be seen in estimations of N_{mn} for Norway, Denmark and The Netherlands, in appendix (b).

The best case scenario (3) is based on all known accidents, returning an absolute lower limit of occurred accidents, and subsequently proving that underreporting exists. It is not a good measure of underreporting, but provides a definite limit.

The method of using Cefor data (4) to scale up ratios found in the NoMIS database has strong limitations with regards to validity of the required assumptions. The subset of vessels found in NoMIS is not representative of a flag state's whole fleet, as the underwriting goal of the insurance companies in Cefor is to avoid clients with bad records, and insure the more high quality fleets and companies. However, with that in mind, it stands to reason that the Cefor subset should have a lower accident ratio than the real fleet. Should the results from some of the flag states indicate a higher accident ratio among the NoMIS data than the best case scenario, which contains all known accidents, this could indicate that there is in fact a certain amount of accidents that go unreported. In the case of the UK, USA and Canada, the results go beyond 100% reporting performance, as the method returns a smaller amount of occurred accidents than what is the de facto amount, as proven by the best case scenario. This may be due to insufficient data, or the data not being representative. However, for Norway, Sweden and Denmark, the results using method (4) returns a lower reporting performance than even Chao, indicating major underreporting.

It should be pointed out that Cefor only receives claims for vessel damages exceeding the deductible, meaning there is even some "underreporting" to NoMIS. The necessary assumptions of uniform distribution and representative subsets are also not entirely correct, meaning this method should only be used as an indicator, not a proper estimation technique. That said, it is easy to explain why the results may be too low, but if the NoMIS accident ratio is higher than the flag state or Sea-Web accident ratio, it is an indication of underreporting, as the NoMIS subset is deemed to be "above-average".

Based on the author's assessment of the data quality provided from the flag state authorities, most emphasis should be placed on the results for Norway, The United Kingdom and Canada. The data from the other countries in this study is wrought with significantly more uncertainty and has at times a more variable and debatable quality.

This study estimates the amount of accidents that are not reported to flag state authorities or found in IHS Fairplay Sea-Web. However, there are often little or no formal link between flag state authorities and IHS Fairplay, meaning the flag state reporting performance is really the most interesting value. Nevertheless, commercial actors purchasing data from casualty databases such as IHS Fairplay should be aware of their reporting performance as well. The best flag state in this study, Canada, is missing roughly a quarter of all accidents occurring in their area of responsibility. Norway and The United Kingdom covers just over a third of all occurred accidents in their area of responsibility, despite these flag state authorities' high level of perceived competency and quality. These findings confirm those made by Psarros et al. (2010).

The large amount of underreporting uncovered by this study indicates that all users of statistical data should assume a certain amount of underreporting, and adjust their analyses accordingly. Whether they use a correction factor, safety margin or rely more heavily on expert judgement must be decided on a case by case basis.

Chao's lower bound estimator is regarded as the most accurate estimator for this study. The use of capture-recapture estimators to uncover underreporting of maritime accidents should be explored further, preferably by statisticians with more in-depth knowledge of these methods.

6. Discussion

It is hard to draw any general conclusions based on the results from the studied flag states, but a clear pattern across the estimation methods can be found. The conditional probability, Lincoln-Petersen and Chapman estimators returns virtually identical values, while Chao's estimation consistently returns a higher total number of occurred accidents. As previously mentioned, the issue of dependency between data sources is key. Dependency will vary, both between the studied flag states and with time. Future research should try to uncover the relationship between flag state authorities' accident statistics, and commercial sources, such as IHS Fairplay or Seasearcher - Lloyd's MIU.

This study has not had much focus on the dependency and relationship between data sources, or tried to map out how the data sources collect and receive their data. This may be necessary to better understand which estimation technique would return the best results, as all estimators come with a set of requirements to work optimally.

Lack of transparency and knowledge of how the major commercial maritime data providers such as IHS Fairplay and Lloyd's MIU collect and receive information can be a problem. They regard the information gathering process as a trade secret, and are very vague in their reply to any inquiry regarding their input data. It is also worth mentioning that these data sources, widely used and highly regarded as they may be, are subject to faults and omissions, just like any other product. LMIU's terms of service clearly states: *"Please note that although we try to ensure that the content of our website, the online service and the materials is accurate, our website, the online service and/or the materials may contain errors, omissions or inaccuracies. … All materials which are supplied by third parties are published in good faith but we do not ... accept responsibility for the accuracy ... for the use of those materials. You assume total responsibility and risk for your use of the materials and the online service. … We accept no liability for any indirect or consequential loss or damage, or for any loss of data, profit, revenue or business (whether direct or indirect) in each case, however caused, even if foreseeable. In circumstances where you suffer loss or damage arising out of or in connection with the viewing, use or performance of the online service or the materials, we accept no liability for this loss or damage whether due to inaccuracy, error, omission or any other cause and whether on the part of Informa or our servants, agents or any other person or entity."* During this study, a significant error was discovered in IHS Fairplay functionality, purely by chance. How many other errors are still unnoticed? How do they affect research and work done on the basis of information provided from these sources? The discovery is in stark contrast to LMIU's slogan: *"The only reliable marine casualty reporting service in the world!"* (Seasearcher, 2010).

Another aspect of commercial data sources is that their main goal is to make a profit. They have some search capability for the average user/subscriber, but to perform a more complex/detailed SQL type query, the user must buy a manual/custom search, usually costing several thousand pounds (GBP). The Seasearcher LMIU is not even an SQL database, but simply an advanced text list/stream. They have also stated that they view their data as proprietary, and are not keen on users exporting data to their own computers, for research or other uses. Presently, LMIU does not have an export data functionality on the Seasearcher website, and IHS Fairplay have an upper export limit of 2500 rows of data for any particular query. Due to the limitations of their search functionality, it may be problematic to split a large search (>2500 returned rows) into several smaller queries, for the intention of exporting. The only other option is again to pay for a manual search. Research requires free and unbiased access to information, which in turn may be processed and analysed, and finally published. The commercially operated casualty databases have very strong restrictions on what they regard proprietary information. Again, here is another excerpt from LMIU's terms of service: *"2.3 - A single user license means that during the subscription period, in relation to the materials to which you subscribe through the online service:*

(a) *you may:*

- (i) *display such Materials electronically (on a single computer screen, mobile telephone or personal digital assistant) to one concurrent User at any given time;*

