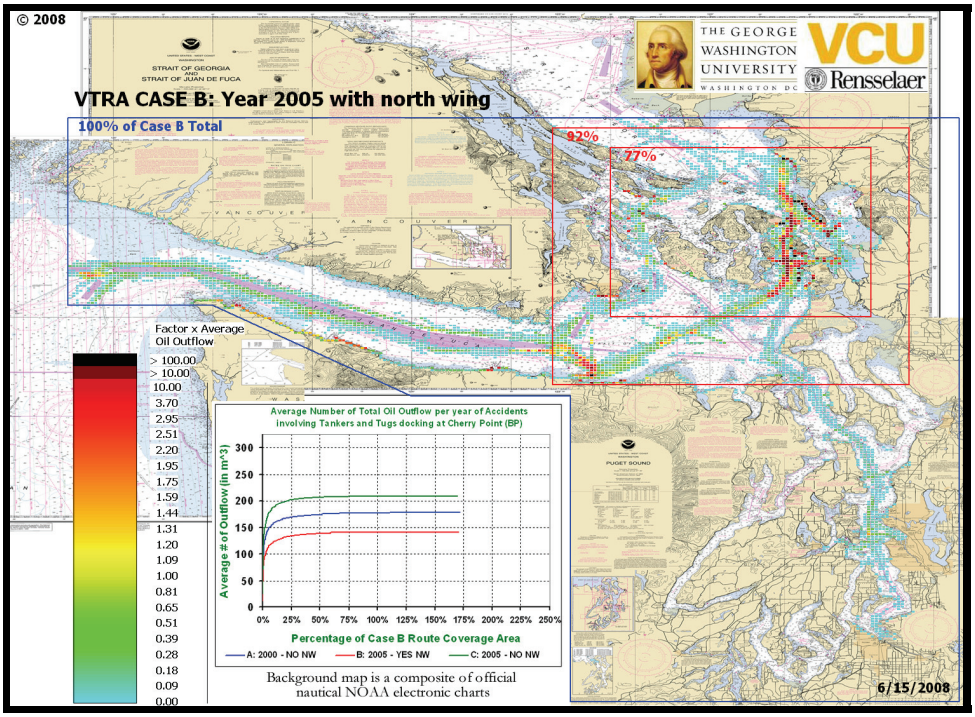
TECHNICAL APPENDIX E:
OIL OUTFLOW MODEL



Assessment of Oil Spill Risk due to Potential Increased Vessel Traffic at Cherry Point, Washington

Submitted by VTRA TEAM:

Johan Rene van Dorp (GWU), John R. Harrald (GWU),

Jason R.. W. Merrick (VCU) and Martha Grabowski (RPI)

August 31, 2008

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF FIGURES (Continued)

TABLE OF TABLES

E-1. The NRC oil outflow report

"The International Maritime Organization (IMO) is responsible for regulating the design of oil tankers to provide for ship safety and environmental protection. ... IMO's first attempt to apply a probabilistic methodology to tankers was in response to the US Oil Pollution Act of 1990 (OPA 90). In OPA 90 the US required that all oil tankers entering US waters must have double hulls. ... IMO responded to this unilateral action by requiring double hulls or their equivalent. Equivalency is determined based on probabilistic oil outflow calculations specified in IMO (1995)." (see, Brown (1995)). The purpose of the IMO model is to measure outflow performance of a particular tanker design. For this model, data was taken from approximately 100 historical collision and grounding scenarios from the period 1980-1990 to establish probability density functions (PDFs) for the location and extent of damage in a collision or grounding scenario (see Figure E-1). Based on these distributions, each unique combination of tanks or compartments in a given tanker design can be associated with a probability of being damaged.

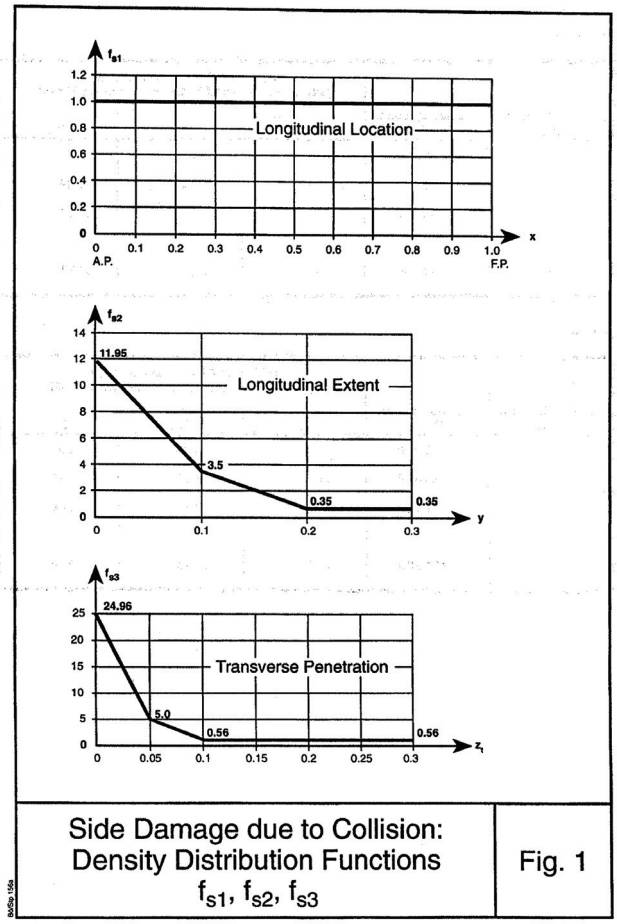


Figure E-1. Damage Extend PDFs, IMO Model (1995)

Unfortunately, the IMO model suffers from a number of fundamental limitations. Some of the objections raised by for example by Van der Laan(1997) and Brown (1998) (amongst others) are :

- The model uses a single set of damage extent PDFs from limited single hull data applied to all ships, independent of structural design; realistically, however, this data should only be used to model single hull accidents.
- Damage PDFs only consider damage that is significant enough to breach the outer hull. This penalizes structures able to resist rupture.
- Damage extents are treated as independent random variables when they are actually dependent variables, and should be described using a joint PDF.
- The IMO model does not have the ability to take the specifics of an accident scenario into account. Damage extents are sampled independently from the PDF's in Figue E-1.

In 2001, the Marine Board of the National Academy of Science published a report (see Figure E.2) assessing a methodology to compare double hull tanker designs to alternative designs NAS (2001). It too noted that the IMO (1995) model was insufficient for the goals outlined by the NAS (2001) report and that, consequently, further research was necessary: "Given the status of previous efforts to establish a methodology for comparing the environmental performance of alternative tanker designs, the committee concluded that the development of a new approach was warranted - NAS (2001)."

E-2. Developing an oil outflow model

The report NAS (2001) evaluates single hull and double tanker designs for both collisions and groundings. For their purpose they use physical simulations of accident damage inflicted on a tanker as developed by Brown (2001) and Tikka (2001) using the simulation programs SIMCOL resp. DAMAGE. For the Marine Board research, 10,000 collision and grounding scenarios were randomly generated and put through these simulation programs four times; each time using a different tanker design. This resulted in a data set of 40,000 collisions and 40,000 groundings, describing input (i.e. ship speed, displacement, collision angle) and output variables (i.e. damage length, outflow volume). The specific tanker designs that were evaluated by the NAS(2001) report are provided in Figures E-3 and Figures E-4. The goal of having these large data sets was for the NAS(2001) to compare typical outflow performance between single hull and double hull tankers.

While the physical simulation programs SIMCOL resp. DAMAGE by Brown (2001) and Tikka (2001), respectively, were used to develop the input and output data for 40,000 collision scenarios and 40,000 grounding scenarios of single hull and double hull tankers, an evaluation of a single scenario is quite computationally extensive on its own.

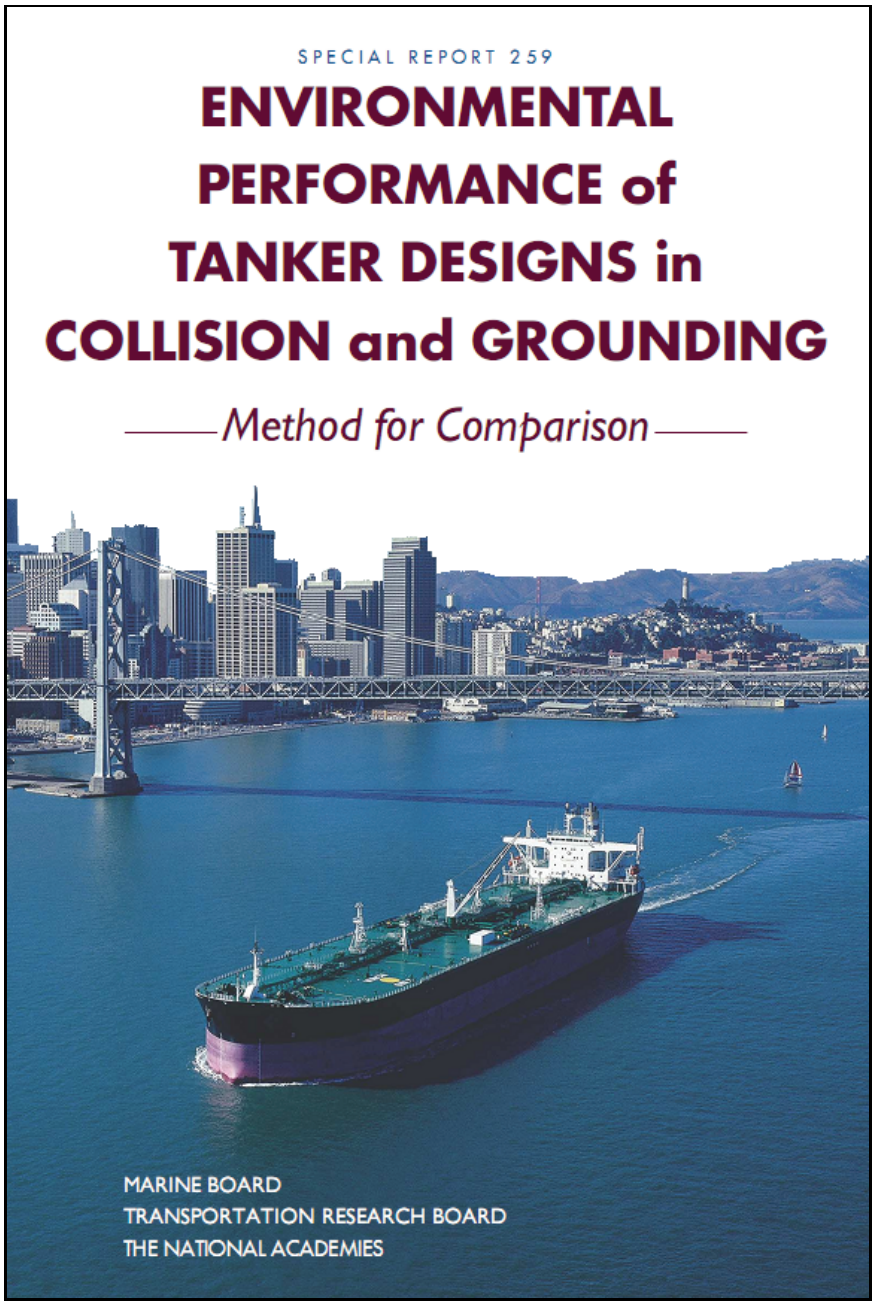


Figure E-2. Cover of Special Report 259 published by the Marine Board, Transportation Research Board, The National Academies.

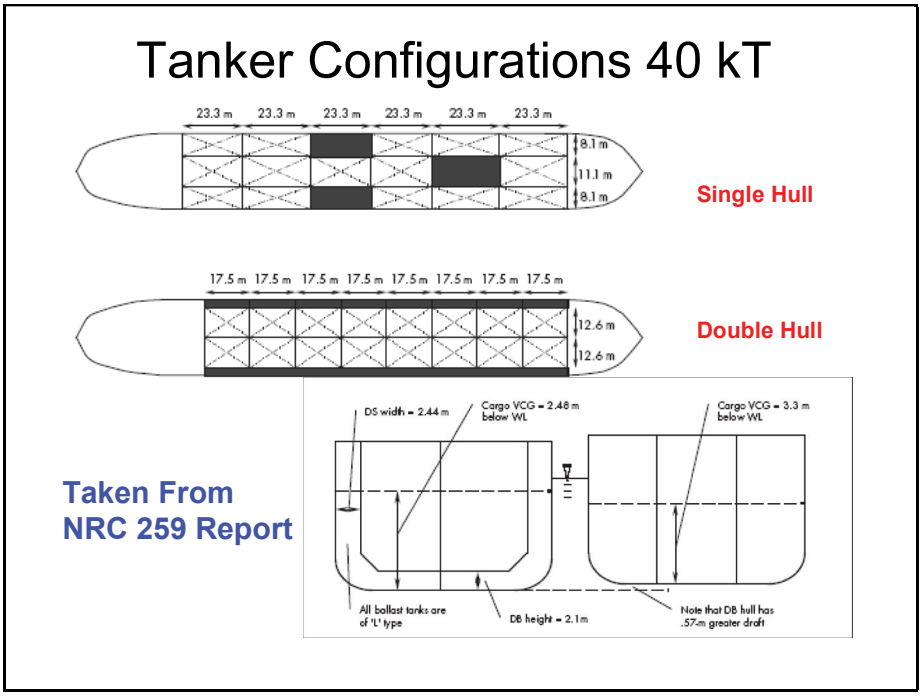


Figure E-3. Tank configurations of 40kT tankers taken from the National Research Council Special Report 259.

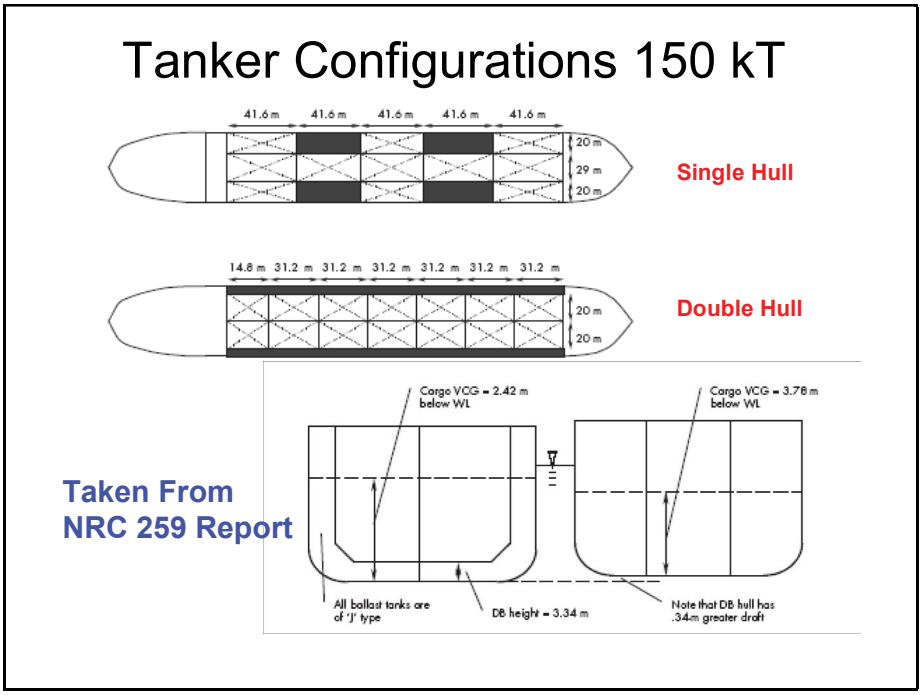


Figure E-4. Tank configurations of 150kT tankers taken from the National Research Council Special Report 259.

Therefore, these software programs at this time do not allow for a seamless integration with such tools as the Vessel Traffic Risk Assessment maritime simulation model that is computationally efficient as well. For example, the calibration scenario VTRA CASE B generates 61,427 vessel to vessel interactions and a future scenario VTRA CASE H has as many as 118,274 vessel to vessel interactions.

However, by carefully studying the relationships between input and output parameters of the large data sets made available through the NAS (2001) report one can "empirically" develop a probabilistic model that determines accident oil outflow based on statistical data analysis techniques rather can computationally intensive physical simulations; one that nevertheless needs to adhere to the same physical principles as the latter. An oil outflow model that explicitly describes the "albeit" statistical relationships between the input parameters and the output parameters can be integrated with the VTRA Maritime Simulation (provided that the simulation records available input data needed to evaluate the oil outflow of collision and grounding scenarios).

Such a model was developed by the Delft University of Technology over the course of this project in close coordination with the George Washington University to ensure a seamless connection between that oil outflow model and the VTRA Maritime Risk Simulation. Its construction is described in detail in the Sub Appendix to this appendix. Chapter 5 in this sub-appendix provides example oil outflow calculations for single hull and double hull tankers for collisions and groundings implemented for this VTRA project.
In this sub-appendix report, twelve accidental outflow models are presented: six collision models and six grounding models: a collision model for the single hull tanker and double hull tankers displayed in Figure E.3 (referred to in the Sub-Appendix as SH40 and DH40), a collision model and grounding model for the single hull and double hull tankers displayed in Figure E.4 (referred to in the Sub-Appendix as SH150 and DH150), a collision model and grounding model that was estimated using all single hull data (referred to as SHCOMB) and, finally, a collision and grounding model using all double hull data (referred to a DHCOMB). The SHCOMB and DHCOMB models allow for an interpolation between the different tankers sized displayed in Figures E-3 and E-4.

These models determine the amount of oil that flows from an oil tanker in case it is struck by another ship or runs aground on a rocky pinnacle. Based on specific scenario data, these models have the ability to evaluate the extent of collision or grounding damage, the probability of rupture and oil spill volume given a set of accident scenario variables. Chapter

5 in this sub-appendix provides example oil outflow calculations for single hull and double hull tankers for collisions and groundings implemented for this VTRA project. Each of these models can be quickly and straightforwardly be implemented in large scale system simulations of tanker movements because they involve formulas using only elementary functions and include an overseeable amount of parameters and coefficients. In short, they combine the power of the physical simulation software of SIMCOL resp. DAMAGE by Brown (2001) and Tikka (2001) with the simplicity of explicit functions. Moreover, these models improve significantly upon the previous IMO model since:

- They are based on a large data set obtained by physically meaningful simulations, rather than a model with simpler assumptions based on a small historic data set;
- They allow for ship size-dependent damage extent and probability of rupture assessments, whereas the old model gave damage and probability independently of ship size;
- Damage extent parameters are dependent on scenario input variables as opposed to independently distributed;
- Damage extent parameters take into account the physical characteristics of the ship designs and accident scenarios, such as speed, mass, collision angle, etc.

While on the outset of our project we set out to evaluate cargo losses from tank vessels that dock at BP Cherry point (referred to hereafter as BP CHPT vessels), we were requested over the course of the project to also consider the potential oil outflow from an interacting vessels when it potentially collides with a BP CHPT Vessel. Specifically, we were requested to separate the total expected oil outflow results by location and size from BP Cherry Point vessels and interacting vessels that potentially collide with them into the following four categories:

- Persistent Expected Oil Outflow results to include crude oil and bunker fuel from BP Cherry Point vessels.
- Persistent Expected Oil Outflow to include crude oil and bunker fuel from interacting vessels.
- Non-Persistent Expected Oil Outflow to include refined products and diesel fuel from BP Cherry Point vessels.
- Non-Persistent Expected Oil Outflow to include refined products and diesel fuel from interacting vessels.

Possibly the need for accounting fuel losses arose from the November 6, 2007 M/V Cosco Busan oil spill in San Francisco Bay. While the models in the sub-appendix were designed specifically for cargo losses from tank vessels, these models offer a flexibility to provide the outflow results above by making some reasonable assumptions.

Given an damage location *a* from mid ship, a damage length *b* and a damage penetration *b*, the model in the Sub-Appendix evaluates the compartments that have been "penetrated" by making two worst case assumptions as depicted in Figure E-5.