(ii) *download and store one copy of such Materials in machine readable form;*
(iii) *print one copy of such Materials; and*
(iv) *use such Materials solely for the internal research purposes of the Licensee.*

(b) *you may not:*
(i) *download, store, reproduce, transmit, display, copy, distribute, commercially exploit or use the Materials other than as expressly permitted in sub-clause 2.3 (a) above;*
(ii) *resell, sub-licence, rent, lease, transfer or attempt to assign the rights in the Materials (in whole or in part) to any other person;*
(iii) *make the Materials available (in whole or in part) on a Computer Network;*
(iv) *distribute the Materials (in whole or in part) via an intranet or global network;*
(v) *use the Materials in any manner, (or transfer or export the Materials or any copies into any country), other than in compliance with applicable laws;*
(vi) *allow any person to use and/or gain access to the Materials other than in accordance with these Terms;*
(vii) *allow any person other than an authorised User to use and/or gain access to the Materials or*
(viii) *modify, alter or create derivative works from such Materials nor may you create a database in electronic or structured manual form by systematically downloading and storing any of the content from such Materials."*

(Seasearcher, 2010).

From a research perspective, this kind of practice is unfortunate, as it limits other scientists to verify and recreate results. Following the terms of service to the letter, it is hardly possible to perform any research at all with data from this source. IHS Fairplay has slightly less stringent terms of use, as they do allow some exporting of data, but they are also relatively restrictive. In contrast, most (western/developed) countries have some form of Freedom Of Information Act legislation, enabling researchers access to nearly all recorded national data.

The lack of standardization and common database structure is another problem for researchers. It makes data-matching more complex, and prone to errors. Admittedly, the reasons behind this is often that the databases have different purposes, and not all are primarily intended as casualty repositories. Nevertheless, the complete lack of quality and content benchmarks should be addressed.

The question of ownership is another issue that should be mentioned. The only working global databases are all commercially operated. There is no global casualty database, except for the IMO's Global Integrated Shipping Information System (GISIS) - Marine Casualties and Incidents Module. As concluded by Hassel and Hole (2009), GISIS is an empty shell of good intentions. No real and tangible information is available, and accident reports are not publicly available, if they even exist. Lately, a European Union (EU) initiative under the European Maritime Safety Agency (EMSA), has emerged with the European Marine Casualty Information Platform (EMCIP), to serve as a European international casualty database. It links together European maritime authorities and is working towards a more standardized model of international casualty recording. The EMCIP initiative looks promising, and has a great potential for improving the current state of fragmented and incomplete casualty records (EMCIP, 2010; EMSA, 2010; Hassel and Hole, 2009; IMO, 2010).

One of the missing functionalities needed to do a better estimation of maritime authorities' reporting performance is to compare all casualties within their area of responsibility. Due to IHS Fairplay's current system of geographical location categorization, this is not possible. The current study has thus only investigated flag states' reporting performance of own flag, and has not included foreign vessels in territorial waters. According to Hole (2010), Norwegian flagged ships in foreign waters often ignore local rules and regulations regarding reporting procedures. There are no reason to believe that foreign vessels in Norwegian waters are any better than their Norwegian counterparts. This paper has not been able to properly study underreporting among vessels outside their national borders. This shortcoming could be circumvented if it was possible to have casualty data integrated with map services, such as Google Maps. The visualization of accidents could provide the necessary filter to allow for more detailed studies. The use of visualization in maps have been done by the International Chamber of Commerce - Commercial Crime Services (ICC-CCS) to plot the dynamic progression of piracy incidents for several years now. It is neither expensive nor difficult, if only the data providers would see the added value it provides to the users.

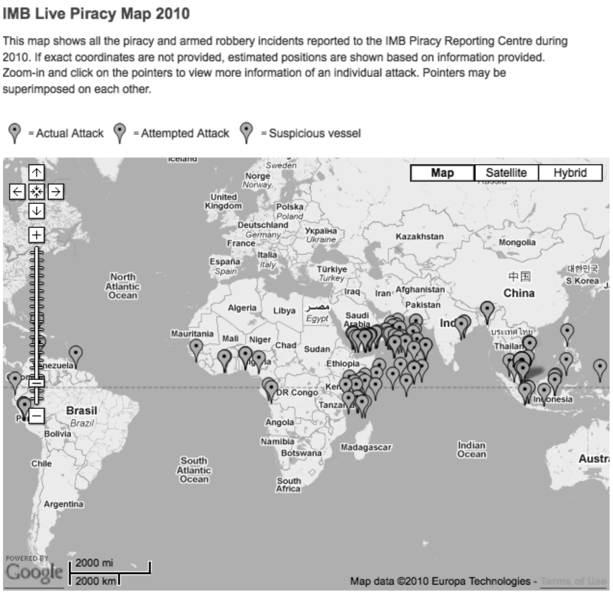


Figure 4.
Example of visual representation of dynamic incident data, enabled by integrating simple mapping technology. (ICC Commercial Crime Services, 2010)

Further research should investigate any relation that may exist between underreporting and ship types, geographical area or other relevant factors. If one is only interested in a certain ship type, and a specific threat, for instance collision between a platform supply ship and an offshore installation, it is possible to use the same methodology employed by this paper, to give more customized answers. The only problem is to gather enough data, as national databases are all sorted by flag. The NMD only covers Norwegian ships, the STA only covers Swedish ships and so on. If one were to look at the world LNG fleet, it would require the researcher to identify what flags the majority of the LNG fleet is carrying, and then gather data from all these flag state authorities. This issue may be relieved some with the implementation of EMCIP, but only time will tell how successful this will be.

A simpler and more instant improvement could be achieved if IHS Fairplay could add a few more relevant attributes to their vessel data cards, and implement map integration with their database. Information about the ship's position, being within national borders of a country or at the high seas, would instantly close the current gap of ship accidents in foreign waters. Sorting by country location would enable researchers to fully investigate a flag state's underreporting. Adding more relevant attributes to casualty data is a whole different story, which is outside the scope of this study. It is briefly discussed by Hassel and Hole (2009), in their comparative analysis of casualty databases.

However, not all risk management companies rely on statistical data as much as one may believe. Simulation models are often based on expert judgement and general ship and operational data, such as vessel specifics, AIS and port information, and not necessarily casualty statistics. Some analyses do require the use of casualty statistics, but in those cases the basic data are usually quality assured and evaluated before use. That said, inaccurate basic data will most often have a significant impact on the end result (Haugen, 2010).