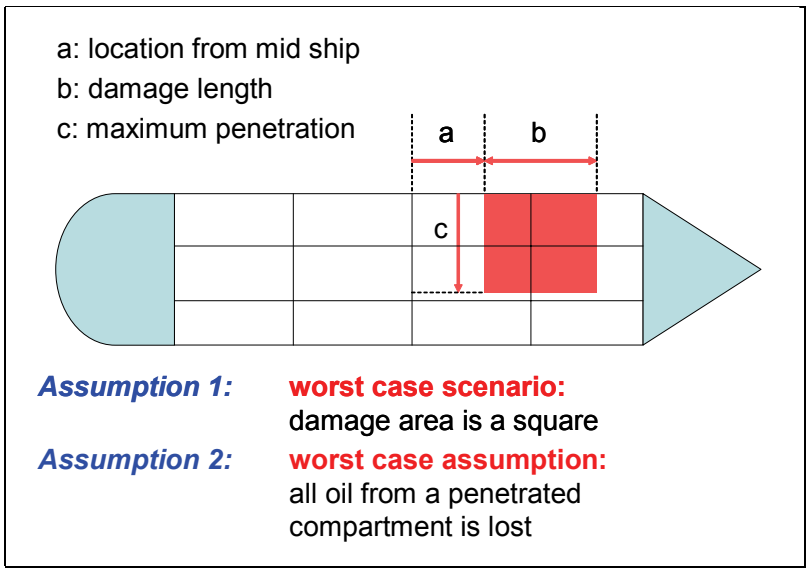


Figure E-5. Worst case assumption of oil outflow volume given a certain damage extent.

Hence, to be able to accomodate diesel fuel and bunker fuel oil outflow calculations for tankers one needs to augment the vessel compartmentalization of Figures E-3 and Figure E-4 with bunker fuel and diesel fuel compartments. While there certainly can be more than two tanks for bunker fuel and two tanks for diesel fuel on a given tanker one could assume (again from a worst case scenario perspective) the following locations for bunker fuel and diesel fuel for tankers as provided in Figure E-6. Note that it follows from Figure E-6 that we continue to provide the double hull tankers the benefit of the double hull for the diesel fuel and bunker fuel compartments. We located the bunker fuel compartments towards the stern (since this is where the main engine compartment is located) and the diesel compartments towards the bow. A reversal of these locations did not seem to make sense given that bunker volumes on deep draft vessels may differ in one order of magnitude (see Section E.2.2). Table E-1 provides the tanker dimensions and the location of the various compartment for

the DH150 tankers design that we used in the VTRA maritime simulation. Similar tables were developed for the SH40, DH40 and SH150 designs.

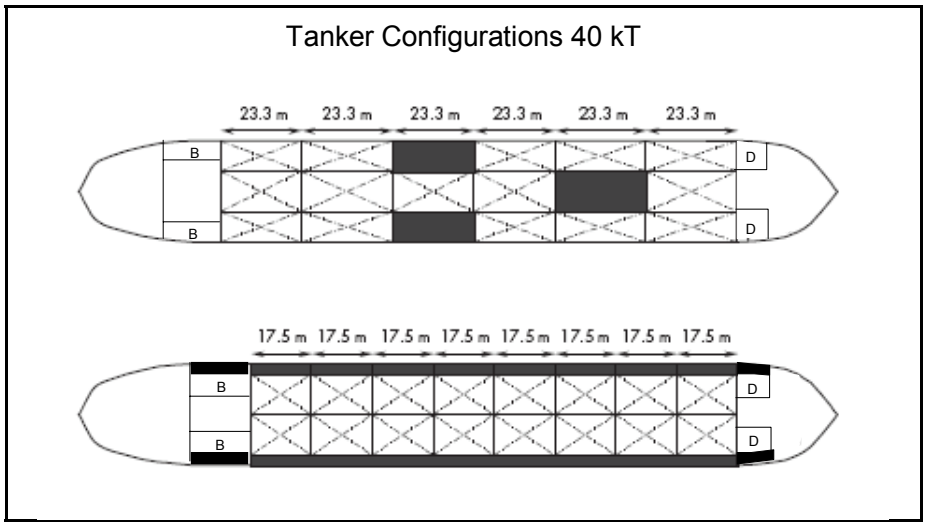


Figure E-6. Worst case assumption locations for bunker fuel tanks and diesel fuel tanks for Tankers.

The capacities provided in Table E-1 are full load capacities for each compartment. The VTRA simulation actually passes to the oil outflow model calculations, whether the tank vessel is carrying product or crude and also the cargo DWT of crude or product that it is carrying. This total capacity is next evenly distributed across the cargo tanks and the tanker is ballasted making the assumption that a tanker is 100% ballasted when the cargo tanks are empty and 0% ballasted when the cargo tanks are complete full (and following a linear relationship in between). Next, using the lightship weight information (also passed by the VTRA maritime simulation) we recalculate the displacement of a partially loaded tanker. The ship's mass (displacement) is one of the required input variables for oil outflow calculations as described in the sub-appendix. Before the above "re balancing" above, the size of a tanker design from the NAS (2001) report and its compartments are rescaled in a linear manner using the ship length and beam also passed to the oil outflow calculation model by the VTRA maritime simulation (but keeping the same format of the compartmentalization of the SH40, DH40, SH150 and DH150 tanker designs). A tank vessel's length is used to evaluate its bunker fuel load and diesel fuel load using a regression model for deep draft vessels (see Section E.2.2).

When restricting oil outflow calculations to those from BP CHPT vessels one could have made a worst case assumption that the BP CHPT vessel was always the stuck vessel in a

vessel interaction. However, with the requirement of evaluating oil outflow from interacting vessels when potentially colliding with a BP CHPT vessel, such an assumption is not reasonable since not both vessels can be "the stuck vessel" in a single vessel interaction scenario at the same time. In the sections below we shall discuss some additional detail regarding the oil outflow model described in the Sub-Appendix, describe a striking-struck ship model and regression models to relate a vessels lengths to its fuel carrying capacity.

Table E-1. Example of modeled tank locations, dimensions and capacities for a 150kT double hull tanker

	X	Y	Z	Length	Width	Capacity (m3)	Content
1	0.00	0.00	3.34	12.30	3.34	710.91	Empty
2	8.29	3.34	3.34	4.01	7.06	159.94	Diesel
3	8.29	39.60	3.34	4.01	7.06	159.94	Diesel
4	0.00	46.66	3.34	12.30	3.34	710.91	Empty
5	12.30	0.00	3.34	31.20	3.34	1803.28	Ballast
6	12.30	3.34	3.34	31.20	21.66	11694.30	Crude
7	12.30	25.00	3.34	31.20	21.66	11694.30	Crude
8	12.30	46.66	3.34	31.20	3.34	1803.28	Ballast
9	43.50	0.00	3.34	31.20	3.34	2262.78	Ballast
10	43.50	3.34	3.34	31.20	21.66	14674.20	Crude
11	43.50	25.00	3.34	31.20	21.66	14674.20	Crude
12	43.50	46.66	3.34	31.20	3.34	2262.78	Ballast
13	74.70	0.00	3.34	31.20	3.34	2259.11	Ballast
14	74.70	3.34	3.34	31.20	21.66	14650.40	Crude
15	74.70	25.00	3.34	31.20	21.66	14650.40	Crude
16	74.70	46.66	3.34	31.20	3.34	2259.11	Ballast
17	105.90	0.00	3.34	31.20	3.34	2259.23	Ballast
18	105.90	3.34	3.34	31.20	21.66	14651.20	Crude
19	105.90	25.00	3.34	31.20	21.66	14651.20	Crude
20	105.90	46.66	3.34	31.20	3.34	2259.23	Ballast
21	137.10	0.00	3.34	31.20	3.34	2259.17	Ballast
22	137.10	3.34	3.34	31.20	21.66	14650.80	Crude
23	137.10	25.00	3.34	31.20	21.66	14650.80	Crude
24	137.10	46.66	3.34	31.20	3.34	2259.17	Ballast
25	168.3	0	3.34	31.2	3.34	2137.522899	Ballast
26	168.30	3.34	3.34	31.20	21.66	13861.90	Crude
27	168.30	25.00	3.34	31.20	21.66	13861.90	Crude
28	168.30	46.66	3.34	31.20	3.34	2137.52	Ballast
29	199.50	0.00	3.34	14.80	3.34	850.37	Ballast
30	199.50	3.34	3.34	14.80	21.66	5514.70	Crude
31	199.50	25.00	3.34	14.80	21.66	5514.70	Crude
32	199.50	46.66	3.34	14.80	3.34	850.37	Ballast
33	214.30	0.00	3.34	52.00	3.34	2987.80	Empty
34	214.30	3.34	3.34	26.79	11.16	2649.93	Heavy Fuel
35	214.30	35.50	3.34	26.79	11.16	2649.93	Heavy Fuel
36	214.30	46.66	3.34	52.00	3.34	2987.80	Ballast

E-2.1. Description of scenario data obtained from the NRC Oil outflow Report

A complete description of the scenario data is provided in the Sub-Appendix, Chapter 2. Tables E-2 and E-3 provide an informal description of those input variables and output

variables from the NAS(2001) report that we were able to link directly within the VTRA maritime simulation. Figure E-5 depicts the input scenario information from Table E-2 graphically for a particular example collision scenario. With the exception of the damage location input variables listed in Tables E-2 and E-3, these input variables are recorded directly into the recording databases from the VTRA simulation. Such a recording was also necessary for the accident attributes of the accident probability models for collisions and groundings described in Appendix D.

Table E-2. Input variables and output results for the collision oil outflow model in the VTRA maritime simulation.

Input Variables	Output Variables
Striking ship velocity	Damage length
Struck ship velocity	Maximum penetration
Collision angle	Oil outflow volume
Displacement of Striking Vessel	
Displacement of Struck Vessel	
Collision location, relative from stern	
Striking ship type	

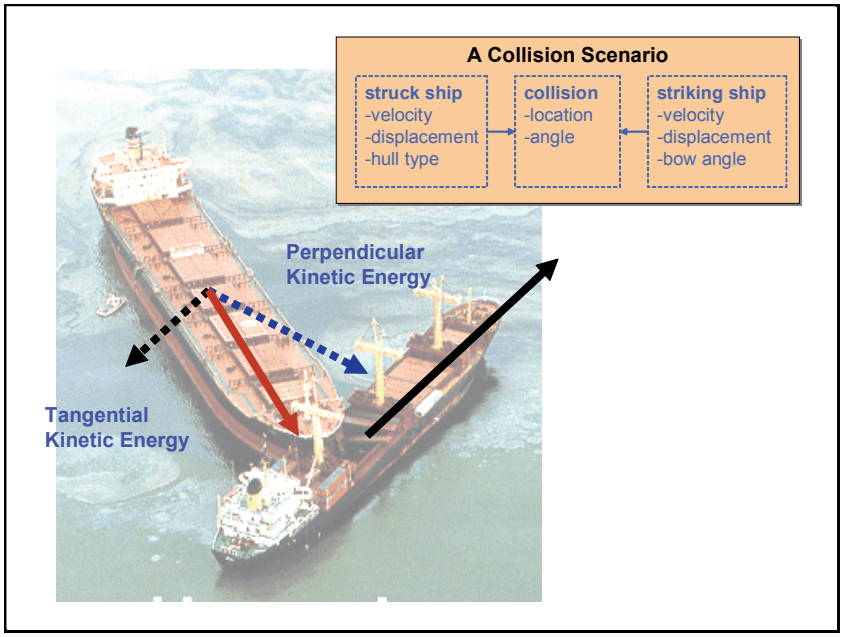


Figure E-7. A schematic of a sticking ship-struck ship collision scenario

Table E-3. Input variables and output results for the grounding oil outflow model in the VTRA maritime simulation.

Input Variables	Output Variables
Ship velocity	Begin Damage length
Displacement of Vessel	End Damage Length
Damage Location from Mid-Ship	Damage Width
	Damage Height
	Outflow Volume

Our VTRA Maritime simulation does not simulate the circumstances and movements of vessels immediately preceding accidents and as a result we cannot record the exact location of a collision or the location of the grounding relative to the dimensions of the vessel. Hence, instead we evaluate the oil outflow distributed over 100 discrete points across a vessel length for collisions and across a vessels half-width for groundings and we evaluate the average oil outflow per collision or per grounding across all these different locations.

The models constructed in the sub-appendix allow for an interpolation between the tanker sizes depicted in Figures E-3 and E-4. To that end, we converted the input variables in Tables E-2 and E-3 to ones that relate to a kinetic energy interpretation. For example, striking ship velocity, struck ship velocity, striking ship displacement and stuck ship displacement are converted into a tangential and perpendicular kinetic energies (see Figure E-7) which are then in turn related to damage length and damage penetration calculations.

As noted previously, in all scenarios in the NAS(2001) report the tanker is assumed to be the struck vessel (contrary to the example photo in Figure E-7.). The next section discusses the striking ship-struck ship model that we developed to account for the possibility that in fact the tank vessel is the striking vessel (as depicted in Figure E-7).

E-2-2. Striking and struck ship model

In the event of two identical ships crossing each others paths at a 90 degree angle traveling at exactly the same speeds, it would be reasonable to assume that their would be a 50-50 chance that either one would the struck or striking vessels. However, this assumption becomes less reasonable when their is a large speed differential or if their ship dimensions are much different. Take, for example, an interaction between a tanker and a recreational vessel. Simply from the point of size it would seem much easier to actually strike the tanker

than striking the recreational vessel. Who strikes and who is stuck has implications with respect to the oil outflow that one evaluates for such a collision scenario. We have developed a conditional probability model that evaluates a probability that either Vessel 1 or Vessel 2 is the struck vessel (given that a collision is about to occur between these two vessels). Needless to say, these two conditional probabilities need to sum up to 1 in that case by definition.

Figure E-6 provides a schematic and a geographic explanation of this striking-stuck ship model. Let L_1, w_1 and v_1 be the length, width and traveling speed of the first vessel. Let L_2, w_2 and v_2 be the length, width and traveling speed of the second vessel and let Φ be the angle of the crossing paths of these two vessels. From these parameters we first evaluate the distance that Vessel 1 is exposed to the potential of a collision which follows as

$$L_{1+}\frac{w_2}{\sin\phi}. \qquad \text{(E-1)}$$

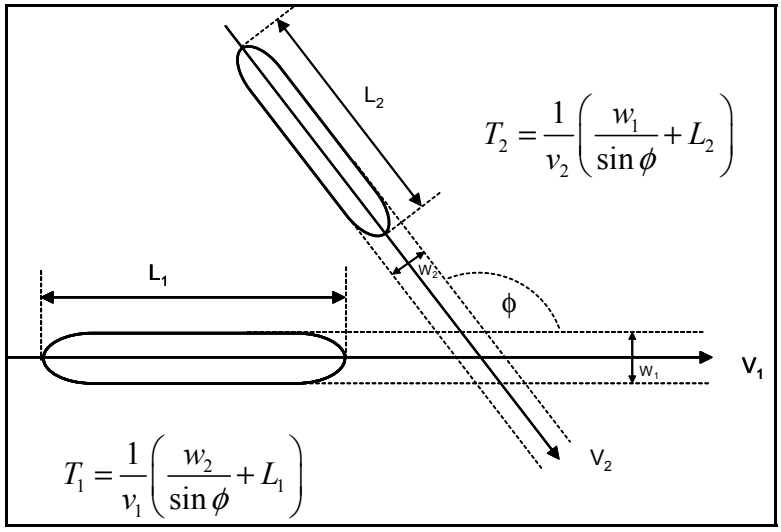


Figure E-8. A schematic of a sticking ship-struck ship probability model.

Dividing this distance by the Vessel 1 speed v_1 yields the length of time T_1 that Vessel 1 is exposed to the potential of a collision given the angle Φ of the tracks of Vessel 1 and Vessel 2 and the width of Vessel 2:

$$T_1 = \frac{1}{v_1}\left(L_{1+}\frac{w_2}{\sin\phi}\right). \qquad \text{(E-2)}$$

Using a symmetry argument we evaluate for the length of time T_2 that Vessel 2 is exposed as:

$$T_2 = \frac{1}{v_2}\left(L_{2+}\frac{w_1}{\sin\phi}\right). \quad \text{(E-3)}$$

Next, we set:

$$Pr(\text{Vessel 1 is struck}) = \frac{T_1}{T_1 + T_2} \text{ and } Pr(\text{Vessel 2 is struck}) = \frac{T_2}{T_1 + T_2}. \quad \text{(E-4)}$$

From expression (E-4) we evaluate that two identical vessels traveling at the same speeds and crossing paths at a 90 degree angles indeed have a 50-50 chance of being the struck vessel. On the other hand we evaluate from (E-4), for example, that a DH150 tanker (with a length 266.3 meters and width of 50 meters) traveling at 8 knots crossing the path of a tug (with a length of 34 meters and a width of 12 meters) traveling at 12 knots at a ϕ=135 degree angle, has approximately an 80% probability of being struck. Hence, in that scenario the tug has approximately a 20% probability of being struck.

In the VTRA maritime simulation, the loss of oil from a struck vessel is weighted by the probability of the vessel being struck evaluated using expression (E-4). It is further assumed that no vessel fuel or oil cargo products is lost from the striking vessel. In the case of a traffic scenario that a small vessel is the struck vessel (in the sense that the length of the smaller vessel is less than or equal the width of the larger vessel) all diesel fuel on board of the smaller vessel is assumed lost. Otherwise the oil outflow models in the sub-appendix are used to evaluate damage length and penetration to determine those cargo or fuel tanks that are penetrated. For non-tankers the single hull parameters settings are used from the sub-appendix to evaluate these damage extents.

E-2.3. Bunker fuel and diesel fuel regression models

The vessels considered in the VTRA range in size and utility between Cherry Point oriented tankers and sailing regattas. The fuel or fuel oil capacities of these vessels are as diverse as the vessels themselves. In order to include diesel fuel and bunker fuel in the outflow models of VTRA maritime simulation multiple sources have been queried in order to develop model fuel oil capacities as a function of size and utility of the vessel being considered. This section outlines the sources of the data queried and the regression models that have been fitted to estimate a vessel's fuel capacity as a function of a vessel's length.

The fuel oil capacities for Cherry Point tankers are source in Vessels Particular Questionnaires (VP's) for each tanker that has made calls at the Cherry Point Facility. The

VP for each vessel offered the fuel oil type and volume capacity of each fuel oil tank. Bunker and diesel fuel vessel for other deep draft vessels in the VTRA maritime simulation were compiled from various regional and global vessel brokerage firm's web-sites (e.g. http://www.ship-technology.com) as well as from the publication Taggart (1980). The data from these data sources were combined to generate the scatter plots in Figure E-9.

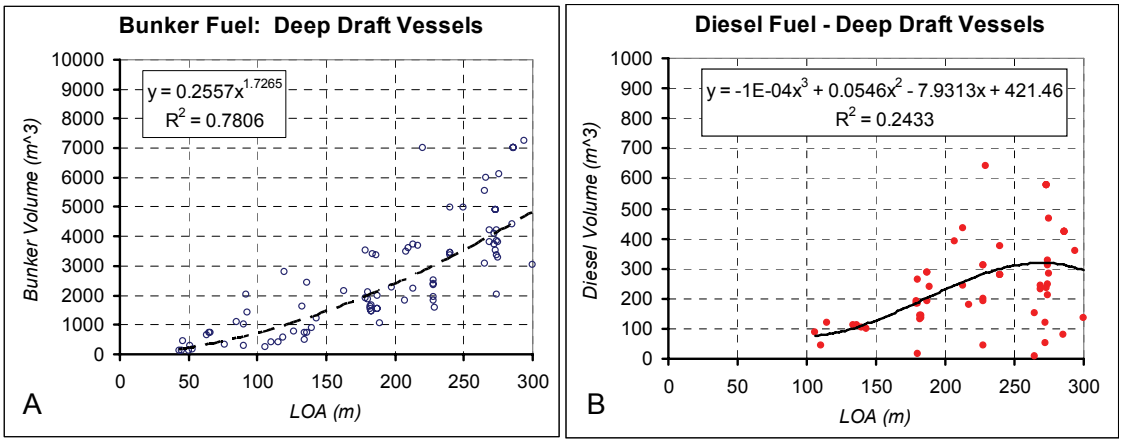


Figure E-9. Deep draft vessel fuel data and least squares regression fits, A: Scatter plot of bunker fuel volume by vessel length, B: Scatter plots of diesel fuel by vessel length.

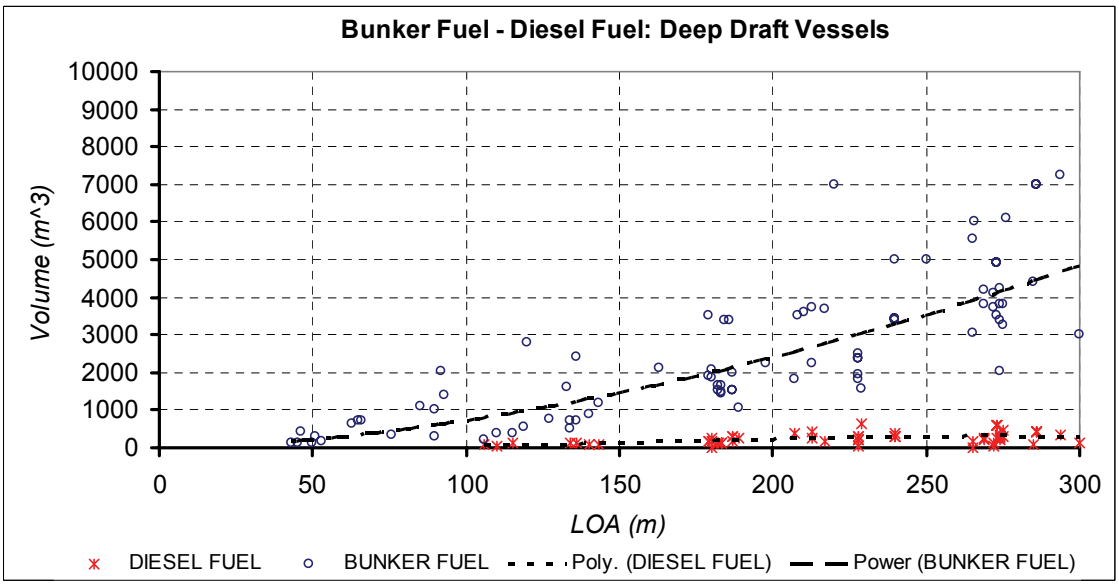


Figure E-10. Scatter plots of deep draft vessel bunker and diesel fuel data and least squares regression fits in a single plot.

Figure E-9A provides a scatter plot of the accumulated bunker fuel data and Figure E-9B provides a scatter plot of the accumulated diesel fuel data for deep draft vessels. Please note that the scale of the y-axis in Figure E-9A is one order of magnitude higher than that of Figure 9B. This becomes more apparent when combining both scatter plots in a single plot in Figure E-10. Figures 9A and 9B contains the equations of the regressions fits linking bunker and diesel fuel to a vessel's length, respectively. Note that the R^2 of 78% value for the bunker fuel is quite respectable, whereas the R^2 value of the diesel fuel is quite low. From Figure E-10 it follows that this lack-of-fit will be masked by the amount of bunker fuel on a tank vessel of a particular length.

The same locations for the bunker fuel tanks and the diesel fuel tanks given in Figure E-6 were assumed for other deep draft vessels than tankers. The parameters of the single hull damage models in the Sub-appendix were used to evaluate damage length and damage penetration for these deep draft vessels after which the analysis exemplified by Figure E-5 was used to determine if these fuel tanks were penetrated. If penetrated, all bunker fuel or diesel fuel in a penetrated tanks was assumed lost.

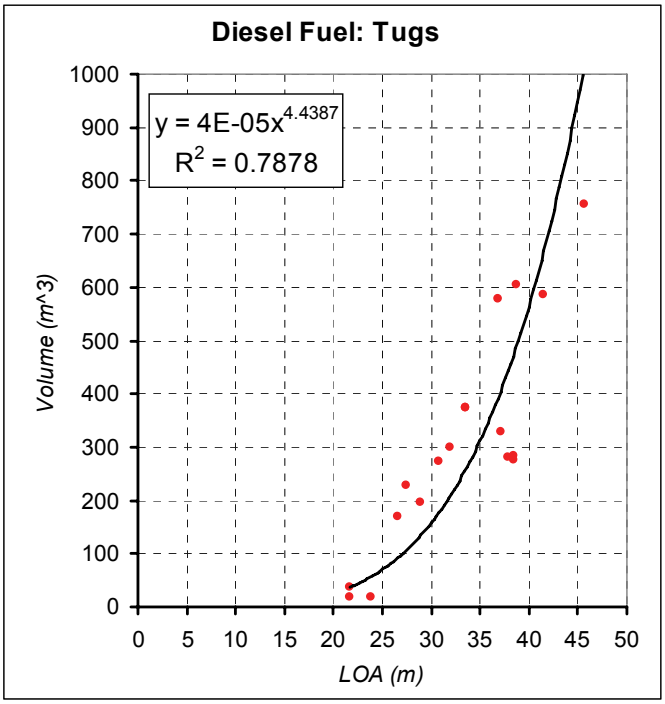


Figure E-11. Scatter plot and least squares regression fit of diesel fuel data for tugs by vessel length.

The fuel oil capacities for Cherry Point oriented ITB's and ATB's are sourced in the VPQ's as well. This information was combined with tug diesel capacity data from more general vessel design sources, specifically, specific vessel schematics made available through the web sites of various vessel brokerage firms, Tug operating companies in the Puget Sound region and Taggert (1980). The resulting scatter plot and regression fit (with an R^2 of about 79%) linking the length of tugs with their diesel carrying capacity are displayed in Figure E-11 above.

Scaled SH40 and DH40 tanker compartmentalizations were also assumed to model the oil outflow from ATB's and ITB's with the exception that the bunker fuel at the stern was replaced with diesel fuel with a carrying capacity determined by the length of the tug and the regression equation in Figure E-11. In the event of a light tug (i.e. a tug traveling by itself without a barge and given that the length of tug is typically smaller than the width of a BP CHPT tank vessel) all diesel fuel from a tug was assumed lost in the event it is the struck vessel. Indeed, DH150 tankers have a width of 50 meters (see Table E-1) whereas the upper bound of the scatter plot E-11 is 50 meters.

In the case that a vessel interaction occurred between a BP CHPT vessel and an oil barge being towed, we accounted for the potential oil loss from the oil barge. There are many different sizes of oil barges that are used within the Puget Sound area. We made a worst case assumption and used the configuration of one of the larger oil product barges depicted in Figure E-12 combined with the single hull oil outflow parameters from the sub-appendix. We modeled the tank locations of the oil barge as per Table E-4. Hence, we evaluated damage lengths and penetration following the oil outflow model in the sub-appendix and used the analysis exemplified in Figure E-5 to evaluate the tanks that were penetrated. All petroleum products from penetrated tanks were assumed lost.

One of the larger participants of the VTRA study area are the Washington State Ferries. Over the course of this project we have requested vessel rides on the Washington States Ferries given their unique distribution of their routes across the VTRA Study area and regular schedule. On every occasion we found the Washington State Ferry system to particularly accommodating and we are obliged for their assistance. When requesting the dimensions of the various WSF's in the system, their diesel fuel carrying capacity and the approximate locations of the fuel tanks we once again found the Washington State Ferries management to very responsive and our data request was honored in a matter of two weeks. We would like to thank the Washington State Ferry system for their participation as they too

(similar to the experts that participated in the expert judgment elicitation described in Appendix B) had no benefit to participating in this study other than that it possibly could enhance the safety of the waterway within the VTRA study area.

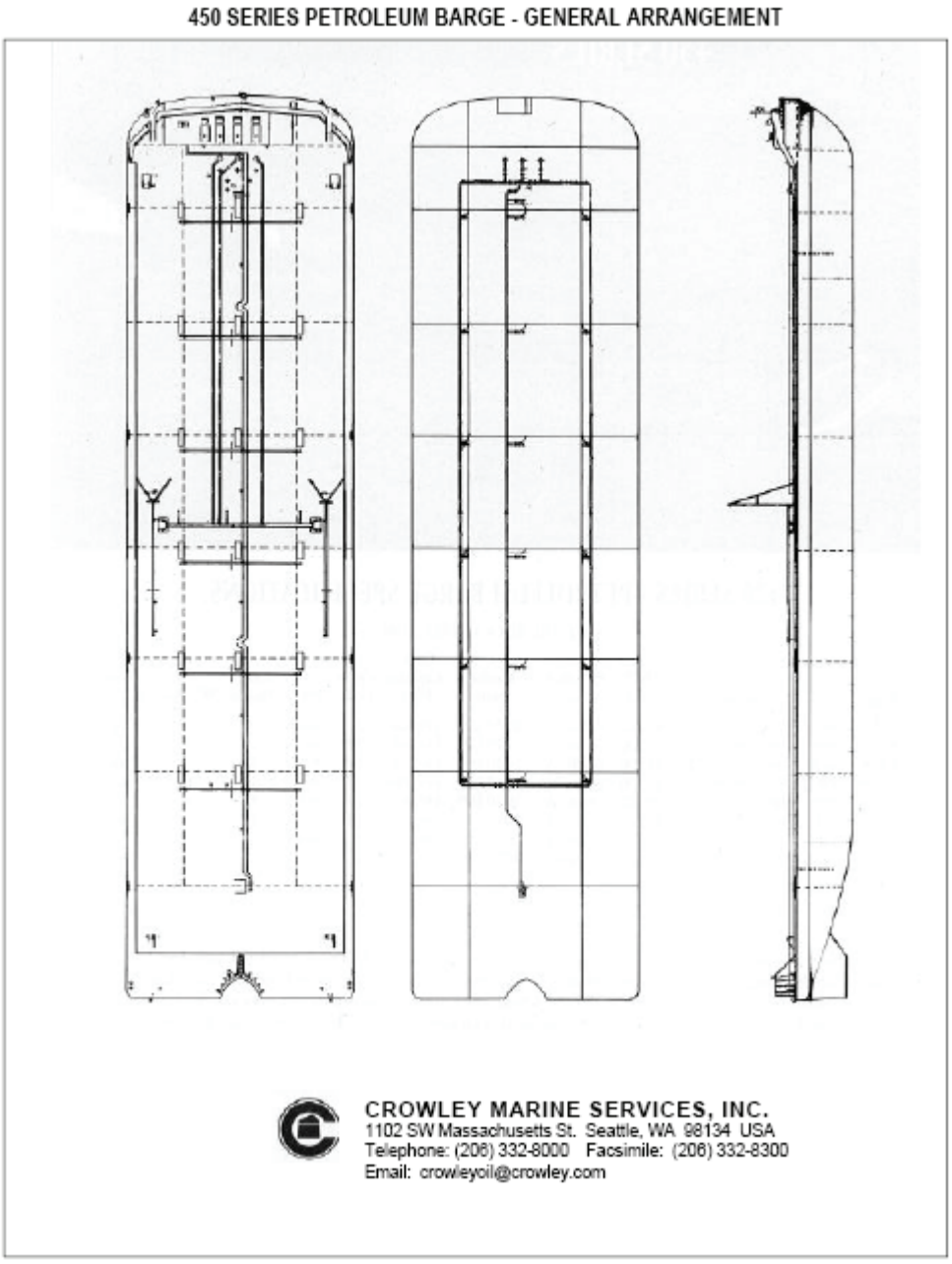


Figure E-12. A 450 Series petroleum barge.

The information that we received from the Washington Ferries System regarding the dimensions of their vessels and the approximate location of the fuel tanks in their vessels are

Table E-4. Example of modeled tank locations, dimensions and capacities for a worst case oil barge.

	X	Y	Z	Length	Width	Capacity (m3)	Content
1	0.00	0.00	0.00	7.62	30.30	0.00	Empty
2	7.62	0.00	0.00	7.62	7.58	473.82	Product
3	15.24	0.00	0.00	15.24	7.58	947.64	Product
4	30.48	0.00	0.00	15.24	7.58	947.64	Product
5	45.72	0.00	0.00	15.24	7.58	947.64	Product
6	60.96	0.00	0.00	15.24	7.58	947.64	Product
7	76.20	0.00	0.00	15.24	7.58	947.64	Product
8	91.44	0.00	0.00	15.24	7.58	947.64	Product
9	7.62	12.93	0.00	4.44	4.44	87.32	Diesel
10	15.24	7.58	0.00	15.24	15.15	1895.28	Product
11	30.48	7.58	0.00	15.24	15.15	1895.28	Product
12	45.72	7.58	0.00	15.24	15.15	1895.28	Product
13	60.96	7.58	0.00	15.24	15.15	1895.28	Product
14	76.20	7.58	0.00	15.24	15.15	1895.28	Product
15	91.44	7.58	0.00	15.24	15.15	1895.28	Product
16	7.62	22.73	0.00	7.62	15.15	473.82	Product
17	15.24	22.73	0.00	15.24	7.58	947.64	Product
18	30.48	7.58	0.00	15.24	7.58	947.64	Product
19	45.72	0.00	0.00	15.24	7.58	947.64	Product
20	60.96	0.00	0.00	15.24	7.58	947.64	Product
21	76.20	0.00	0.00	15.24	7.58	947.64	Product
22	91.44	0.00	0.00	15.24	7.58	947.64	Product
23	106.68	0.00	0.00	15.24	30.30	0.00	Empty

Table E-5. Vessel dimension of Washington State Ferries.

WSF Ferry	Class	Length	Beam	Draft	Speed	Displacement (Mtons)
Puyallup	Jumbo Mark II	460'2"	90'	17'3"	18	10690
Tacome	Jumbo Mark II	460'2"	90'	17'3"	18	10690
Wenatchee	Jumbo Mark II	460'2"	90'	17'3"	18	10690
Spokane	Jumbo	440"	87"	16'	18	9913
Walla Walla	Jumbo	440"	87"	16'	18	9913
Elwha	Super	382'2"	73'2"	18'9"	20	8005
Hyak	Super	382'2"	73'2"	18'9"	17	8005
Kaleetan	Super	382'2"	73'2"	18'9"	17	8005
Yakima	Super	382'2"	73'2"	18'9"	17	8005
Cathlamet	Issaquah 130	328'	78'8"	16'6"	16	6234
Chelan	Issaquah 130	328'	78'8"	16'6"	16	6234
Issaquah	Issaquah 130	328'	78'8"	16'6"	16	6234
Kitsap	Issaquah 130	328'	78'8"	16'6"	16	6234
Kittitas	Issaquah 130	328'	78'8"	16'6"	16	6234
Sealth	Issaquah 100	328'	78'8"	15'6"	16	6234
Evergreen State	Evergreen	310'	73'2"	15'10"	13	5466
Klahowya	Evergreen	310'	73'2"	15'10"	13	5466
Tillikum	Evergreen	310'	73'2"	15'10"	13	5466
Illahee	Steel Electric	256'2"	73'10"	12'9"	12	3550
Klickitat	Steel Electric	256'2"	73'10"	12'9"	12	3550
Nisqually	Steel Electric	256'2"	73'10"	12'9"	12	3550
Quinault	Steel Electric	256'2"	73'10"	12'9"	12	3550
Rhodondendron	Rhodondendron	227'6"	62'	10'	11	2423
Hiyu	Hiyu	162'	63'1"	11'3"	10	2043
Kalama	POV	112'	25"	8'	25	508
Skagit	POV	112'	25"	8'	25	508

Table E-6. Approximate fuel tank locations and capacities for WSF's.

WSF Ferry	Class	Total Fuel Capacity (in Gallons)	Number of Fuel Tanks	Location Fuel Tank (Mid-Ship, Starboard, Port)	Approximate length Fuel Tank	Approximate width Fuel Tank
Puyallup	Jumbo Mark II	110385	2	#1 Centerline #2 Centerline	37	30
Tacome	Jumbo Mark II	110385	2	#1 Centerline #2 Centerline	37	30
Wenatchee	Jumbo Mark II	110385	2	#1 Centerline #2 Centerline	37	30
Spokane	Jumbo	125000	2	#1 Centerline #2 Centerline	40	35
Walla Walla	Jumbo	125000	2	#1 Centerline #2 Centerline	40	35
Elwha	Super	62372	3	Port Center STB (MID)	27	24
Hyak	Super	77683	3	Port Center STB (MID)	27	24
Kaleetan	Super	77683	3	Port Center STB (MID)	27	24
Yakima	Super	77683	3	Port Center STB (MID)	27	24
Cathlamet	Issaquah 130	115400	4	Wing Port, Deep Port Deep STB Wing STB (MID)	1&4 -- 13'6" 2&3 -- 27'	#2&3 fuel oil tks. 22'-6"W, 1&4 fuel oil tks.14'-0"W
Chelan	Issaquah 130	115400	4	Wing Port, Deep Port Deep STB Wing STB (MID)	1&4 -- 13'6" 2&3 -- 27'	#2&3 fuel oil tks. 22'-6"W, 1&4 fuel oil tks.14'-0"W
Issaquah	Issaquah 130	115400	4	Wing Port, Deep Port Deep STB Wing STB (MID)	1&4 -- 13'6" 2&3 -- 27'	#2&3 fuel oil tks. 22'-6"W, 1&4 fuel oil tks.14'-0"W
Kitsap	Issaquah 130	115400	4	Wing Port, Deep Port Deep STB Wing STB (MID)	1&4 -- 13'6" 2&3 -- 27'	#2&3 fuel oil tks. 22'-6"W, 1&4 fuel oil tks.14'-0"W
Kittitas	Issaquah 130	115400	4	Wing Port, Deep Port Deep STB Wing STB (MID)	1&4 -- 13'6" 2&3 -- 27'	#2&3 fuel oil tks. 22'-6"W, 1&4 fuel oil tks.14'-0"W
Sealth	Issaquah 100	115400	4	Wing Port, Deep Port Deep STB Wing STB (MID)	1&4 -- 13'6" 2&3 -- 27'	#2&3 fuel oil tks. 22'-6"W, 1&4 fuel oil tks.14'-0"W
Evergreen State	Evergreen	30600	2	Port STB (MID)	13.5	14
Klahowya	Evergreen	30600	2	Port STB (MID)	13.5	14
Tillikum	Evergreen	30600	2	Port STB (MID)	13.5	14
Illahee	Steel Electric	9000	2	Port STB (MID)	12	6' Diameter
Klickitat	Steel Electric	9000	2	Port STB (MID)	12	6' Diameter
Nisqually	Steel Electric	9000	2	Port STB (MID)	12	6' Diameter
Quinault	Steel Electric	9000	2	Port STB (MID)	12	6' Diameter
Rhodondendron	Rhodondendron	11397	2	Center Line #1 end #2 end	20	12'
Hiyu	Hiyu	10000	2	Port STB #1 end	12'	NA
Kalama	POV	6714	2	Port STB (MID)	6	6
Skagit	POV	6714	2	Port STB (MID)	6	6

Table E-7. Example of modeled fuel tank locations of a Jumbo ferry.

	X	Y	Z	Length	Width	Capacity (m3)	Content
1	54.86	7.92	0.00	12.19	10.67	236.64	Diesel
2	67.06	7.92	0.00	12.19	10.67	236.64	Diesel

Table E-8. Example of modeled fuel tank locations of an Issaquah ferry.

	X	Y	Z	Length	Width	Capacity (m3)	Content
1	47.93	0.86	0.00	4.11	4.27	51.84	Diesel
2	45.87	5.13	0.00	8.23	6.86	166.63	Diesel
3	45.87	11.99	0.00	8.23	6.86	166.63	Diesel
4	47.93	18.85	0.00	4.11	4.27	51.84	Diesel

summarized in Tables E-5 and Tables E-6. This information was used to develop the the locations of the fuel tanks within a ferry for the purposes of oil outflow calculation as per the model described in the sub-appendix. Here too, we used the single hull parameters settings for the evaluation of oil outflow from WSF's. As examples, Tables E-7 and E-8 provide, respectively, our modeled locations of the two fuel tanks on a Jumbo Ferry (which has the larges fuel carrying capacity) and the four fuel tanks of Issaquah Ferry.

Figures E-13, E-14 and E-15 provides additional scatter plots of collected data and regressions fits linking vessel lengths to diesel fuel carrying capacity for, respectively, fishing

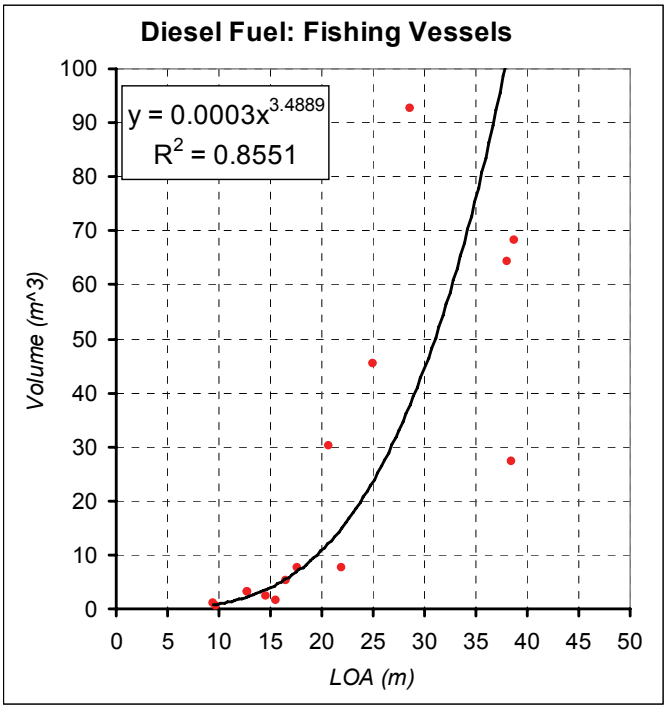


Figure E-13. Scatter plot and least squares regression fit of diesel fuel data for fishing vessels by vessel length.