The flag state databases in this study have not been subject to in-depth analysis, but through the interaction, communication and study of returned data, it is possible to give a general brief impression on some of the flag states.

Canada, the United Kingdom and Norway all showed a high level of structure, proper data management, adequate completeness of casualty records, easy access and public transparency. The U.S.A., Denmark and Greece did not impress, as they had more inaccurate data, poor structure and completeness and a more cumbersome process of data retrieval.

The issue of underreporting is not merely one of quantifying an unknown amount, but also the investigation of the underlying reasons why underreporting occurs. This is addressed in more detail by Hole (2010). Only when both the extent and causality of underreporting is known, can proper measures be taken to adjust legislation and incentives to improve the situation.

Acknowledgements

The author would like to thank the Norwegian Research Council for their financial support to the MARRISK and FARGE research projects, in which this study is part of the preliminary research work. The author would also like to thank all persons and organisations that have contributed with data and expertise to this paper.

Bibliography

Alsop J. and Langley J., 2001. Under-reporting of motor vehicle traffic crash victims in New Zealand. Accident Analysis & Prevention 33 (3), 353-359.

Amoros E., Martin J.L. and Laumon B., 2006. Under-reporting of road crash casualties in France. Accident Analysis & Prevention 38 (4), 627-635.

Amoros E., Martin J.L. and Laumon B., 2007. Estimating non-fatal road casualties in a large French county, using the capture-recapture method. Accident Analysis & Prevention 39 (3), 483-490.

Amoros E., Martin J.L., Lafont S. and Laumon B., 2008. Actual incidences of road casualties, and their injury severity, modeled from police and hospital data, France. European Journal of Public Health 18 (4), 360-365.

Aptel I., Salmi L.R., Masson F., Bourdé A., Henrion G. and Erny P., 1999. Road accident statistics: discrepancies between police and hospital data in a French island. Accident Analysis & Prevention 31 (1-2), 101-108.

Brittain S. and Böhning D., 2009. Estimators in capture–recapture studies with two sources. AStA Advances in Statistical Analysis 93 (1), 23-47.

Böhning D., 2010. Professor in Applied Statistics. University of Reading. (E-mail correspondence)

CARE (Community database on Accidents on the Roads in Europe) 2010.
Available at:
http://ec.europa.eu/transport/road_safety/index_en.htm

Cefor's Nordic Marine Insurance Statistics Database (NoMIS), 2010. Ms Astrid Seltmann.
(E-mail correspondence)

Danish Maritime Authority, 2010. Ms Lea-Valeska Giebel and Mr Christian Hauge Alminde.
(E-mail correspondence)

Desenclos J.C. and Hubert B., 1994. Limitations to the universal use of capture-recapture methods. International Journal of Epidemiology 23 (6), 1322–1323.

Devanney J., 2009. Uses and Abuses of Ship Casualty Data. Center for Tanker Excellence.

Elvik R. and Mysen A. B., 2007. Incomplete Accident Reporting: Meta-Analysis of Studies Made in 13 Countries. Transportation Research Record: Journal of the Transportation Research Board 1665 (1999), 133-140.

European Marine Casualty Information Platform, 2010. Available at:
http://emcipportal.jrc.ec.europa.eu

European Maritime Safety Agency, 2010.
Available at: http://www.emsa.europa.eu

Galea E.R., Finney K.M., Dixon A.J.P., Siddiqui A. and Cooney D.P., 2006. Aircraft Accident Statistics and Knowledge Database: Analyzing Passenger Behavior in Aviation Accidents. Journal of Aircraft 43 (5), 1272-1281.

Harris S., 1990. The real number of road traffic accident casualties in the Netherlands: A year-long survey. Accident Analysis & Prevention 22 (4), 371-378.

Hassel M. and Hole L.P., 2009. Marine Databases - A comparative study, from a risk management perspective. Project Thesis, The Norwegian University of Science and Technology, Department of Marine Technology. (To be published)

Hauer E. and Hakkert A.S., 1988. Extent and some implications of incomplete accident reporting. Transportation Research Record 1185, 1-10.

Haugen S. (Prof. Dr.), 2010. Chief Executive Research & Development, Safetec Nordic AS. (E-mail correspondence)

Hellenic Coast Guard, Marine Casualties Department. 2010. Lt Constantinos Moutzouvis (Phone conversation and e-mail correspondence)

Hole L.P., 2010. Mapping the reporting procedures of accidents in the maritime transportation sector. Master Thesis, The Norwegian University of Science and Technology, Department of Marine Technology. (To be published)

Hook, E. and Regal, R., 1995. Capture–recapture methods in epidemiology: methods and limitations. Epidemiologic Reviews 17 (2), 243–264.

IHS Fairplay - Sea-Web, Casualty Database. 2010. (February and April versions) Available at: http://www.sea-web.com

IHS Fairplay - Sea-Web, Customer Services. 2010. Helen Kent, Senior Technical Services Analyst. (Phone conversation and e-mail correspondence)

IMO, 2005. Reports on marine casualties and incidents. MSC-MEPC.3/Circ.1 - 26th September 2005, International Maritime Organization.

IMO, 2007. Guidelines for formal Safety Assessment (FSA) MSC 83/INF. 2, 14th May 2007. MSC/Circ.1023 – MEPC/Circ.392, International Maritime Organization.

IMO, 2010. Global Integrated Shipping Information System - Marine Casualties and Incidents Module. Available at: http://gisis.imo.org/Public

Inspectie Verkeer en Waterstaat (The [Dutch] Transport and Water Management Inspectorate), 2010. Mr Sander Walpot. (E-mail correspondence)

International Chamber of Commerce - Commercial Crime Services, 2010. Piracy Map. Available at: http://www.icc-ccs.org/index.php?option=com_fabrik&view=visualization&controller=visualization.googlemap&Itemid=219

IRTAD (International Traffic Safety Data and Analysis Group), 2010. Available at: http://www.internationaltransportforum.org/irtad/about.html

Japanese Transport Safety Board, 2010. Recommendation and Opinion Section, Mr Hideshige Imamura. (E-mail correspondence)

Johnson C., 2002. Software tools to support incident reporting in safety-critical systems. Safety Science 40, 765–780.