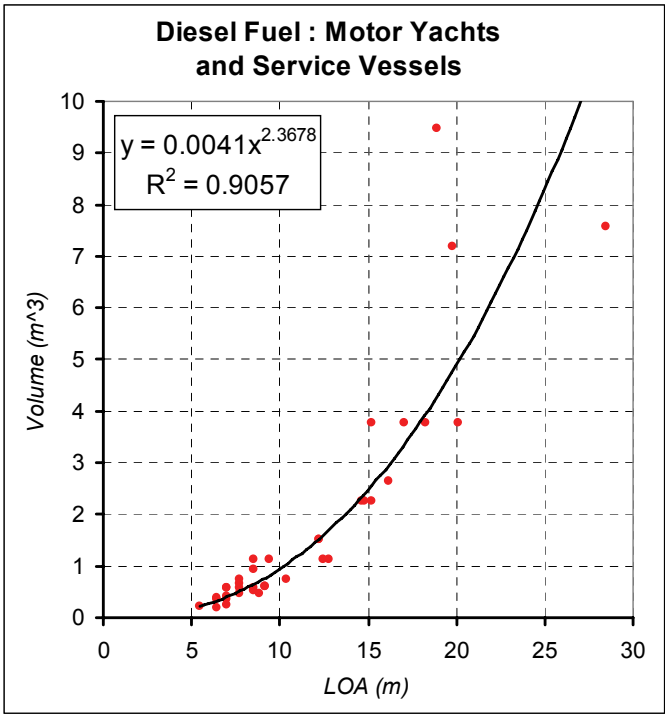


Figure E-14. Scatter plot and least squares regression fit of diesel fuel data for motor yachts and service vessels by vessel length.

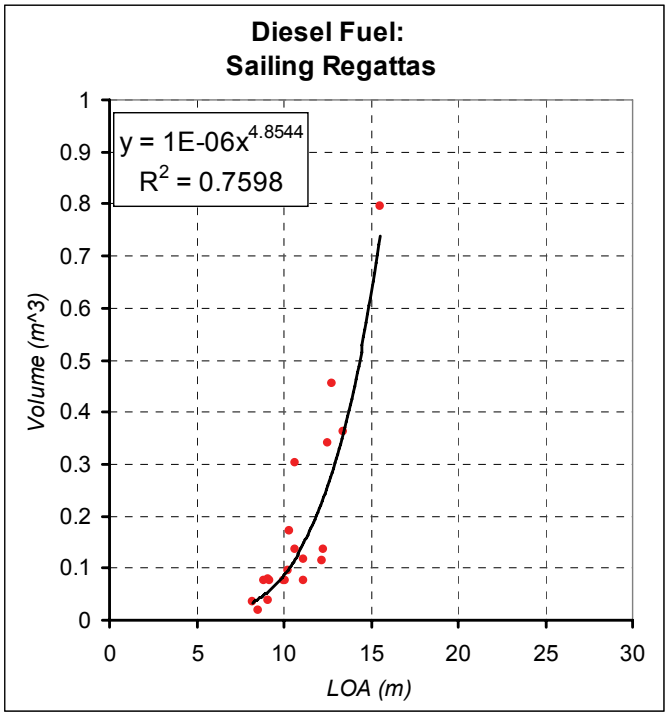


Figure E-15. Scatter plot and least squares regression fit of diesel fuel data for sailing yachts by vessel length.

vessels, motor yachts and service vessels and sailing regattas. The data for Figures E-13 through E-15 were compiled through the web sites of various regional vessel brokerage firms (see, e.g., http://www.yachts.com). The R^2 values for these regressions fits are in order 86%, 90% and 76%, which are all quite high. Please observe that in going from Figure E-13 to Figure E-15 the order of magnitudes of the y-axis goes down by 1 each time. The order of magnitude of the y-axis in Figure E-13 for fishing vessels is in turn one less than that of the y-axis in Figure E-12 for tugs. Finally, the order of magnitude of the y-axis in Figure E-13 for tugs vessels is one less than that of the y-axis in Figure E-11 for deep draft vessels. Moreover, whereas Figures E-12 through E-15 relate to diesel fuel, Figure E-11 relates to both bunker (heavy) fuel and diesel fuel.

E-3. Representative results from the oil outflow model

Similar to the recording of accident attributes for the accident probability models in Appendix D, the parameters for the oil outflow calculation are recorded by the VTRA maritime simulation program. Figure E-16 displays a screen shot of this recording process for the transit of the vessel of interest 134 identified in Figure E-16. The colored cells indicate the vessel interactions that have occurred thus far during its transit, while the

database on the lower left corner shows the recording of the specific accident attributes and input parameters for the oil outflow models during these vessel interactions. The oil outflow model in the sub-appendix together with its augmentations described in this appendix above are used to evaluate the oil outflow in terms of crude oil, petroleum products, heavy fuel and diesel fuel. Next, the crude oil and heavy-fuel outflows are combined into the category "persistent oil" and the petroleum (refined) products and diesel fuel are combined into the category "non-persistent oil". In addition, our analysis is able to separate these later two categories in terms of the originating sources BP CHPT vessels and interacting vessels that potentially collide with a BP CHPT vessels. Table E-9 and E-10 summarize the aggregate annual average oil outflow results that we have analyzed for calibration VTRA Case B.

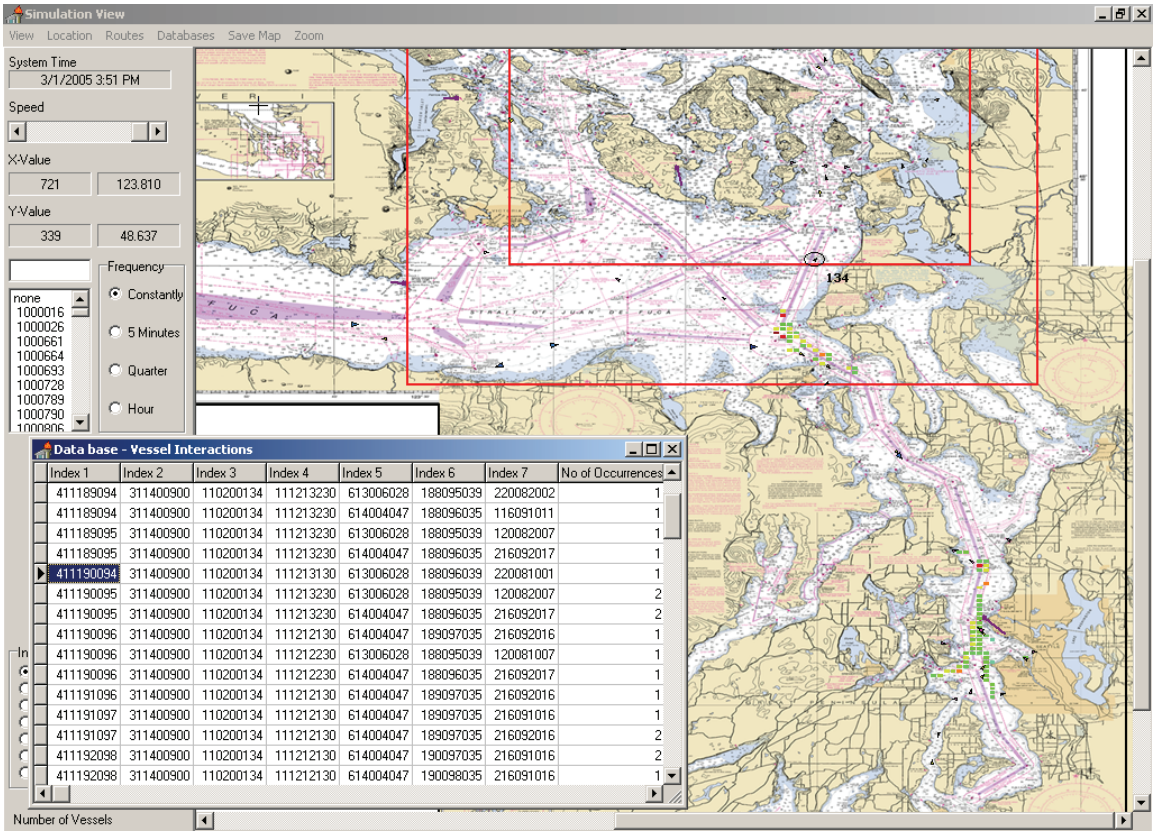


Figure E-16. Encoding of interactions by the VTRA maritime simulation.

From Table E-10 we observe that about 33% of the overall average yearly oil outflow for the calibration VTRA Case B can be attributed to collisions, 62% to powered groundings, 4% to drift grounding and 1% to allisions. Moreover, 97.5% can be attributed to the BP CHPT vessels and only 2.5% to the interacting vessels that potentially collide with BP CHPT Vessel. It is important to point out here that this study was only to consider the oil outflow

from BP Cherry Point vessels and those that potentially collide with them. Hence, of the total 33% attributed to collisions, 30.5% originated from the BP CHPT vessels, which perhaps should not be a surprise given that the interacting vessels are not necessarily tank vessels and hence carry much less oil. Finally, of the total annual average oil outflow we evaluate that 87.3% is persistent oil and 12.7% non-persistent. While these percentages are of interest by themselves, of at least an equal interest would be the comparison of theseoil outflow across the different VTRA Cases. This is not a topic of this appendix, but is described in the main report and Appendix G.

Table E-9. Average oil outflows per year by accident type for the calibration VTRA Case B (amounts are in cubic meters)

	Collisions	Powered Grounding	Drift Grounding	Allisions	Total Oil Outflow
BP CHPT Persistent	31.2	84.5	5.3	1.1	122.1
BP CHPT Non-Persistent	12.2	2.8	0.2	0.1	15.3
IV Persistent	1.0	N/A	N/A	N/A	1.0
IV Non - Persistent	2.6	N/A	N/A	N/A	2.6
Total Oil Outflow	47.0	87.3	5.5	1.2	141.0

Table E-10. Percentages of average oil outflows per year by accident type for the calibration VTRA Case B (% of total average oil outflows)

	Collisions	Powered Grounding	Drift Grounding	Allisions	Total Oil Outflow
BP CHPT Persistent	22.1%	59.9%	3.8%	0.8%	86.6%
BP CHPT Non-Persistent	8.6%	2.0%	0.2%	0.1%	10.9%
IV Persistent	0.7%	N/A	N/A	N/A	0.7%
IV Non - Persistent	1.8%	N/A	N/A	N/A	1.8%
Total Oil Outflow	33.4%	61.9%	3.9%	0.9%	100.0%

Aside from the aggregate results in Tables E-9 and E-10 we are able to develop geographic profiles of average oil outflow by grid cell similar to the geographic profiles of interactions and accident frequencies presented in Appendix D. Appendix G will provide the geographic profiles for each case for the different oil types: BP CHPT Persistent, BP CHPT Non-Persistent, Interacting Vessel (IV) persistent and Interacting Vessel Non-Persistent. In this appendix we shall suffice by showing the accident frequency geographic profile results by accident type (for the calibration VTRA Case B) followed by its aggregate geographic oil outflow profile. The comparison of these two profiles illustrates geographically the effect of the additional oil outflow analysis layer on top of the accident frequency layer.

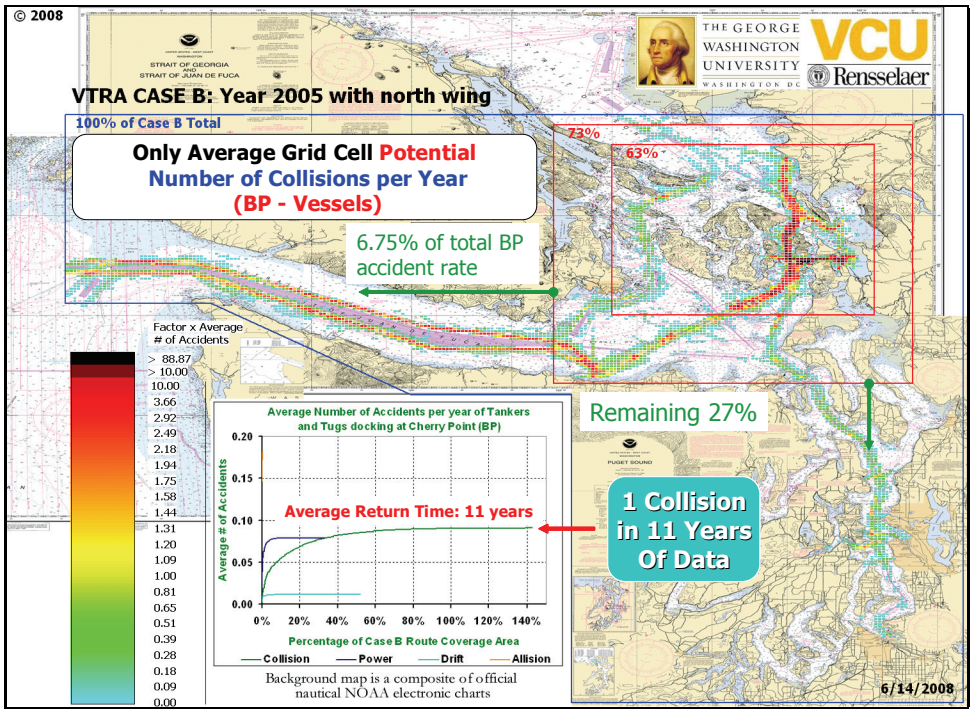


Figure E-17. Annual average collision frequencies of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

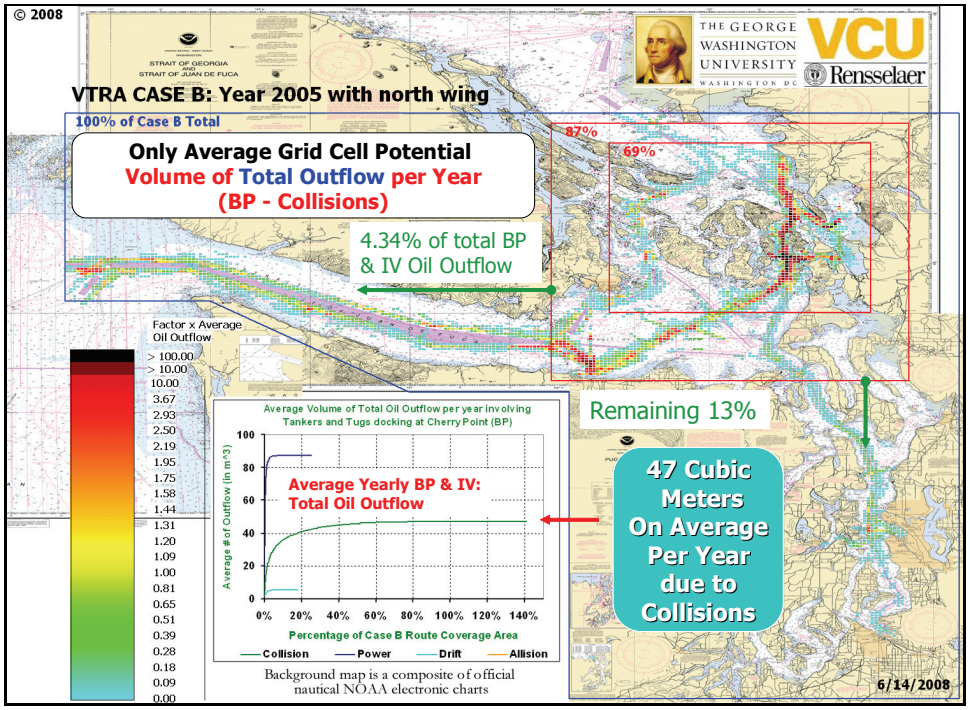


Figure E-18. Aggregate average oil outflow from collision with Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

Figures E-17 and E-18 respectively display the geographic profiles for collisions for the calibration VTRA Case B in terms of accident frequency and oil outflow. Firstly, we observe that the largest red square in Figure E-18 indicates 87% of the total oil outflow within this area whereas in terms of accident frequency this red square contains 73% of the accident frequency (a difference of 14%). Hence, we see a further concentration within this largest red-square when going from accident frequency to oil outflow. This is largely explained by the lightening of the colors in the West Strait of Juan de Fuca in case of the geographic oil outflow profile when compared to the geographic accident frequency profile. Finally, we observe only a difference of 6% in the percentage of accident frequency in the smallest red-square when going from oil outflow to accident frequency. This is exemplified by a darkening effect within the Rosario Strait area when going from accident frequency to oil outflow. While this too reflects a further concentration within this smaller red-square, the earlier difference of the 14% (when comparing the larger red-square) reflects a larger concentration effect outside the smallest red-square (but within the largest one). Indeed we do observe quite a darkening of color in front of the Port Angelas area when going from accident frequency geographic profile to oil outflow geographic profile.

Figures E-19 and E-20 respectively display the geographic profiles for drift groundings for the calibration VTRA Case B in terms of accident frequency and oil outflow. Figures E-21 and E-22 respectively display the geographic profiles for powered groundings for the calibration VTRA Case B in terms of accident frequency and oil outflow. While we have an overall annual frequency of groundings of ≈ 0.09 (average return time of 11 years), we obtain for average annual frequencies of drift grounding and powered grounding for VTRA Case B:

Drift Grounding: ≈ 0.012 (average return time of ≈ 85 years),
Powered Grounding: ≈ 0.079 (average return time of ≈ 13 years),

This coincides with a ratio of 6.8 of powered groundings to drift groundings. This ratio was explained in more detail in Appendix D. If we now evaluate the total average oil outflow for drift groundings and powered grounding we have (see Table E-9):

Drift Grounding: ≈ 5.5 (in cubic meters),
Powered Grounding: ≈ 87 (in cubic meters).

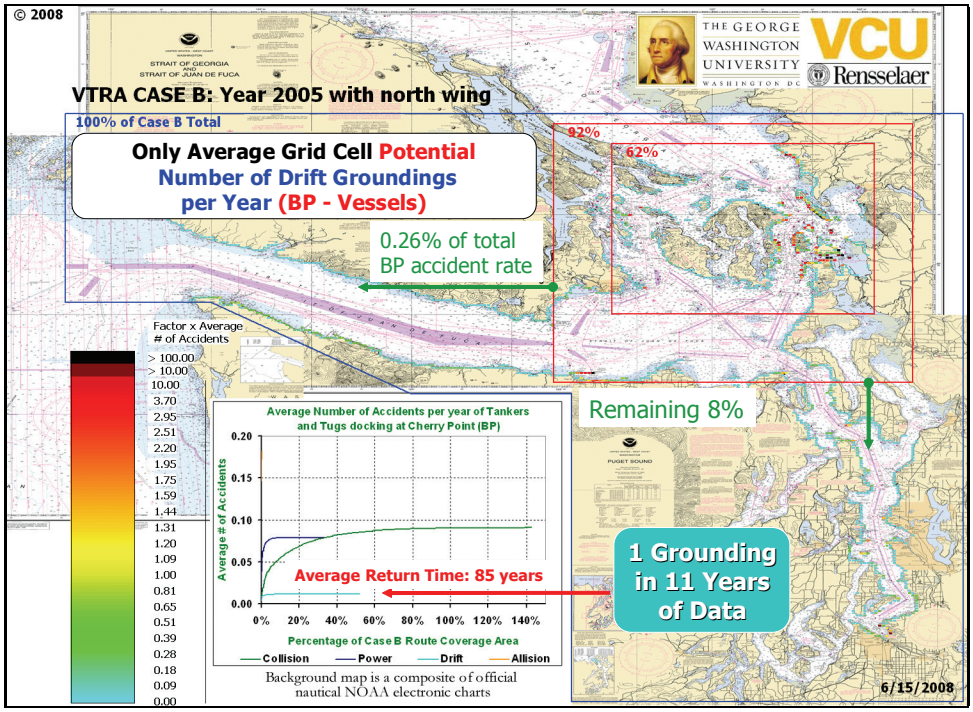


Figure E-19. Annual average drift grounding frequency of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

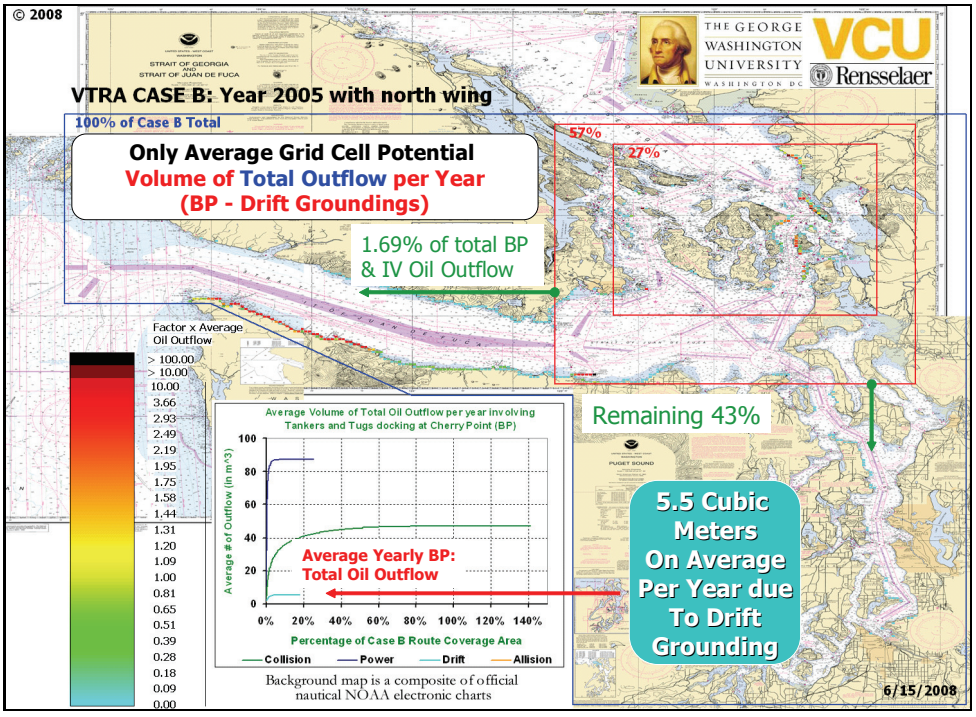


Figure E-20. Aggregate average oil outflow due to drift groundings of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

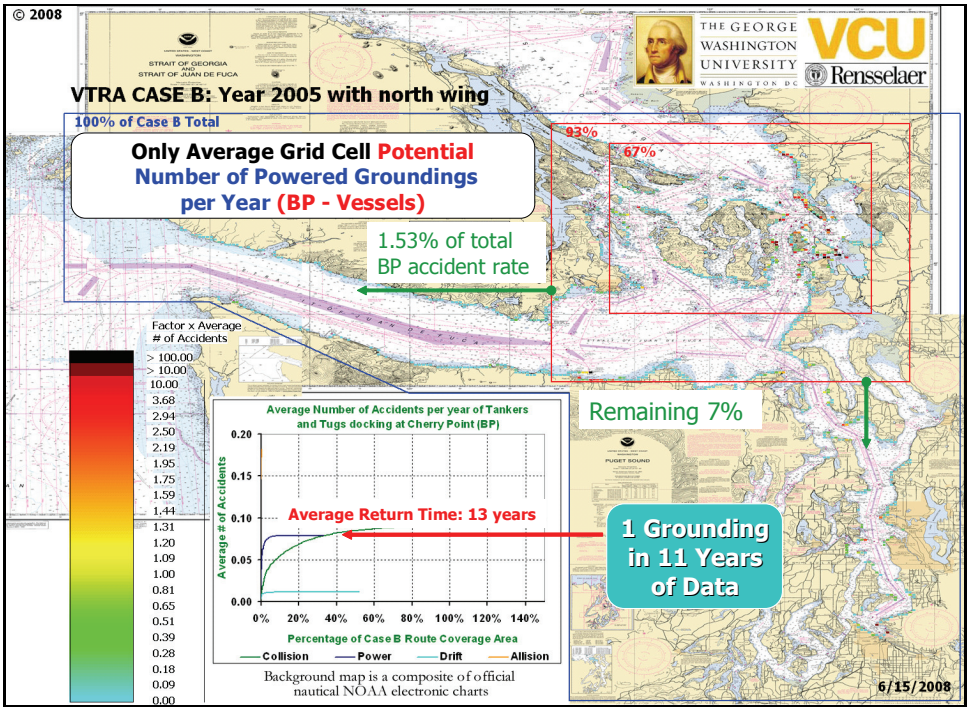


Figure E-21. Annual average powered grounding frequency of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

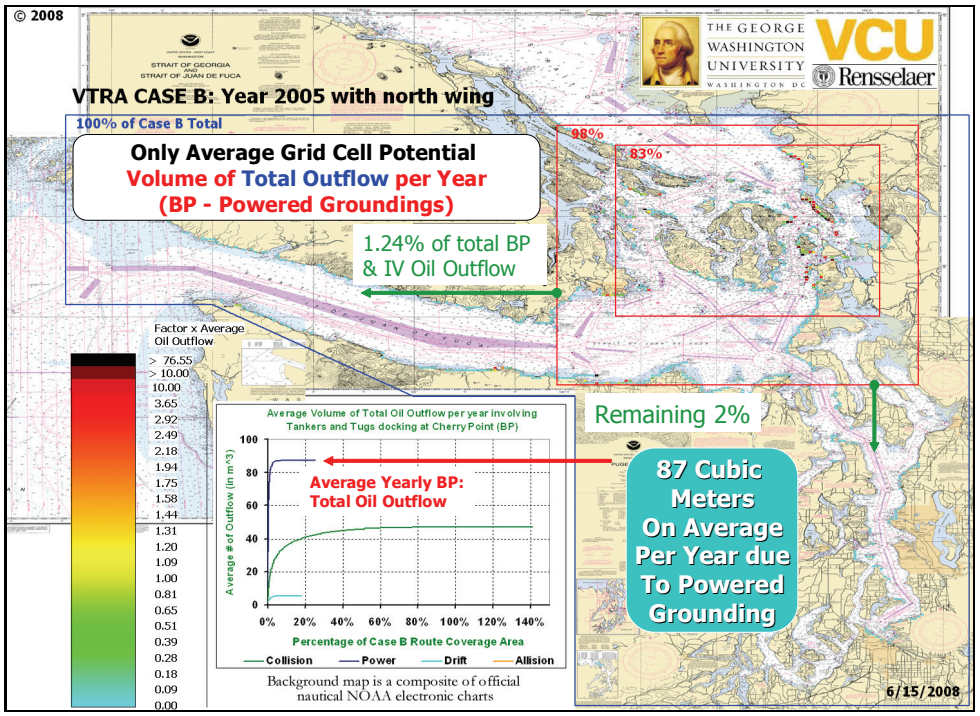


Figure E-22. Aggregate average oil outflow due to powered groundings of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

This coincides with a ratio of 15.8 of powered groundings to drift groundings. The increase from the ratio 6.8 in terms of accident frequency is explained here by the higher speeds at the time of grounding when under power as compared to when drifting. This results in a higher kinetic energy at the time of impact, larger damage extents and thus higher oil outflows in the case of power groundings as compared to drift groundings.

Observe from Figures E-19 and E-20 that we go from 92% to 57% for drift groundings in the largest red square when going from accident frequency to oil outflow. Observe from Figures E-21 and E-22 that we go from 93% to 98% for powered groundings in the largest red square when going from accident frequency to oil outflow. Hence, we see a reversal in behavior with respect to this red square when we go from drift groundings to power groundings.

This is partially explained by the distribution of accidents in the West Strait of Juan de Fuca. While we see somewhat of an even distribution in case of powered groundings to the north and to the south in the West Strait of Juan de Fuca, one observes a higher propensity to the south of West Strait of Juan de Fuca in case of drift groundings. This is primarily explained by the drifting patterns as a results of prevailing winds and currents in this area. Combined with the fact that the inbound traffic in West Strait of Juan de Fuca contains the laden BP CHPT tankers, whereas the outbound tankers are part of the outbound traffic, we see an effect on the oil outflow redistribution relative to the largest red square as above. Perhaps a larger explanation of this redistribution is due to modeling assumption that we have applied to the speed of impact in case of a drift grounding when the tanker is tethered. We have applied an additional speed reduction at the time of impact of on average 0.44 knots per minute of the time-to-shore recorded variable along the drifting path when the tanker is tethered. We evaluated this average speed reduction per minute from the "Strait of Georgia Full-Scale Trials" report by Wingard and Gray (1997). Tethering is primarily practiced within the area of the largest read square.

However, if we combine with the information above the data from Table E-10 that in our analysis about 62% of the total average oil outflow arises from powered grounding and about 4% arises from drift grounding, we still arrive at the same conclusion towards the end of appendix D that the predominant oil outflows over the entire VTRA study area are confined to the largest red square. Indeed, when aggregating the average oil outflows from all accident types in a single plot we still arrive at a total percentage of 92% of average oil

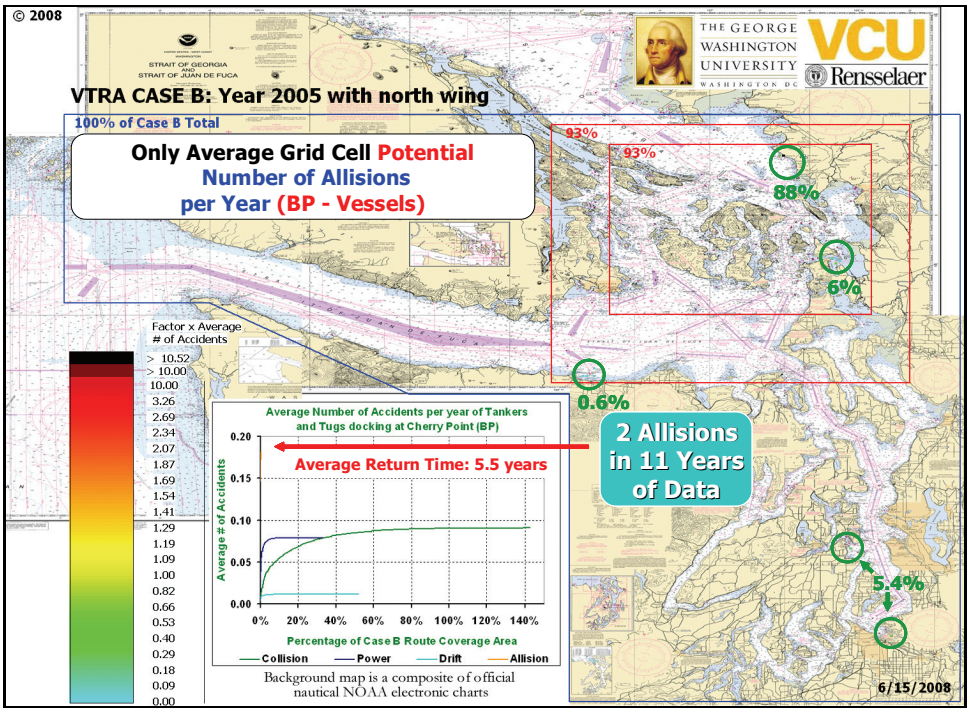


Figure E-23. Annual average allision frequency of BP Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

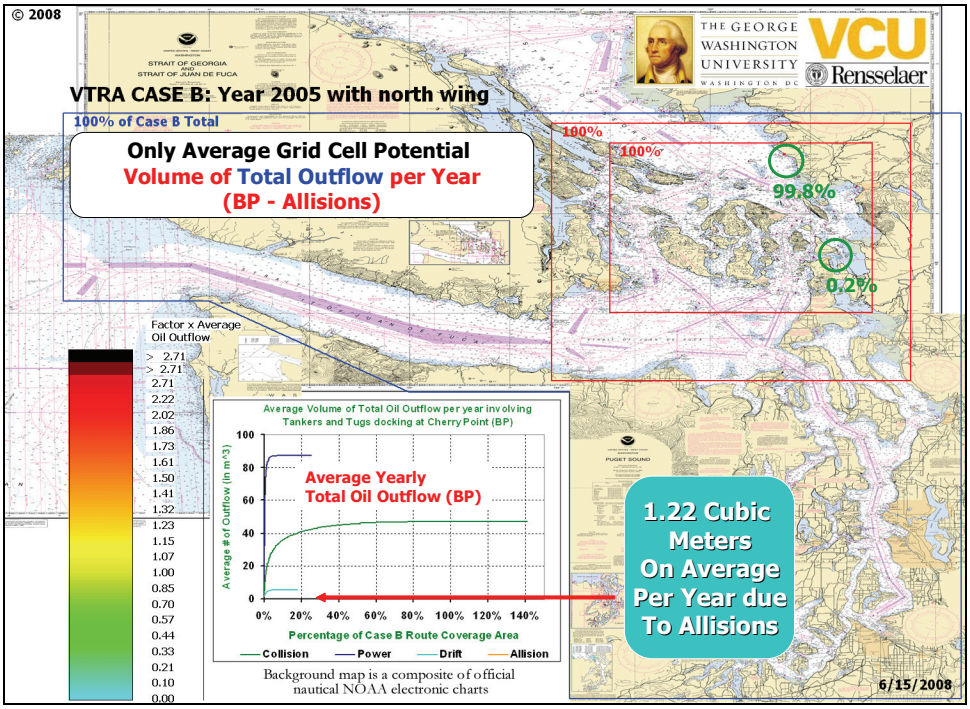


Figure E-24. Aggregate average oil outflow due to allisions of BP Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

outflow (see Figure E-24) within the largest red square (and thus 8% outside of it, which is not negligible).

Observe from Figures E-23 and E-24 that we go from 88% to 99.8% for allisions at the BP Cherry Point dock when going from accident frequency to oil outflow. This is primarily explained by the fact that when tankers dock at the BP Cherry Point dock they are fully laden whereas the other docks involve a mix of partially laden and even unladen tank vessels. While this change seems to be a dramatic one needs to bear in mind that of the total analyzed average annual oil outflow of about 141 cubic meters for the calibration VTRA Case B, only 1.22 cubic meters originates on average from allisions, which represents just about 1% of the total average oil outflow analyzed.

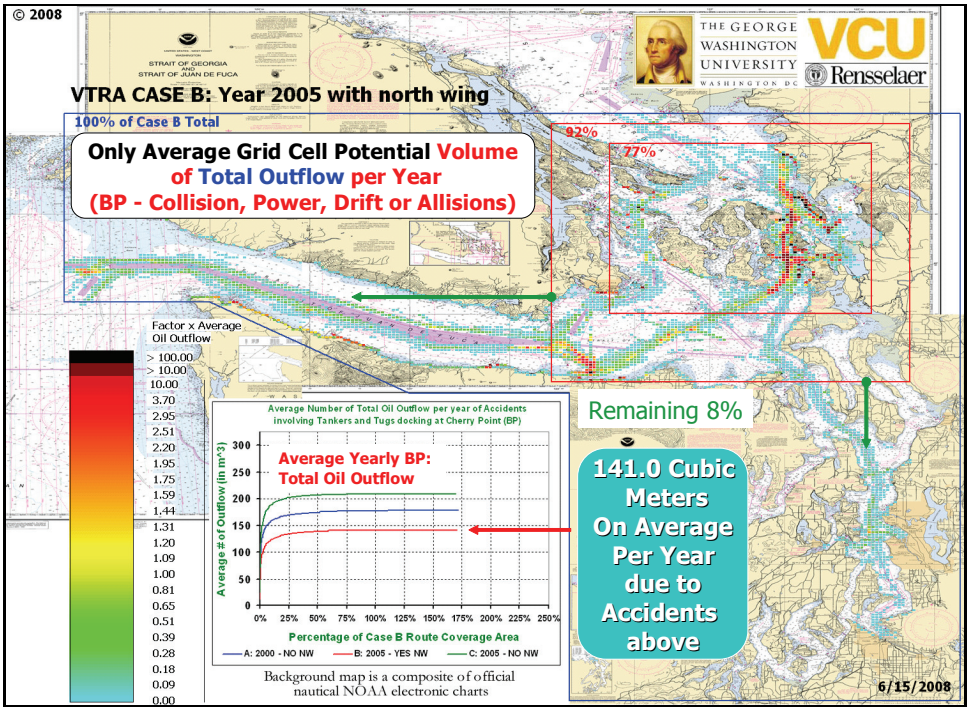


Figure E-25. Aggregate average oil outflow from accident types involving Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

With the VTRA Case B calibrated for CHPT vessels, the VTRA Case B simulation generates on average the same frequencies of incidents and accidents as observed in the accident-incident database analysis described in detail in Appendix A. Modifications can now be made to this VTRA Case B simulation to represent various alternatives and scenarios. For example, VTRA Case B represents the 2005 year with the BP Cherry Point North wing dock in operation. We can simulate the behavior of the CHPT vessel traffic as if this North wing

dock was not there. This case is labeled VTRA Case C. The case using a modification of the VTRA simulation to represent 2000 traffic levels is designated VTRA Case A. Next, we can compare the aggregate analysis results of VTRA Cases A and C to those of VTRA Case B and draw overall conclusions regarding the aggregate effect of potentially removing the North Wing in our model.

While the analysis above demonstrates that it is informative for the planning of potential future risk interventions where "average oil outflow risk is coming from" (both from an exposure, accident frequency and an oil outflow perspective), it also demonstrates that a comparison of different VTRA Cases ought to be based on average aggregate results for the entire VTRA study area. Such a comparison is provided also in the plot within Figure E-25. The red line indicates the aggregate oil outflow distribution over the grid cells that have oil outflow for VTRA Case B (2005 with North Wing) and the blue and green line respectively provide these results for VTRA Case A (2000 without North Wing) and VTRA Case C (2005 without North Wing). From this plot if follows that we analyzed the least annual aggregate average oil outflow over the entire VTRA Study area for the VTRA Case B (2005 with North Wing).

The geographic profiles allow us to further zoom-in on these aggregate effects by comparing those of VTRA Case B (see Figures E-17 to E-25) to those of VTRA Cases A and C (provided in Appendix G). By zooming in, one obtains a better general understanding about where this aggregate change in level (and possibly migration) of accident frequency or oil outflow from one case to another comes from. Visual comparison of these geographic profiles allows one to draw high level conclusions regarding general tendencies about the changing "risk" behavior from case to case or alternative to alternative.

It should be noted, however, that the maritime transportation modeled within the VTRA simulation is highly dynamic (as demonstrated by a running simulation) and relatively sparse. Even though we evaluate a total of 61427 vessel interactions for VTRA Case B distributed over a total of 3454 grid cells, this results on average annually in about 18 interactions per grid cell. Hence, when making changes to the VTRA Case B simulation this may results in high relative differences from grid cell to grid cell (especially in those with an even smaller number of interactions). In fact, from case to case one may experience an increase in one grid cell and a decrease in grid cells immediate adjacent to it. Hence, our general position is that these geographic profile analyses should not be used to perform grid cell by grid cell

comparisons from case to case, but should only be used to observe general tendencies of change for larger areas.

References

A. J. Brown (1995). "Oil Tanker Design Methodology Considering Probabilistic Accident Damage", *International Journal Of Advanced Manufacturing Systems*. Vol.4, Issue 1(2001) 25-34.

A. J. Brown (1998). Assessing the environmental performance of tankers in accidental grounding and collision", *SNAME Transactions,* Vol. 106: pp. 41-58.

A. J. Brown (2001). *Alternative Tanker Designs, Collision Analysis*. NRC Marine Board Committee on Evaluationg Double-Hull Tanker Design Alternatives.

International Marine Board (1995). "Interim Guidelines for Approval of Alternative Methods of Design and Construction of Oil Tankers under Regulation 13F(5) of Annex I of Marpol 73/78", Resolution MEPC.66 (37), Adopted September 14, 1995.

M. van der Laan (1997). *Environmental Tanker Design*. Delft University of Technology.

National Research Council (2001). "Environmental Performance of Tanker Designs in Collision and Grounding", Special Report 259, Marine Board, Transportation Research Board, The National Academies.