Liberian Registry, 2010. Vice President Marine Investigations & Assistant General Counsel, Mr Michael Davies-Sekle. (E-mail correspondence)

Lloyd's Register (Fairplay), 2009. World Casualty Statistics 2008. ISBN 978-1-096313-22-7

Maritime Accident Investigation Branch, 2005. The Merchant Shipping (Accident Reporting and Investigation) Regulations.

Maritime Accident Investigation Branch, 2010. Mr Luke Grant. (E-mail correspondence)

National Transportation Safety Board, 2002. Transportation Safety Databases. NTSB/SR-02/02.

Norwegian Maritime Directorate, 2010. Strategic Safety Division. Mr Håvard Gåseidnes. (Phone conversation and e-mail correspondence)

OECD-IRTAD (International Road Traffic and Accident Database), 2010. Available at: http://www.oecd.org/document/53/0,2340,en_2649_34351_2002165_1_1_1_1,00.html

OECD-IRTAD, 1994. Under-reporting of road traffic accidents recorded by the police at the international level. Special report by Hvoslef H., Public Roads Administration of Norway, Oslo.

Psarros, G., Skjong R. and Eide M.S., 2010. Under-reporting of maritime accidents. Accident Analysis and Prevention 42 (2), 619-625.

Rose A., 2006. Measuring operational safety in aviation. Aircraft Engineering and Aerospace Technology 78 (1) 26-31.

Ryan B., Hutchings J. and Lowe E., 2010. An analysis of the content of questions and responses in incident investigations: Self reports in the investigation of signals passed at danger (SPADs). Safety Science 48, 372–381.

Seasearcher - Lloyd's List Intelligence, 2010. James Dean, Sales Director.
(E-mail correspondence)

Seasearcher - Lloyd's MIU, Casualty Database. 2010. Available at: http://www.seasearcher.com

Seasearcher - Terms of service, 2010. Available at: http://www.seasearcher.com/lmiu/terms.htm

Soares C.G. and Teixeira A.P., 2001. Risk assessment in maritime transportation. Reliability Engineering and System Safety 74 (3), 299-309.

Sciortino S., Vassar M., Radetskya M. and Knudson M.M., 2005. San Francisco pedestrian injury surveillance: Mapping, under-reporting, and injury severity in police and hospital records. Accident Analysis & Prevention 37 (6), 1102-1113.

Swedish Transport Agency, 2010.
Accident Investigation Office. Ms Gunnel Persson.
(E-mail correspondence)

Thomas, M. and Skjong, R., 2009. Cost benefit analysis of inert gas systems for chemical and product tankers. In: Proceedings of the 28th International Conference on Ocean, Offshore and Arctic Engineering, Honolulu, OMAE 2009-79919.

Transportation Safety Board of Canada, 2010.
Ms Francine Rochefort.
(Phone conversation and e-mail correspondence)

Trépanier M., Leroux M.H. and de Marcellis-Warin N., 2009. Cross-analysis of hazmat road accidents using multiple databases. Accident Analysis & Prevention 41 (6), 1192-1198.

Tzannatos E. and Kokotos D., 2009. Analysis of accidents in Greek shipping during the pre- and post-ISM period. Marine Policy 33, 679–684.

United Nations Conference on Trade and Development, 2005-2009. Review of Maritime Transport. Available at: http://www.unctad.org/Templates/Page.asp?intItemID=2618

United States Coast Guard, 2010.
Ms Dawn Patterson. (E-mail correspondence)

Wang J. and Foinikis P., 2001. Formal safety assessment of containerships. Marine Policy 25 (2), 143-157.

Yamamoto T., Hashiji J. and Shankar V.N., 2008. Underreporting in traffic accident data, bias in parameters and the structure of injury severity models. Accident Analysis & Prevention 40 (4), 1320-1329.

Appendix (a) - Chao estimates

Figure (i).

Distribution of accident reporting for Norwegian flagged vessels (NOR/NIS), from 01.01.2005 to 31.12.2009

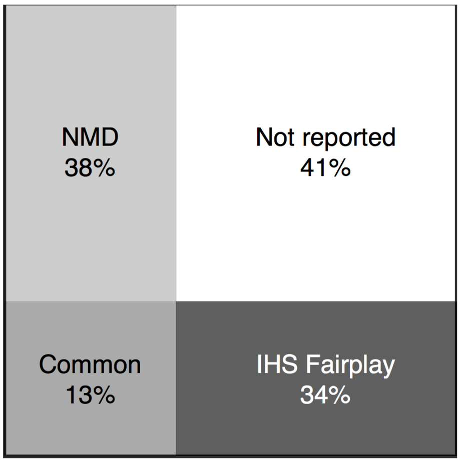


Figure (ii).

Distribution of accident reporting for Swedish flagged vessels, from 01.01.2005 to 31.12.2009

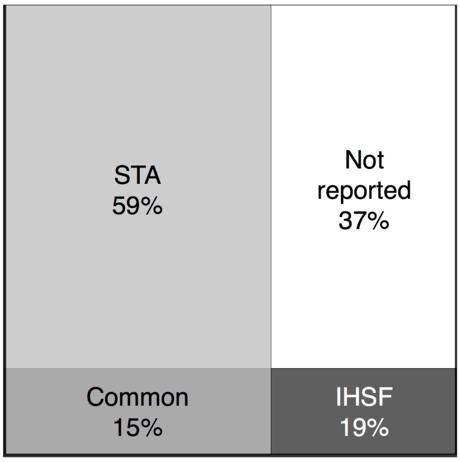


Figure (iii).

Distribution of accident reporting for Danish flagged vessels (DS/DIS), from 01.01.2005 to 31.12.2009

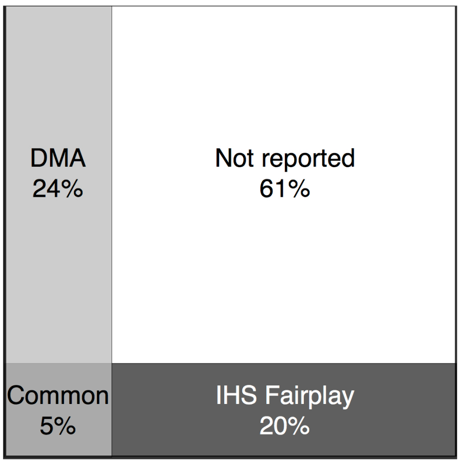


Figure (iv).

Distribution of accident reporting for British flagged vessels, from 01.01.2005 to 31.12.2009

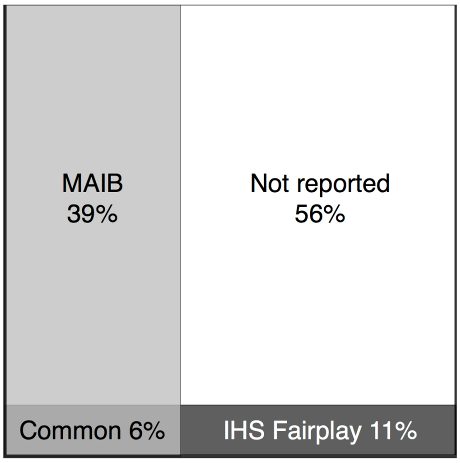


Figure (v).

Distribution of accident reporting for American flagged vessels, from 01.01.2005 to 31.12.2009

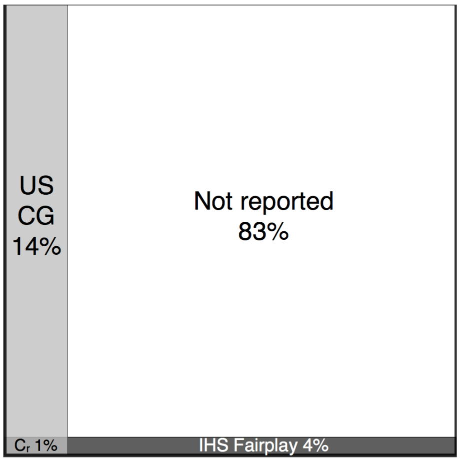


Figure (vi).

Distribution of accident reporting for Canadian flagged vessels, from 01.01.2005 to 31.12.2009

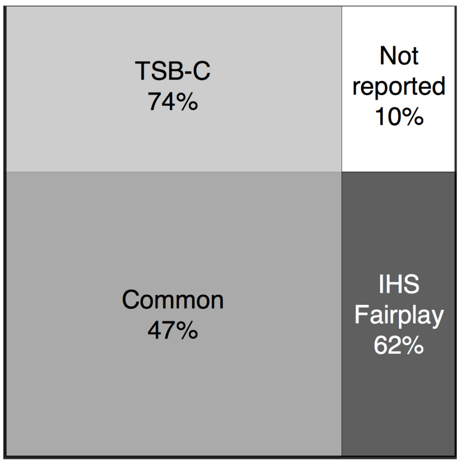


Figure (vii).
Distribution of accident reporting for Greek flagged vessels, from 01.01.2009 to 31.12.2009

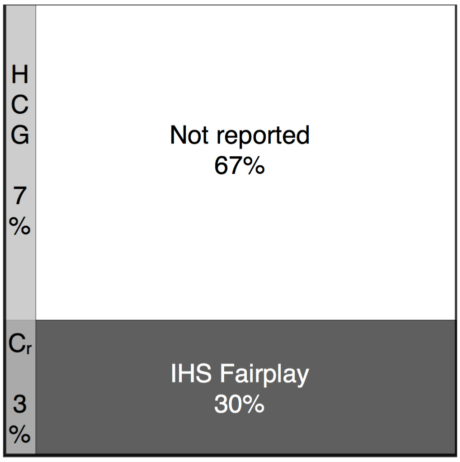
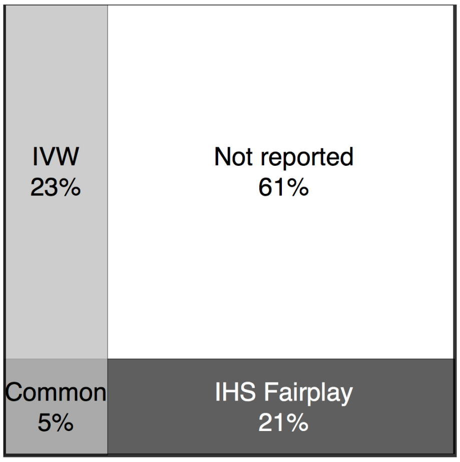


Figure (viii).
Distribution of accident reporting for Dutch flagged vessels, from 01.01.2005 to 31.12.2009



Appendix (b) - Reporting performance

Conditional probability (1) | Capture/recapture - Chao (2) | Best Case (3) | Cefor (4)

Figure (ix).

Estimated reporting level for Norwegian flagged vessels (NOR/NIS), from 01.01.2005 to 31.12.2009

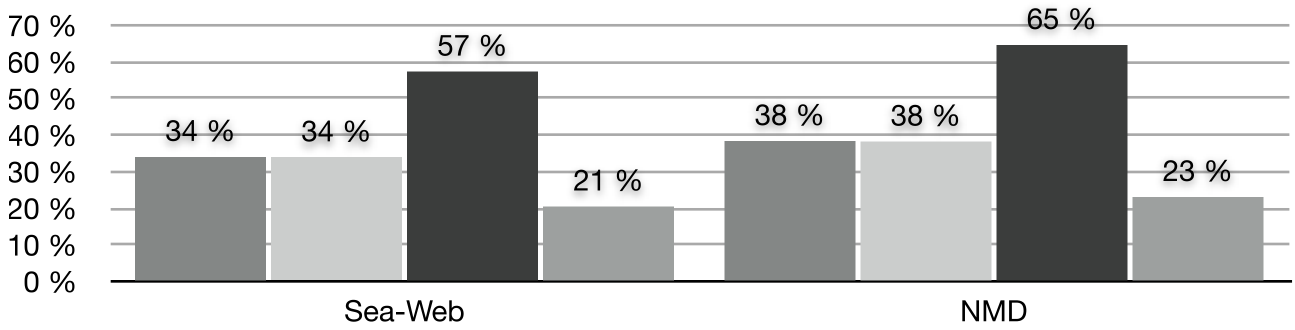


Figure (x).