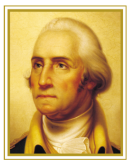
K. K. Tikka (2001). *Alternative Tanker Designs, Grounding Analysis*. NRC Marine Board Committee on Evaluationg Double-Hull Tanker Design Alternatives.

R. Taggart (1980). *Ship Design and Construction*, The Society of Naval Architects and Marine Engineers.

P.A. Wingard and D.L. Gray (1997). *Strait of Georgia Full-Scale Trials,* The Glosten Associates, Inc. File No. 97022.

SUB-APPENDIX:

G.F. van de Wiel (2008). "A Probabilistic Model for Oil Spill Volume in Tanker Collisions and Groundings", Master's Thesis: Applied Mathematics, Delft University of Technology, The Netherlands.



THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON DC

TU Delft

Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Department of Risk Analysis

A Probabilistic Model for Oil Spill Volume in Tanker Collisions and Groundings

A thesis submitted to the
department of Risk Analysis
in partial fulfillment of the requirements

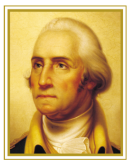
for the degree

MASTER OF SCIENCE
in
APPLIED MATHEMATICS

by

GIEL F. VAN DE WIEL

Delft, The Netherlands
June 2008



TU Delft

MSc THESIS APPLIED MATHEMATICS

"A Probabilistic Model for Oil Spill Volume in Tanker Collisions and Groundings"

GIEL F. VAN DE WIEL

Delft University of Technology

Daily supervisor

Dr. D. Kurowicka

Responsible professor

Prof. dr. R.M.Cooke

Other thesis committee members

Dr. ir. J.R. van Dorp

Dr. ir. P.H.A.J.M. van Gelder

June 2008

Delft, The Netherlands

Preface

This document presents the results from my thesis research at the George Washington University (GWU) in Washington, D.C., U.S.A. It was written for the purpose of completing the Master's Programme in Applied Mathematics at Delft University of Technology (DUT), The Netherlands. The research was supervised by Dr. Rene van Dorp at the GWU Department of Engineering Management and Systems Engineering and by Dorota Kurowicka at the DUT Department of Risk Analysis. Partial funding for this work was provided by The George Washington University.

Acknowledgements

Without the support of many people this report would not have been possible. First off I would like to thank Roger Cooke, Dorota Kurowicka, Rene van Dorp and Pieter van Gelder, for taking place in my thesis committee. In this regard, my thanks go out in particular to Roger Cooke and Thomas Mazzuchi for their support and advice throughout the thesis, both in practical as well as in academic terms; and a very special thanks to Rene van Dorp, who has given me the opportunity to take on this project and has supported me from its very start; whose knowledge, perseverance and enthusiasm I could not have found in any other supervisor.

Secondly, my gratitude goes out to all the friends I have come to know throughout my years of study and whose presence has played a big part in completing this thesis. I would like to thank the International Student House in Washington, D.C. and its residents for making my stay in the USA a truly memorable experience.

Lastly, I would like to thank my family: especially my parents, whose unconditional support helped me accomplish what I have today.

Contents

List of Tables

List of Figures

List of Symbols

Collisions

Symbol	Description
$e_{k,p}$	Perpendicular kinetic energy
$e_{k,t}$	Tangential kinetic energy
m_1	Striking ship mass
m_2	Stuck ship mass
v_1	Striking ship velocity
v_2	Struck ship velocity
ϕ	Collision angle
l	Relative collision location
l'	Absolute collision location from center
t	Striking ship type
η	Striking ship bow half entrance angle
y_l	Damage length
y_t	Maximum penetration
z	Oil outflow volume
z'	Occurrence of oil outflow

Groundings

Symbol	Description
e_k	Kinetic energy
o_d	Obstruction depth
o_a	Obstruction apex angle
o_r	Obstruction tip radius
c	Obstruction eccentricity
y_l	Longitudinal damage extent
y_t	Transversal damage extent
y_v	Obstruction elevation
z	Oil outflow volume
z'	Occurrence of oil outflow

Chapter 1

Introduction

1.1 Background

Maritime transportation plays an unreplaceable and ever-growing role in the global economy, taking up 96% of the world's global freight in terms of weight [17]. In 2006, seaborne trade grew 5.5% to 30,686 billion ton-miles. Of goods loaded, crude oil and petroleum products represented 36% [22]. Of course, transportation of goods by sea carries the risk of marine accidents, i.e. an event where a ship adversely interacts with its environment, possibly causing damage to either the ship, the environment, or both. When oil tankers are involved in accidents, a typical consequence of resulting damage is the release of crude oil or petroleum products into the sea.

Seaborne oil spills from tanker ships have the potential to cause major environmental damage, interfering with marine and coastal biology and influencing human livelihoods for decades after a spill occurs. These spills are usually accidental in nature; from 1995 to 2004, over three quarters of spills greater than 7 tons were caused by collisions and groundings [8]. Although the trend in both frequency and volume of spills has gone down significantly over the decades, the environmental risk of a spill remains significant and severe because of both the immensity of worldwide maritime transportation, the large amounts of oil transported by a typical tanker, and the increased likelihood of vessels interacting with each other due to traffic growth in harbors and waterways.

The context of this study was a Vessel Traffic Risk Assessment in which The George Washington University was tasked to evaluate incremental oil transportation risk as a result of potential traffic increases due to a dock expansion of a refinery in Washington State. Oil transportation routes traverse through the San Juan Islands and the Straits of Juan de Fuca. The San Juan Islands area is considered an environmentally pristine area and serves as a habitat for an Orca Whale family. Moreover, The San Juan Islands and the Strait of Juan de Fuca are fishing grounds for both commercial and tribal salmon, crab and shrimp fisheries.

Figure 1.1: Left: the Exxon Valdez, grounded in Prince William Sound. Right: pooled oil stranded between rocks after the Exxon Valdez grounding. (Source: National Oceanic and Athmospheric Administration)

1.1.1 The Exxon Valdez Grounding

On March 24, 1989, the oil tanker *Exxon Valdez* ran aground shortly after leaving the Valdez oil terminal in Alaska, spilling 36,000 metric tons of crude oil into Prince William Sound and beyond, in total affecting 1,500 miles of coastline (see Figure 1.1). Although only the 28th largest historical spill by volume [12], this accident became world news as the spilled oil contaminated the Prince William Sound coastline, seriously affecting the health and abundance of local shoreline biology as well as compromising the economic and public value of Prince William Sound. In its aftermath, Exxon —the company owning the Exxon Valdez— payed about US$ 2 billion in cleanup costs and court settlements and was sentenced to pay US$ 2.5 billion in punitive damages. In response to the spill, the United States Congress passed the 1990 Oil Pollution Act to prevent further oil spills from occurring in the United States.

1.1.2 Modelling Oil Spill Risk

To improve prevention of future oil spills after Exxon Valdez, numerous models for analyzing oil spill risk were developed. In the Prince William Sound Risk Assessment [14], a system simulation of Prince William Sound that integrated shipping fleet, traffic rules and operating procedures was run to generate a dataset of accident types and locations over a timespan of 25 years. This assessment was based on Probabilistic Risk Analysis [1], which:

1. Identifies the series of events leading to an accident;

2. Estimates the probabilities of these events;

3. Evaluates the consequences of the accident.

Brown and Amrozowicz [4] propose a model that consecutively determines

1. Accident probability (grounding, collision, structural failure etc.);
2. Probability of zero outflow and mean outflow volume given a spill;
3. Immediate response to contain the spill;
4. Spill consequence.

A similar methodology is provided by the software package GRACAT [5], short for Grounding and Collision Analysis Toolbox, which has the following modelling capabilities:

1. Frequency: estimation of grounding or collision probability for a vessel operating on a specified route;
2. Damage: establishment of models for calculating the resulting grounding and collision damage;
3. Consequence: analysis of the conditions of the damaged vessel;
4. Mitigation: identification and evaluation of remedial measures for the considered consequences.

Looking at these methodologies, to model the risk of an individual tanker spill, one can argue that in general one has to:

1. Determine the probability of an accident *given* the state of the surrounding environment;
2. Determine the oil outflow volume *given* an accident;
3. Determine the spill consequence *given* the outflow volume.

This report focuses entirely on the 2^{nd} item: the modelling of oil outflow volume from an oil tanker given that an accident involving the tanker has occurred.

1.1.3 IMO Outflow Model

A widely accepted model used in determining the oil outflow volume in tanker accidents was drafted by the International Maritime Organization [9]. The purpose of the model is to measure outflow performance of a particular tanker design against a reference double hull design.

For this model, data was taken from approximately 100 historical collision and grounding scenarios from the period 1980-1990 to establish probability density functions (PDFs) for the location and extent of damage in a collision or grounding scenario (see Figure 1.2). Based on these distributions, each unique combination of tanks or compartments in a given tanker design can be associated with a probability of being damaged.

In a collision, the assumption is made that all oil is lost from a damaged compartment. Hence the sum of cargo volumes of damaged compartments represent the total volume of spilled oil. In a grounding, a pressure balance calculation is carried out, where the water level surrounding the tanker determines the amount of oil that flows out.

After this calculation step, the probability of damage and outflow volume for each unique combination of compartments is known. Using these numbers, three parameters describe the environmental performance of the tanker design in question:

- Probability of no outflow P_O: the cumulative probability for all damage combinations for which there is no oil outflow.
- Mean outflow parameter O_M: the weighted average of outflow volumes of all combinations.
- Extreme outflow parameter O_E: the weighted average of outflow volumes of the damage combinations falling within the cumulative probability range between 0.9 and 1.0.

These parameters are then combined into a "pollution prevention index" E:

$$E = k_1 \frac{P_O}{P_{OR}} + k_2 \frac{0.01 + O_{MR}}{0.01 + O_M} + k_3 \frac{0.025 + O_{ER}}{0.025 + O_E} \tag{1.1}$$

where $k_1 = 0.4$, $k_2 = 0.5$ and $k_3 = 0.1$; and where P_{OR}, O_{MR} and O_{ER} are respectively the probability of no outflow, mean outflow parameter and extreme outflow parameter of the reference double hull design. If $E > 1$, then the design in question has "satisfactory characteristics". An analysis using this methodology was used by the Herbert Engineering Corporation [6] to evaluate 96 different tanker designs to propose a standard tanker design.

Unfortunately, the IMO model suffers from a number of fundamental limitations. The following objections are raised as such [16, 23]:

- The model uses a single set of damage extent PDFs from limited single hull data applied to all ships, independent of structural design; realistically, however, this data should only be used to model single hull accidents.

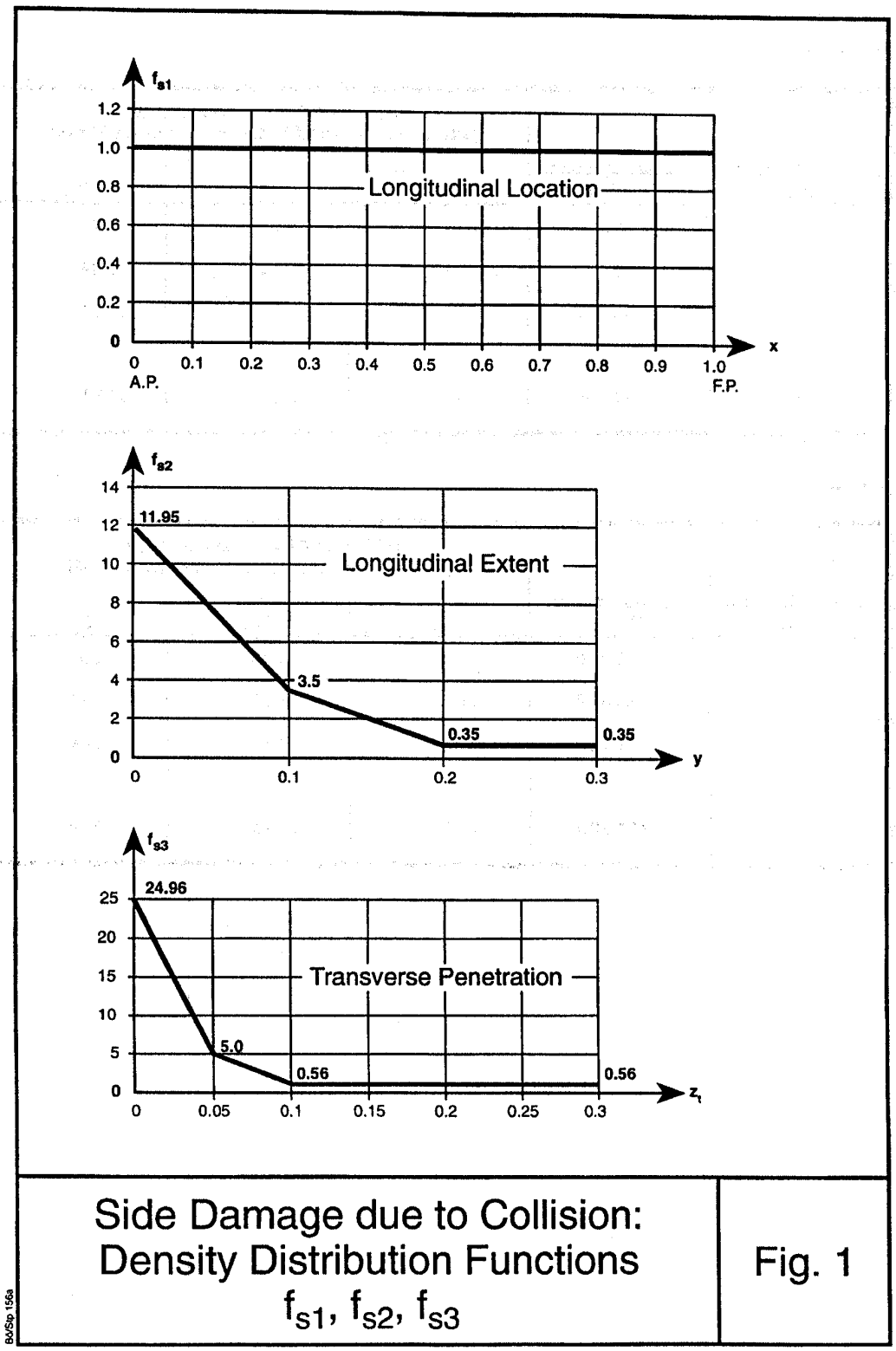


Figure 1.2: Damage extent PDFs, IMO model (source: IMO [9])

- Damage PDFs only consider damage that is significant enough to breach the outer hull. This penalizes structures able to resist rupture.
- Damage extents are treated as independent random variables when they are actually dependent variables, and ideally should be described using a joint PDF.
- Damage PDFs are normalized with respect to ship length, breadth and depth when damage may depend to a large extent on local structural features and scantlings. Most notably, Simonsen and Hansen [19] conclude that relative damage length in groundings is higher for larger ships than for smaller ones.

1.1.4 Collision and Grounding Models

In 2001, the Marine Board of the National Academy of Science published a report assessing a methodology to compare double hull tanker designs to alternative designs [20]. It noted that the IMO model was insufficient for the goals outlined by the report and that, consequently, further research was necessary. A risk-based methodology was therefore developed that included a model for generating probabilistic accident scenarios.

For both collisions and groundings this model is based on the physical simulation of accident damage inflicted on a tanker as developed by Brown [3] and Tikka [21] using the simulation programs SIMCOL resp. DAMAGE. For the Marine Board research, 10,000 collision and grounding scenarios were randomly generated and put through a simulation four times; each time using a different tanker design. This resulted in a dataset of 40,000 collisions and 40,000 groundings, describing input (i.e. ship speed, displacement, collision angle) and output variables (i.e. damage length, outflow volume).

The goal of having this large dataset was to compare outflow performance between single hull and double hull tankers; however, by carefully studying the relationships between input and output parameters of this large data set one can "empirically" develop a probabilistic model that determines accident oil outflow based on statistical data analysis techniques rather can computationally intensive physical simulations; one that nevertheless needs to adhere to the same physical principles as the latter.

The model is envisioned to be used in similar tools as the Prince William Sound Risk Assessment simulation [14]. These tools generate a large number of scenarios and hence the oil outflow volume evaluation needs to computationally efficient. Without oil outflow analysis, multiple year simulation runs take 8 hours or more, just to evaluate accident frequencies. Combining such a simulation tool with the physical damage simulations developed for the Marine Board is from a computational point of view impossible at this time. An explicit oil outflow model, however, that describes a statistical relationship between scenario input characteristics and oil outflow output

characteristics could very well be combined with such a simulation tool. These statistical relationships are estimated using the physical simulation data of the Marine Board report containing 80,000 collision and grounding scenarios.

1.2 Thesis Goal

The research goal of this thesis is to

- Develop a new method for modelling the oil spill volume of an oil tanker in a collision or grounding accident scenario, based on the simulation data as obtained from [3, 21];
- for both single hull and double hull tankers of specific designs;
- emphasizing on the practicality of implementation of the outflow model into large scale system simulations.

1.3 Thesis Outline

In the first chapter, the dataset generated by the collision and grounding simulations (as discussed above) are described. Next, the collision outflow model based on this data is explained and discussed extensively; following this, the grounding outflow model is treated. Because it adheres to the same principles as the collision model, only changes to the grounding methodology as opposed to collisions are mentioned. Third, a concise, practical example of the model is given to demonstrate its use in determining accidental oil outflow. Finally, the conclusions to the thesis goal and recommendations for further research are presented.

Chapter 2

Simulation Data

In the aforementioned research, 10,000 sets of input variables for both collisions and groundings were generated, and subsequently fed into a physical simulation model. These simulations were performed on four different tanker designs, resulting in a total of 80,000 sets of output variables; hence in total 80,000 pairs of input and output variables ('scenarios') are available. In this chapter, the ship designs, input variables, collision and grounding simulations, and resulting output variables are described and discussed in detail. It must be noted that there are differences between the ship designs used in the collision and grounding studies, which will be discussed when relevant.

2.1 Tanker Designs

An oil tanker is mainly characterized by its cargo area, which consists of one or more tanks or compartments. The cargo capacity is measured in deadweight tonnage (DWT) representing cargo mass. The displacement equals the water mass that the ship displaces. Among tankers, single-hull and double-hull designs are the most widespread used. As the name implies, in a single-hull design only one wall separates the cargo compartments from the surrounding water; in a double-hull design, these compartments are protected by ballast tanks. The four different tanker designs are designated by hull type and tonnage: SH40, SH150, DH40 and DH150. Their schematic designs can be found in Figures 2.1 and 2.2.

2.2 Collisions

In a collision, an oil tanker is struck by a striking ship (see Figure 2.3). The collision transforms translational motion mainly into rotational motion, elastic deformation and plastic deformation. It is assumed that the striking ship does not experience any damage. When a collision is severe enough, the hull of the oil tanker is penetrated and ruptured, resulting in a damaged area. If the damaged area overlaps with a compartment, all contents from this compartment are assumed spilled.