Estimated reporting level for Swedish flagged vessels, from 01.01.2005 to 31.12.2009

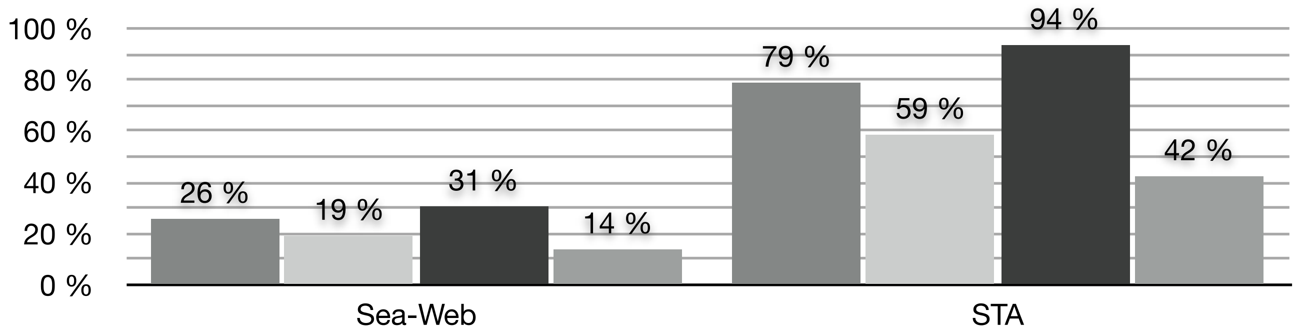


Figure (xi).

Estimated reporting level for Danish flagged vessels (DS/DIS), from 01.01.2005 to 31.12.2009

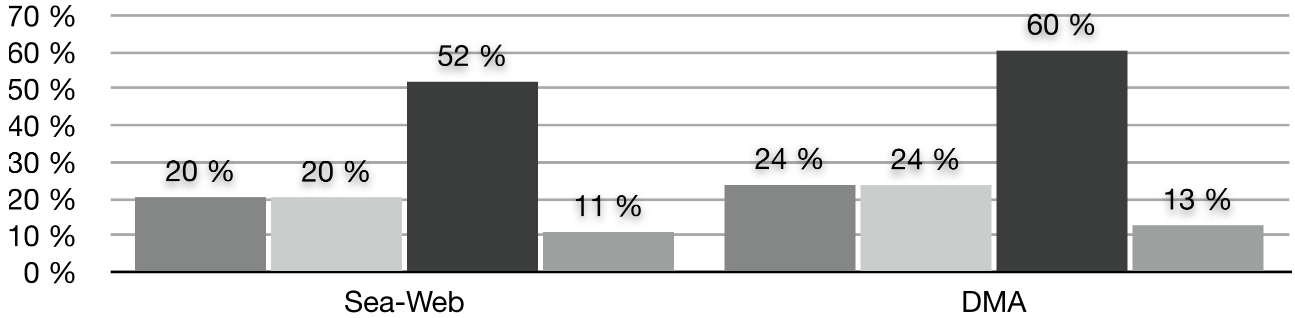


Figure (xii).

Estimated reporting level for British flagged vessels, from 01.01.2005 to 31.12.2009

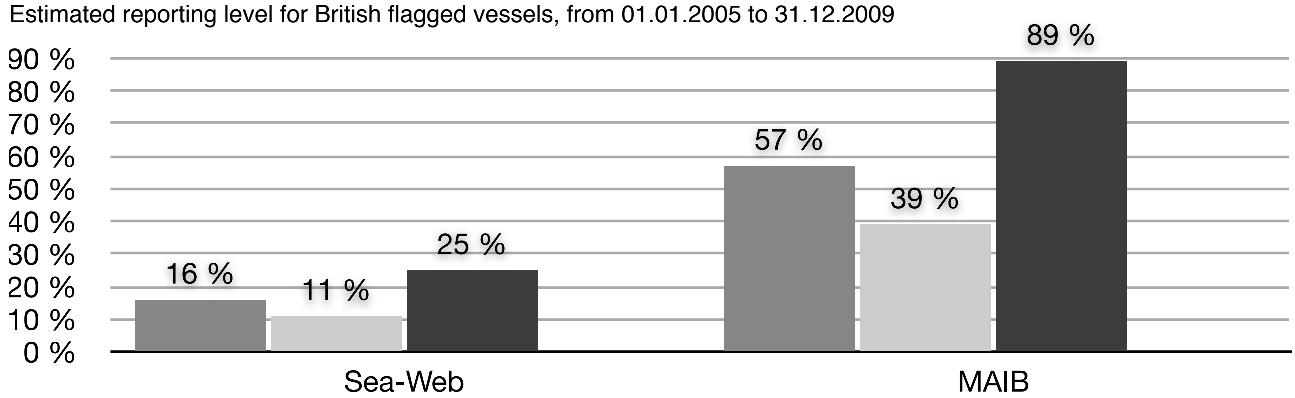


Figure (xiii).

Estimated reporting level for U.S. flagged vessels, from 01.01.2005 to 31.12.2009

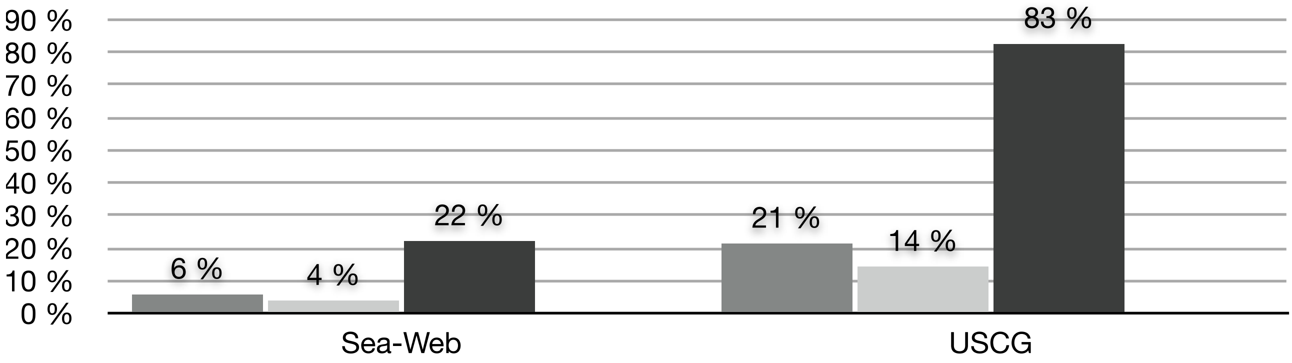


Figure (xiv).

Estimated reporting level for Canadian flagged vessels, from 01.01.2005 to 31.12.2009

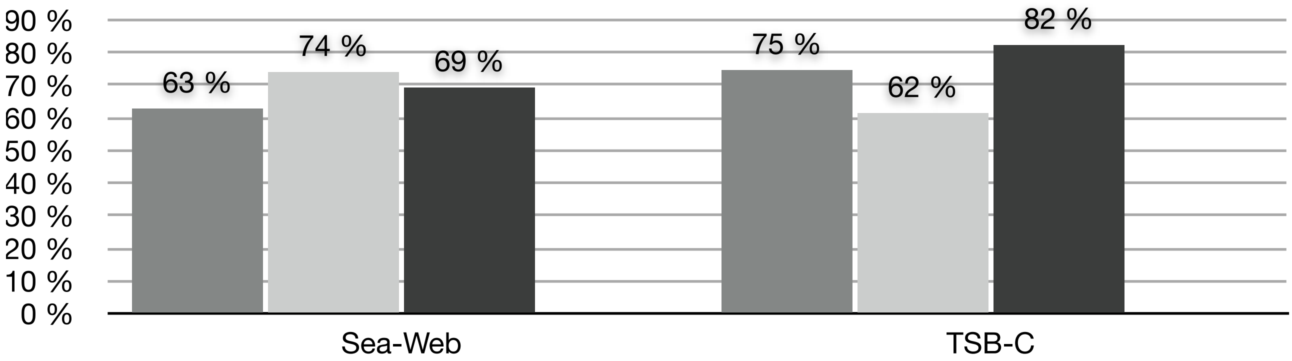


Figure (xv).

Estimated reporting level for Greek flagged vessels, from 01.01.2009 to 31.12.2009

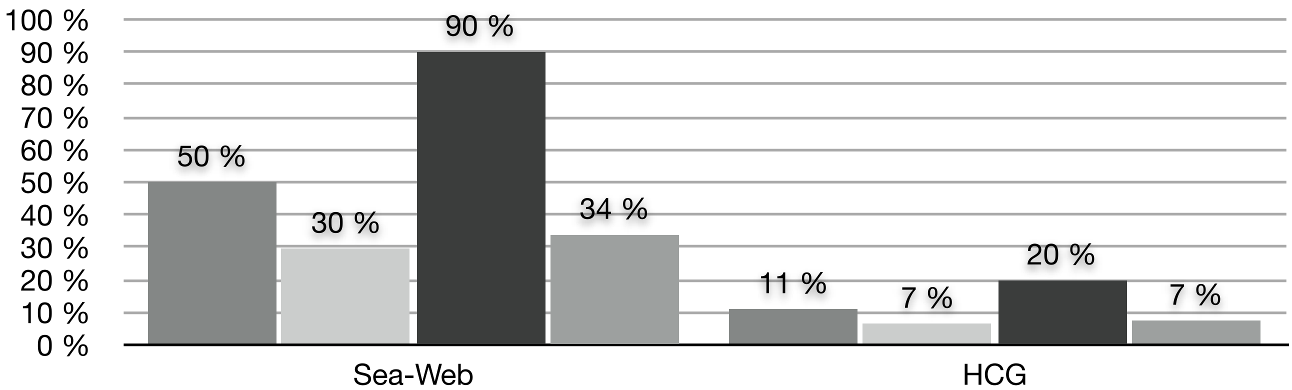
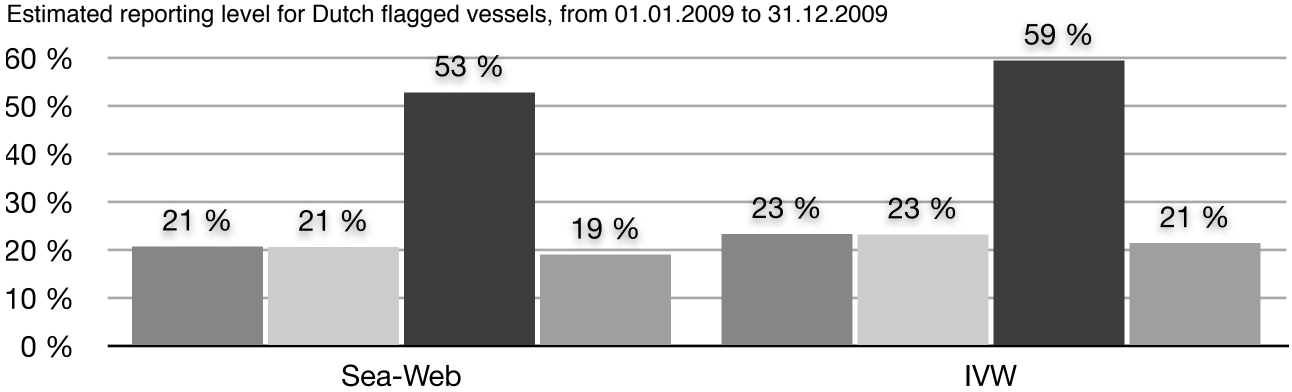


Figure (xvi).

Estimated reporting level for Dutch flagged vessels, from 01.01.2009 to 31.12.2009



Appendix (c) - Chao's estimated distribution

Table (i).
Chao's lower bound estimate of unreported accidents, for Norwegian flagged vessels, from 01.01.2005-31.12.2009

Accidents in		NMD		
		Yes	No	Sum
Sea-Web	No	25 %	41 %	
	Yes	13 %	21 %	34 %
	Sum	38 %		

Table (ii).
Chao's lower bound estimate of unreported accidents, for Swedish flagged vessels, from 01.01.2005 to 31.12.2009

Accidents in		STA		
		Yes	No	Sum
Sea-Web	No	43 %	37 %	
	Yes	15 %	4 %	19 %
	Sum	59 %		

Table (iii).
Chao's lower bound estimate of unreported accidents, for Danish flagged vessels, from 01.01.2005 to 31.12.2009

Accidents in		DMA		
		Yes	No	Sum
Sea-Web	No	19 %	61 %	
	Yes	5 %	15 %	20 %
	Sum	24 %		

Table (iv).
Chao's lower bound estimate of unreported accidents, for British flagged vessels, from 01.01.2005 to 31.12.2009

Accidents in		MAIB		
		Yes	No	Sum
Sea-Web	No	33 %	56 %	
	Yes	6 %	5 %	11 %
	Sum	39 %		