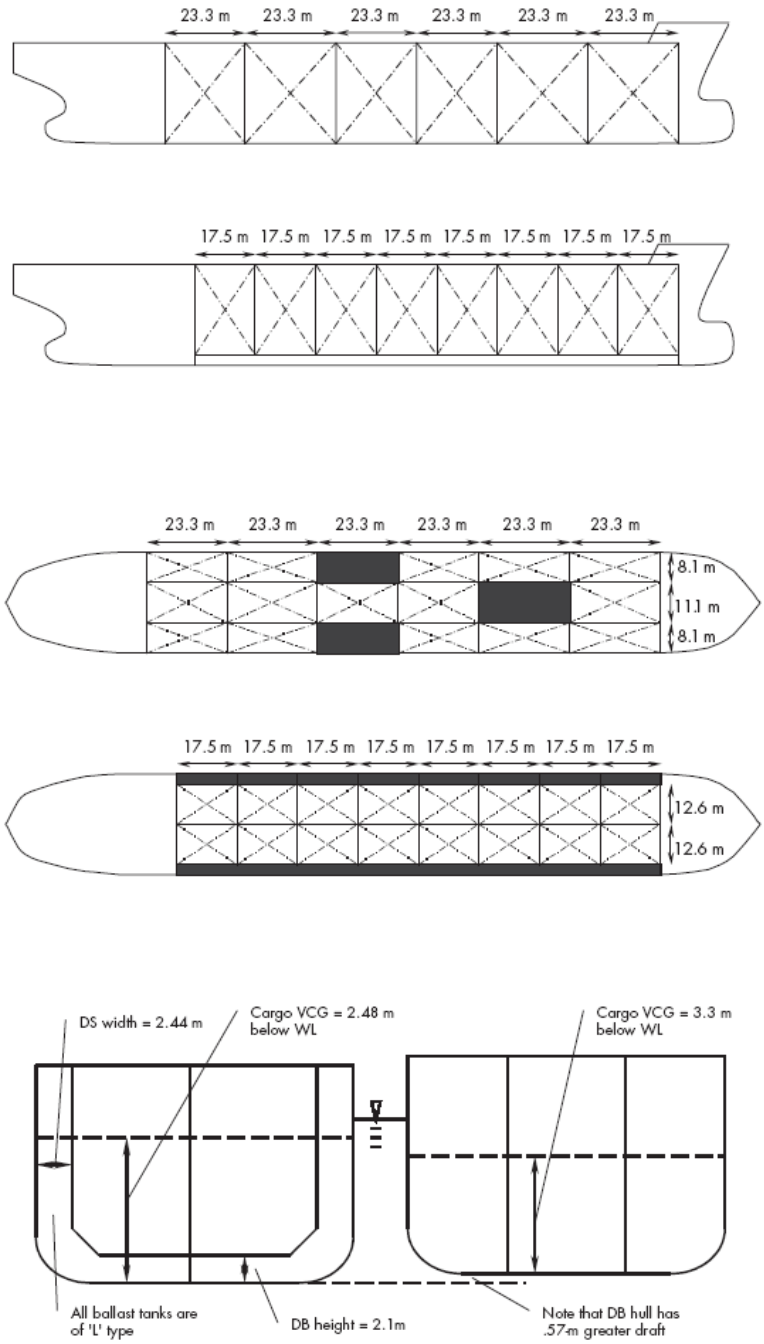


FIGURE 4-2 Profile, plan, and midship section for 40,000-DWT ships (VCG = vertical location of the center of gravity; WL = waterline).

Figure 2.1: 40,000 DWT tanker designs (source: National Academies Press [20])

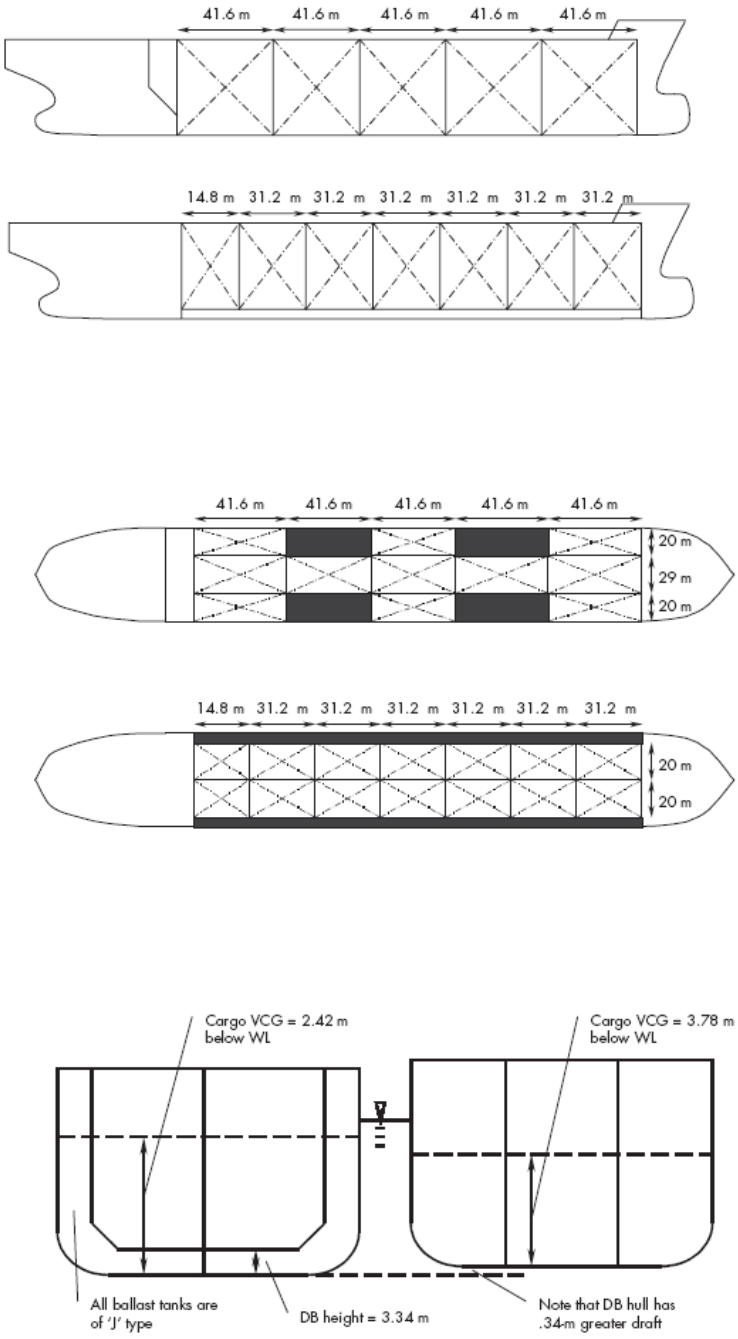


FIGURE 4-1 Profile, plan, and midship section for 150,000-DWT ships (VCG = vertical location of the center of gravity; WL = waterline).

Figure 2.2: 150,000 DWT tanker designs (source: National Academies Press [20])

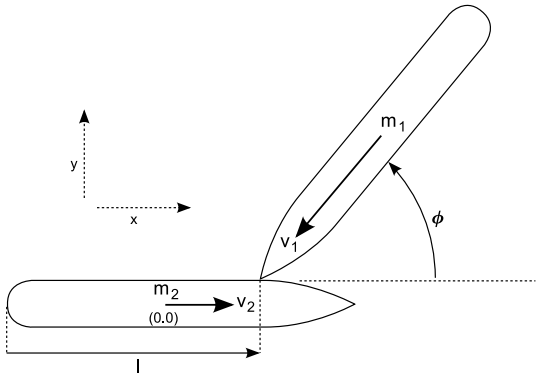


Figure 2.3: Two ships at the moment of collision

Name	Hull Type	Length (Meters)	Breadth (Meters)	Draft (Meters)	Deadweight Tonnage (Metric Tons)	Displacement (Metric Tons)
SH40	Single	201.168	27.432	10.603	40,000	47,547
SH150	Single	266.3	50.0	16.76	150,000	175,882
DH40	Double	190.5	29.26	10.58	40,000	47,448
DH150	Double	261.0	50.0	16.76	150,000	175,759

Table 2.1: Tanker specifications, collisions

2.2.1 Input Data

The specifications for the different tanker designs[1] that were used in the collision simulations are described in Table 2.1; an overview of compartment volumes for these ships is given in Table C.1 in the Appendix.nThe input variables in Table 2.2 are realizations of random variables with specific probability distributions. Together with other (fixed) parameters, like ship dimensions, plate thickness, compartment configurations etc. they define a collision scenario at the moment of impact. It is assumed that these variables are realizations of random variables which are defined by parametric distributions.[2]

- V_1 is characterized by a Weibull distribution with shape parameter $\alpha = 2.2$ and scale parameter $\beta = 6.5$;
- V_2 is given by an exponential distribution with parameter $\mu = 0.584$;
- Φ is the angle between port bows: if vessels travel in the same direction,

[1] Brown [3] is ambiguous as to whether the small designs (SH40 and DH40) have a deadweight tonnage (DWT) of 40,000 or 45,000; however, Tikka [21] gives a DWT of 40,000 for these designs. Therefore the decision was made to assume that the ships in the collision model also have a DWT of 40,000. Also, Brown mentions a length of 261.0m for the double hull in the report where the accompanying simulation file says 266.3m.

[2] By convention, random variables are denoted with capital letters; realizations of random variables are lowercase.

Input Variable	Symbol	Unit
Striking ship velocity	v_1	Knots
Struck ship velocity	v_2	Knots
Collision angle	ϕ	Degrees
Displacement of striking vessel	m_1	1000 metric tons
Collision location, relative from the stern	l	-
Striking ship type	t	-

Table 2.2: Input variables, collisions

$\Phi > 90°$; if not, $\Phi \leq 90°$. The distribution of Φ is approximated by a truncated Normal ($\mu = 90, \sigma = 28.97$) distribution; realizations are selected using Monte Carlo simulation on the interval $[0, 180]$. Although the use of Monte Carlo on a bounded support is only mentioned for Φ in the report, it is believed that this method is applied to other variables as well when a bounded support is imposed on distributions with infinite support.

- L gives the relative distance of the collision location from the Aft Perpendicular (AP) of the ship. $L = 0$ means the collision takes place at the AP, where $L = 1$ represents a collision at the FP[3]. It follows a $Beta(1.25, 1.45)$ distribution with support on $[0, 1]$ (see Appendix B for an explanation on distributions).

- T is one out of five types of striking ships: tanker, bulk cargo, freighter, passenger or container. Each type has its own characteristics; among the distinctions taken into account in the simulations is the bow half entrance angle η, which is the angle between bow and the longitudinal axis of the ship and is given for each type, and displacement M_1 which is a Weibull-distributed random variable. See Table 2.3 for the probability of occurrence of each striking ship type and Table 2.4 for the distribution of each type's displacement. Note that lower and upper bounds are given for displacement, whereas a Weibull distribution has support on $(0, \infty)$. Again, Monte Carlo simulation was probably used in selecting realizations of the Weibull distribution within the given bounds.

The aforementioned randomly generated variables are put into the collision simulation together with other parameters such as ship dimensions, struck ship displacement, compartment design, plate thickness, etcetera.

[3]The AP is the aftmost point of the bottom plane of the ship; the Forward Perpendicular (FP) is defined likewise.

Type t	Name	Probability of Occurrence	η (degrees)
1	Tanker	0.252	38
2	Bulk Carrier	0.176	20
3	Freighter	0.424	20
4	Passenger	0.014	17
5	Container	0.135	17

Table 2.3: Striking ship type distribution

		Weibull		Bounds (MT)	
Type t	Name	α	β	Lower	Upper
1	Tanker	0.84	11.2	699	273550
2	Bulk Carrier	1.20	21.0	1082	129325
3	Freighter	2.00	11.0	500	41600
4	Passenger	0.92	12.0	997	76049
5	Container	0.67	15.0	1137	58889

Table 2.4: Striking ship displacement distribution, by type

2.2.2 Output Data

When the simulation is over, three output variables are generated:

- Damage length y_l, meters
- Maximum penetration y_t, meters
- Oil outflow volume z, cubic meters

Damage length is the extent of the damaged area in the struck ship's longitudinal direction. Maximum penetration is the maximum extent of the damage in transversal direction. Oil outflow is the total sum of volumes of damaged compartments, i.e. compartments that coincide with the damaged area. See Figure 2.4 for a schematic view of an example of the damaged area. The distribution of the resulting output variables for all ship types are presented in Table 2.5 and Figures 2.5 and 2.6.

It must be noted that, when outflow occurs, y_l and y_t are nonzero; however, the reverse is not always the case. Therefore, there may be collision scenarios where there is damage but no outflow, for example in the case of plastic deformation without hull breach, or the rupture of ballast tanks (which contain no oil) but no oil compartments. This is especially likely in double hull tankers, where all oil compartments are seperated from the outer hull by ballast tanks. In Table 2.6, the number of nonzero values of y_l and y_t from the collision scenario are given as well as the number of cases of zero outflow for each ship type to show how many times this occurs.

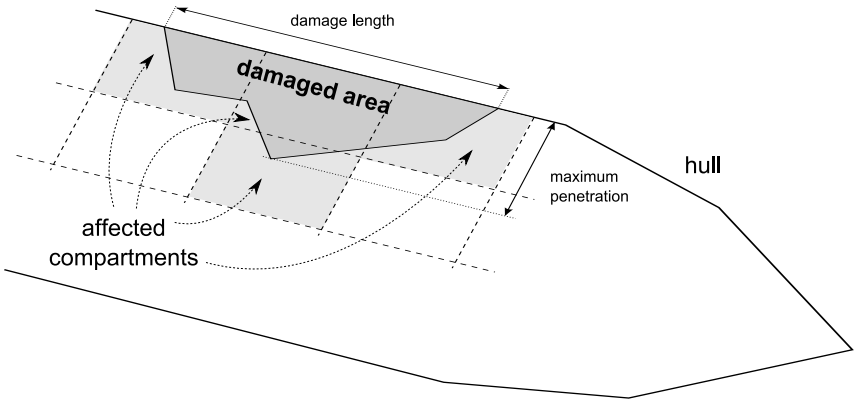


Figure 2.4: Collision damage

SH150		DH150		SH40		DH40	
Volume	Count	Volume	Count	Volume	Count	Volume	Count
0	6817	0	8974	0	5955	0	8596
3820	358	5515	84	1865	488	2270	97
8365	682	11694	86	2529	869	2277	133
12185	168	13862	129	2641	522	2670	189
13103	723	14650	119	2668	844	2825	84
15311	1150	14651	274	2674	797	2846	471
18864	7	14674	87	3644	20	5095	44
21567	13	19377	56	4506	116	5122	107
23479	5	26369	30	5197	156	5515	74
23676	1	28513	47	5314	131	5671	47
28023	2	29302	79	5507	10	5692	155
30882	8	29325	34	6171	21	7968	1
36875	11	43976	1	6312	12	10244	1
46888	13			6320	12	11383	1
51502	11			8147	6		
52449	8			8960	12		
55739	10			9956	13		
58441	12			9964	5		
70367	1			12483	8		
				12638	1		
				14275	1		
				16127	1		
Total	10,000	Total	10,000	Total	10,000	Total	10,000

Table 2.5: Outflow volume distribution, collisions

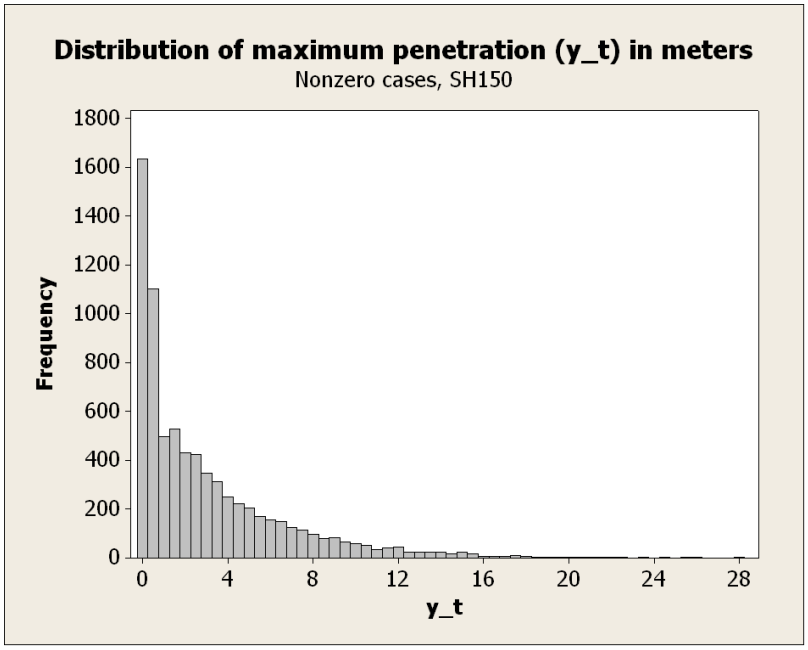


Figure 2.5: Maximum penetration histogram, collisions

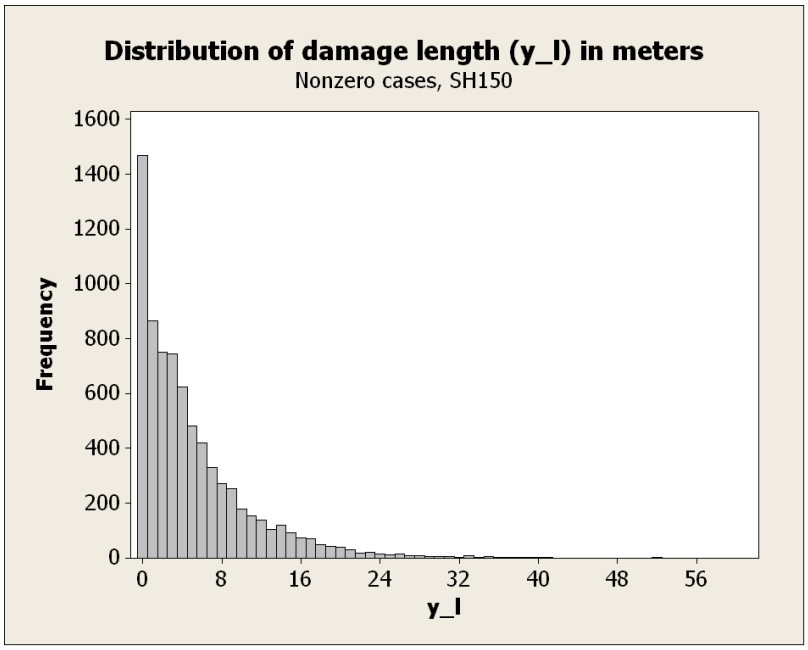


Figure 2.6: Damage length histogram, collisions

Type	SH40	SH150	DH40	DH150
number of nonzero z	4045	3183	1404	1026
number of nonzero y_l	7467	7473	7454	7466
number of nonzero y_t	7470	7478	7455	7467

Table 2.6: Nonzero output values from collision simulations

Name	Hull Type	Draft (Meters)	Deadweight Tonnage (Metric Tons)	Displacement (Metric Tons)
SH40	Single	10.58	40,000	47,448
SH150	Single	16.78	150,000	175,907
DH40	Double	11.17	40,000	49,410
DH150	Double	17.12	150,000	175,940

Table 2.7: Tanker specifications, groundings

2.3 Groundings

In a grounding, a tanker collides at the bottom with an obstacle, in this case a cone-shaped rocky pinnacle with a rounded tip (see Figure 2.7). The rock is assumed fixed and strong enough never to suffer any damage. Specifications for the struck ships in the grounding simulations differ slightly from those in collisions (see Table 2.7). An overview of compartment volumes for these ships is given in Tables C.5 through C.8 in the Appendix.

2.3.1 Input Data

The input variables in Table 2.8, along with fixed parameters such as ship dimensions, plate thickness etc. are put into the grounding simulation. They are realizations of random variables with specific probability distributions to form a specific grounding scenario at the moment of impact.

- V is distributed as in Table 2.9.
- In the report accompanying the grounding study [21], the distribution mentioned for O_d is different than the one found in the data. Therefore the latter distribution will be used later on to get a correct fit.
- O_a is distributed along a 'truncated' Normal distribution with support on the interval $[15, 50]$. Since the original report doesn't state the mean nor the variance of this normal distribution, it is assumed unknown and therefore a fit for this variable will also be determined later on.
- O_r is also characterized by a truncated Normal distribution on $[0, 10]$. Based on the data, it is assumed that the mean of the original distribution is 5, meaning

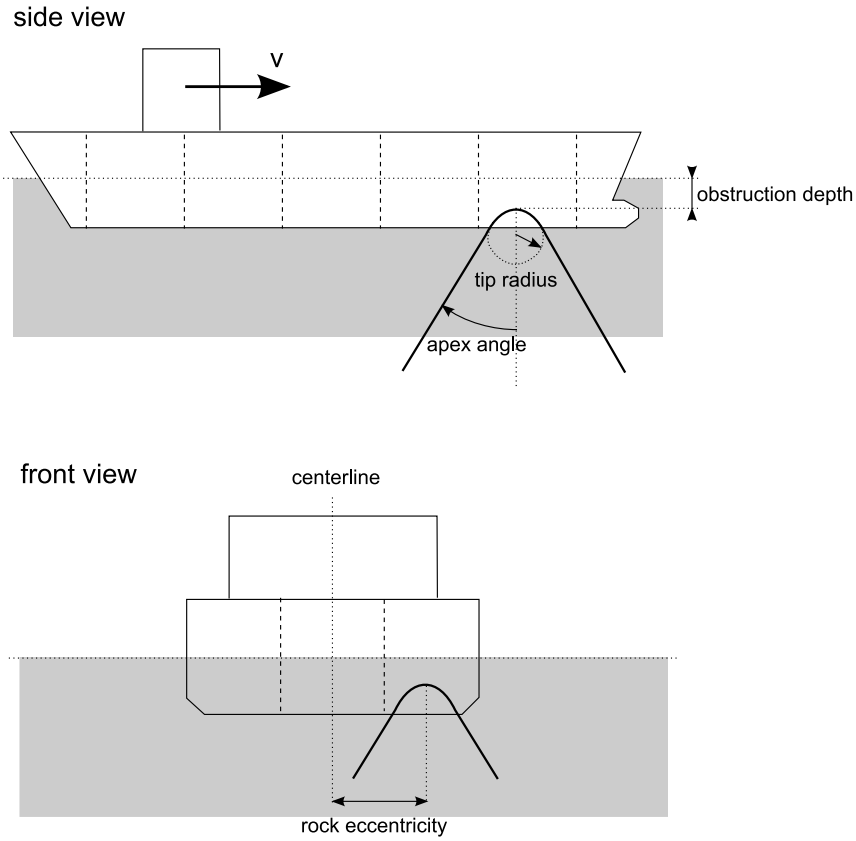


Figure 2.7: Grounding simulation

Input Variable	Symbol	Unit
Struck ship velocity	v	Knots
Obstruction depth from mean low water	o_d	Meters
Obstruction apex angle	o_a	Degrees
Obstruction tip radius	o_r	Meters
Rock eccentricity	c	-
Tidal variation from mean low water	τ	Meters
Inert tank pressure	p	mm water gauge
Capture in ballast tanks	b	% of tank volume
Minimum outflow	ν	% of ruptured tank volume

Table 2.8: Grounding input variables

Bin Bounds		
Lower	Upper	Probability
0	5	0.25
5	8	0.45
8	15	0.08
15	16	0.20
16	20	0.02

Table 2.9: Velocity distribution, groundings

Bin Bounds		
Lower	Upper	Probability
0	0.7	0.50
0.7	1.7	0.35
1.7	2.5	0.15

Table 2.10: Tidal variation distribution

$P(O_r \leq x) = 1 - P(O_r > 10 - x)$ for $x \in [0, 10]$. This variable will also be fitted later on.