Table (v).
Chao's lower bound estimate of unreported accidents, for U.S. flagged vessels, from 01.01.2005 to 31.12.2009

Accidents in		USCG		
		Yes	No	Sum
Sea-Web	No	13 %	83 %	
	Yes	1 %	3 %	4 %
	Sum	14 %		

Table (vi).
Chao's lower bound estimate of unreported accidents, for Canadian flagged vessels, from 01.01.2005-31.12.2009

Accidents in		TSB-C		
		Yes	No	Sum
Sea-Web	No	28 %	10 %	
	Yes	47 %	16 %	62 %
	Sum	74 %		

Table (vii).
Chao's lower bound estimate of unreported accidents, for Greek flagged vessels, from 01.01.2009 to 31.12.2009

Accidents in		HCG		
		Yes	No	Sum
Sea-Web	No	3 %	67 %	
	Yes	3 %	26 %	30 %
	Sum	7 %		

Table (viii).
Chao's lower bound estimate of unreported accidents, for Dutch flagged vessels, from 01.01.2005-31.12.2009

Accidents in		IVW		
		Yes	No	Sum
Sea-Web	No	18 %	61 %	
	Yes	5 %	16 %	21 %
	Sum	23 %		

Appendix (d) - Cefor data

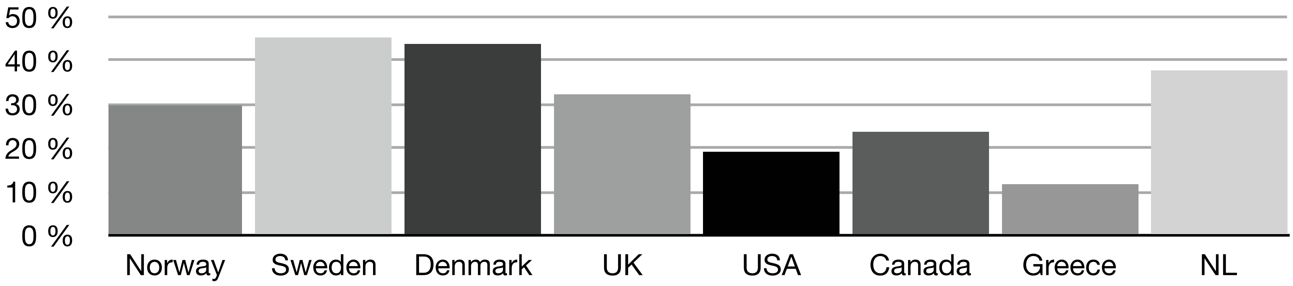
Table (ix).

Number of registered vessels and claims in Cefor's Nordic Marine Insurance Statistics Database, from 2005-2009

		2005	2006	2007	2008	2009	Total	Average
Norway	Vessels	1 142	1 144	1 195	1 201	1 170	5 852	1 170
	Claims/accidents	327	302	421	373	321	1 744	349
Sweden	Vessels	382	416	429	436	445	2 108	422
	Claims/accidents	180	203	198	194	148	923	185
Denmark	Vessels	163	181	193	200	170	907	181
	Claims/accidents	103	81	78	90	59	411	82
UK	Vessels	301	322	351	369	354	1 697	339
	Claims/accidents	95	104	138	133	78	548	110
USA	Vessels	160	157	177	162	122	778	156
	Claims/accidents	24	29	42	40	14	149	30
Canada	Vessels	50	48	39	34	32	203	41
	Claims/accidents	10	7	6	17	8	48	10
Greece	Vessels	444	428	396	393	369	2 030	406
	Claims/accidents	63	47	44	40	44	238	48
NL	Vessels	368	358	338	277	195	1 536	307
	Claims/accidents	146	124	141	113	55	579	116

Figure (xvii).

Vessel accident ratio in Cefor's Nordic Marine Insurance Statistics Database, average values from 2005-2009



Appendix (e) - Basic data

Table (x).

Number of registered accident in IHS Fairplay and flag state autorities, from 01.01.2005 to 31.12.2009

		Total number of registered accidents		Total common entries
Flag state		Sea-Web	Flag state	
Norway (NOR/NIS)		529	596	203
	2005	89	105	20
	2006	76	93	27
	2007	106	132	52
	2008	123	115	41
	2009	135	151	63
Sweden		109	333	86
	2005	18	70	16
	2006	24	54	21
	2007	23	78	17
	2008	21	65	14
	2009	23	66	18
Denmark (DS/DIS)		189	220	45
	2005	26	34	2
	2006	50	46	13
	2007	39	40	10
	2008	39	58	11
	2009	35	42	9
United Kingdom		401	1 428	229
	2005	84	318	42
	2006	75	297	47
	2007	81	286	58
	2008	79	252	44
	2009	80	274	38
United States		632	2 362	135
	2005	132	452	28
	2006	128	499	21
	2007	128	447	33
	2008	150	487	32
	2009	94	477	21
Canada		608	722	454
	2005	146	159	107
	2006	117	139	93
	2007	118	137	85
	2008	115	149	84
	2009	112	138	85
Greece (2009 only)		109	24	12
The Netherlands		304	342	71
	2005	43	61	6
	2006	59	70	14
	2007	78	84	22
	2008	82	94	21
	2009	42	33	8

Appendix (f) - Fleet sizes

Table (xi).

Flag states' fleet size, according to United Nations Conference on Trade And Development, Review of Maritime Transport 2005-2009. (UNCTAD 2005-2009)

Flag state	2005	2006	2007	2008	2009	Average
Norway (NOR/NIS)	1 589	1 665	1 810	1 827	2 027	1 784
Sweden	322	342	346	365	367	348
Denmark (DS/DIS)	646	744	781	861	914	789
United Kingdom	885	779	855	876	918	863
United States	1 633	1 679	1 766	1 769	1 782	1 726
Canada	325	356	340	419	413	371
Greece	2 984	2 318	3 084	3 115	3 064	2 913
The Netherlands	705	722	739	762	1 296	845

Additional material

(I) Electronic copy of raw data

The attached CD contains (amongst others) the following files:

- Excel spreadsheets with the unprocessed data from all flag states mentioned in this paper.
- Electronic copy of UNCTAD's Review of Maritime Transport 2005-2009
- Relevant IMO documents
- High resolution PDF versions of all equations used in this paper
- Electronic copy of relevant reference material
- Spreadsheets of processed data and interim calculations