- Rock eccentricity C is defined as the obstruction distance relative from the centerline, i.e. it is 0 if the obstacle hits the ship in the middle and 1 if it hits on either port or starboard side. C has a uniform $[0, 1]$ distribution.
- Tidal variation is distributed as in Table 2.10. Tank pressure, minimum outflow and ballast capture are uniformly distributed on intervals $[400, 1000]$, $[0.5, 1.5]$ and $[0, 50]$, respectively.

2.3.2 Output Data

Once a grounding simulation is complete, it generates the output variables described in Table 2.11. 'Elevation' is the height of the obstruction tip above the ship's bottom. If k is the number of cargo compartments, $z = \sum_{j=1}^{k} z_{c,j}$ is the total outflow volume (note that $z_{c,j} = 0$ if compartment j is not damaged). The histograms of the output variables y_{l1}, y_{l2}, y_t, y_v and z are displayed in Figures 2.8 through 2.11. Looking at the histogram of y_v (Figure 2.10), it seems that this variable is directly related to an input variable. Indeed, when plotted as a function of obstruction depth o_d (Figure 2.12), it becomes clear that

$$y_v = \max(0, s_d - o_d) \tag{2.1}$$

where s_d is the ship's depth. From the figure, it can be seen that this holds for all ship types.

Variable	Symbol	Unit
Begin of longitudinal damage extent	y_{l1}	Meters aft from midship
End of longitudinal damage extent	y_{l2}	Meters aft from midship
Transversal damage extent	y_t	Meters
Elevation	y_v	Meters from bottom hull
Outflow volume per cargo compartment j	$z_{c,j}$	Cubic meters
Volume captured in ballast tanks	z_b	Cubic meters

Table 2.11: Grounding output variables

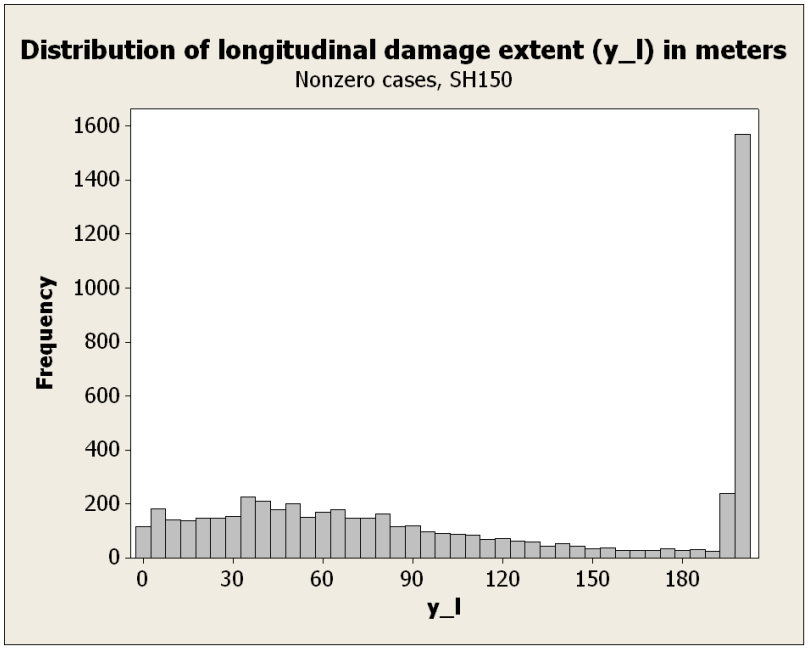


Figure 2.8: Longitudinal damage extent histogram, groundings

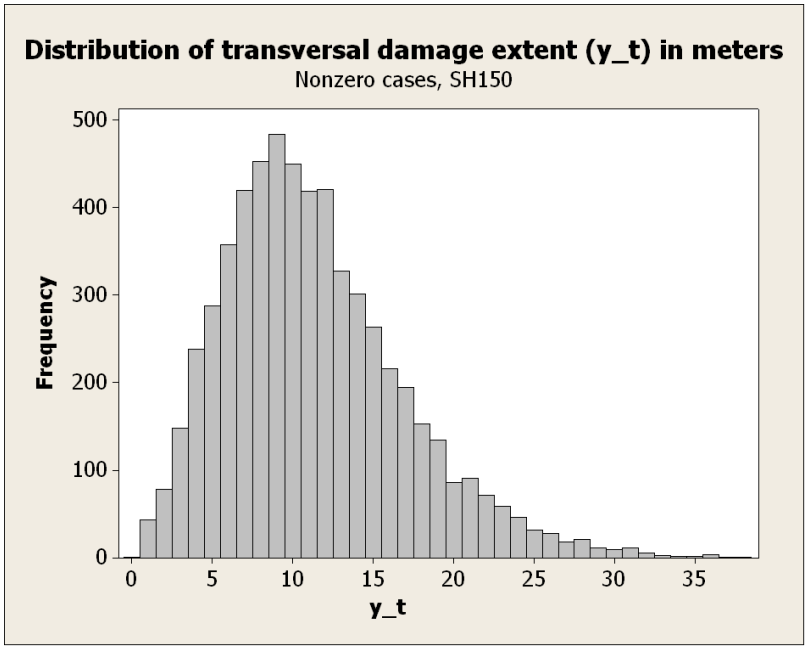


Figure 2.9: Transversal damage extent histogram, groundings

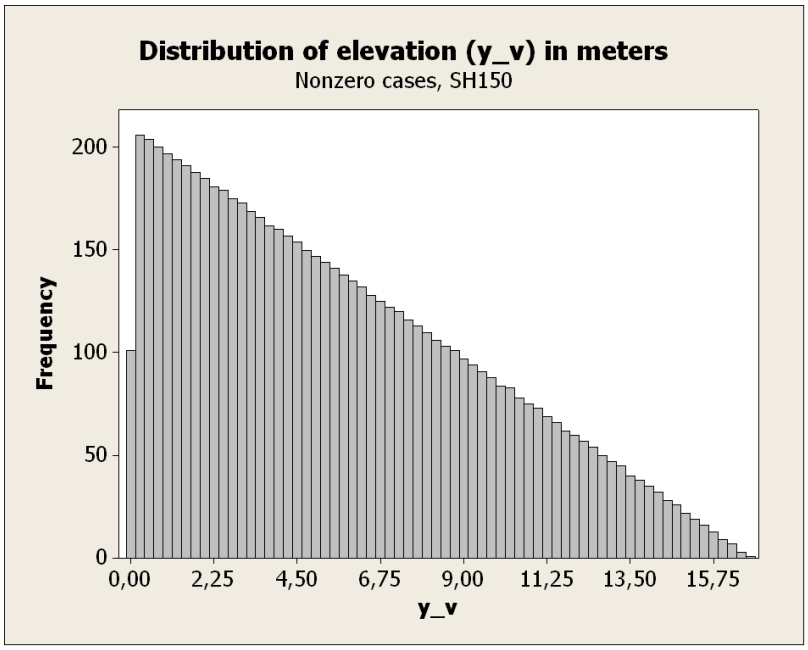


Figure 2.10: Elevation histogram, groundings

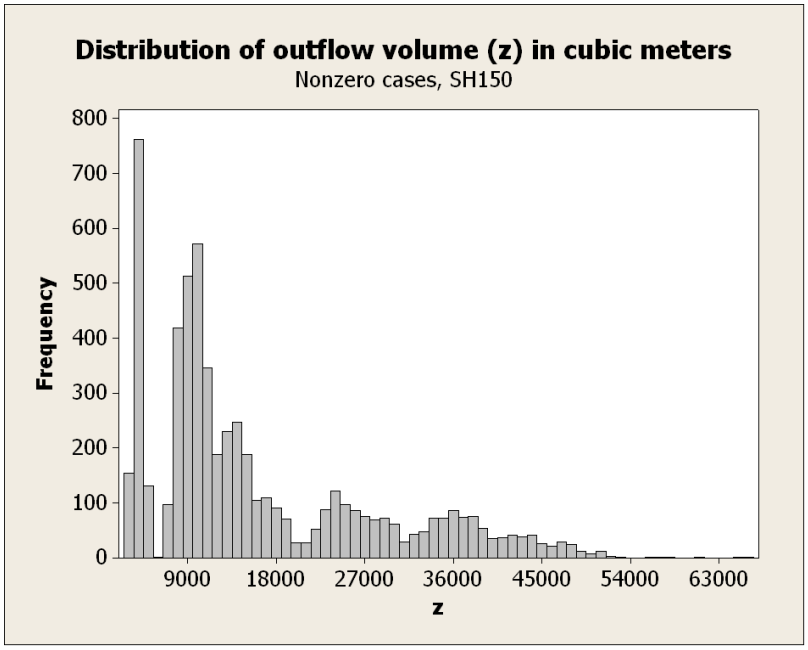


Figure 2.11: Total outflow volume histogram, groundings

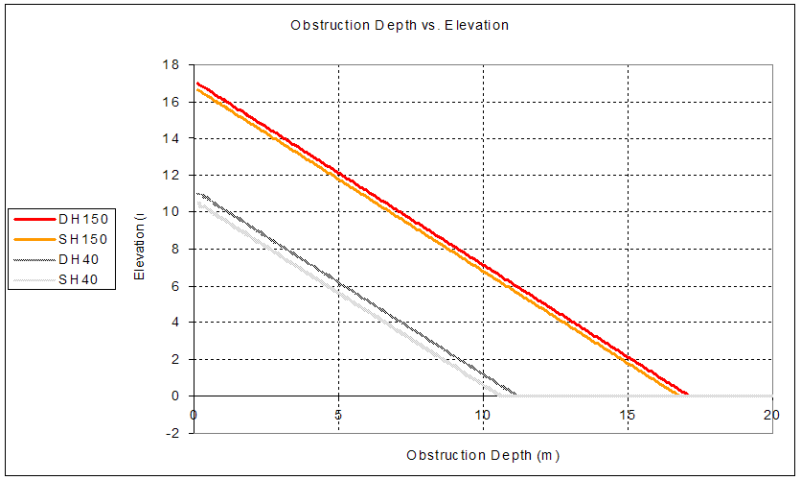


Figure 2.12: Scatterplot of obstruction depth vs. elevation, all ship types

Chapter 3

Collision Model

3.1 Overview

The simulated collision data discussed in Chapter 2 is used to construct a model that calculates outflow volume given a collision scenario. The essence of this model is to establish a relation between known input and output datapoints that are present in the given sample set, i.e. between velocity, collision angle etc. and oil outflow volume, so that outflow can be calculated for any given collsion scenario using these variables.

Just searching through a set of 40,000 datapoints is not practical; furthermore, if the specific scenario is not included in the 40,000 that were simulated, one would need to be able to interpolate between datapoints. A subsequent issue is that directly linking a set of input variables to outflow volume is not ideal. There are only a handful of different outflow values due to the assumption that all oil in a damaged compartment is lost; the limited number of compartments results in limited possible outflow outcomes. Also, in a high number of cases there is no outflow at all.

Since data on the size of the damaged area is available, as well as ship designs used in the simulations, it would be useful to include these aspects into the model.

3.1.1 Model Structure

The collision outflow model is ordered into sequential steps. Given the data obtained from collision simulations, the model should

1. calculate the damage extent to the struck ship given arbitrary scenario input variables;
2. determine the occurrence of rupture given damage extent;
3. calculate the oil spill volume given rupture.

Instead of a model that directly relates outflow volume to input variables, this one is not limited to the scenarios that were generated in the simulations. Also, it makes use of not only outflow data z but also damage data y_l and y_t. Furthermore, it makes use of the different types of data in a sequential fashion. Since data exists for four different ship types, four different collision models will be developed, each estimating the accidental outflow volume based on specific ship type -either single or double hull- and deadweight tonnage -either 40,000 or 150,000. Finally, combining simulation datasets results in generic models for single hull and double hull ships, i.e. models where the struck ship design is not fixed but defined by an additional variable. Thus in total, six models will be developed: four based on a particular design and two a combination of those.

In essence, this model allows interpolation between collision scenarios and between small and large ships of the same type (single hull or double hull).

Developing the outflow model requires several data analyses to be performed. Figure 3.1 gives a schematic overview of this model and the accompanying analysis in three sequential steps. It shows that the available simulation data is fed into different analytical methods in the analysis part (left); each of which is linked to a corresponding calculation method in the calculation part (right).

In the following sections discuss the choice of analytical methods and how they were performed.

3.1.2 Regression Analysis

The usual method of obtaining a relationship between sample sets is through regression analysis. The input variables are known as predictor- or independent variables; the output variable is called the response- or dependent variable. Analysis results in a regression model. Appendix A goes into more detail on various regression models.

3.1.3 Statistical vs. Practical Significance

Goodness-of-fit tests can be useful in determing whether it is suitable to fit a theoretical regression model to a dataset. However, these tests deal with statistical significance, while the practical significance of a model might be a more relevant issue:

> "The question is not whether the input model is absolutely correct; it is whether the input model is adequate for the analysis at hand. [...] The fallacy of the goodness-of-fit test is made obvious when a large real-world data set it fitted to many classical distributions and all are rejected; all are rejected because the large sample size yields large power and the error in the model is indeed statistically significant." [18]

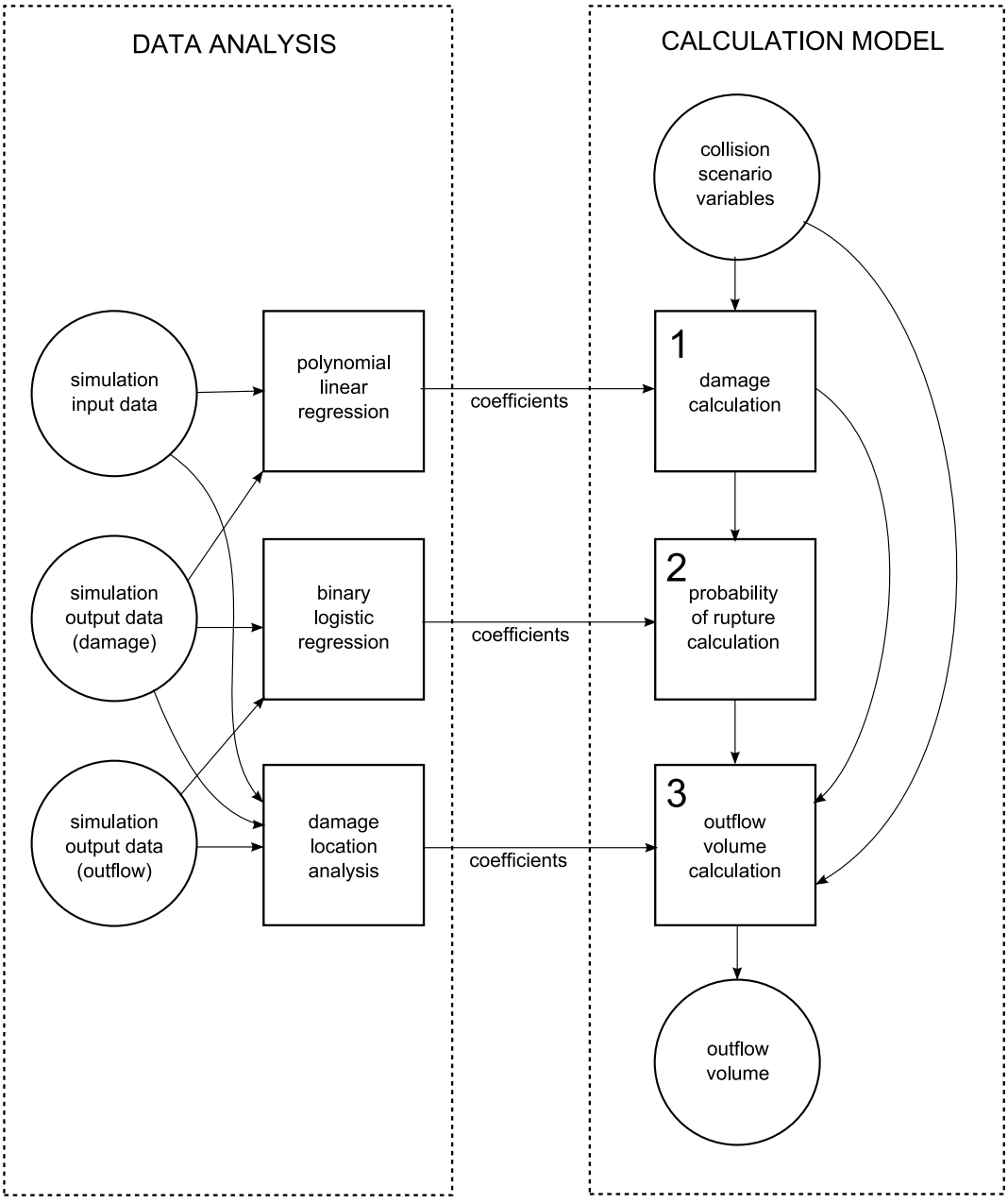


Figure 3.1: Collision outflow model overview

Because this report works with large datasets, validity of the models' significance is based on an "intuitive" judgment rather than statistical tests, although the latter will be taken into account.

3.2 Defining Predictor Variables

The input variables in the collision simulation sample (v_1, v_2, m_1, ϕ, l, t) could be directly used in regression; however, transforming them into other variables might result in a more natural, meaningful representation of a collision scenario. For example, a higher striking ship velocity (v_1) alone does not necessarily lead to a higher outflow probability or a larger damage area; this outcome also depends on the orientation of the striking ship against the struck ship (represented by collision angle ϕ). In this section, the predictor variables to be used in regression are obtained from the variables in the dataset.

Intuitively, when travelling at the same speeds, a heavy ship will release more kinetic energy in a collision than a light one; and and a fast-moving ship will release more kinetic energy than a slow-moving one with the same mass as the former. Therefore it is plausible that damage extent in a collision is related to kinetic energy. A relationship between dissipated energy in a collision and damage volume has been established empirically by Minorsky [15].

Important is the relative direction of motion. If two colliding ships travel in the same direction, less energy is released on collision than when going in the opposite direction. Also, since the striking ship collides under a certain angle, the inflicted damage varies depending on this angle. If it is very oblique, the striking ship will cause less damage than when it strikes perpendicular to the struck ship's longitudinal axis. Hence, it it critical that the energy variable(s) to be developed take into account relative velocities in the travelling direction of the struck ship and the collision angle to be effective in an analysis.

To accomodate this, a decomposition of kinetic energy into a tangential and perpendicular component is proposed.

Kinetic Energy

The kinetic energy of a body represents the amount of energy that is being released when this body is brought from a moving state to a full stop.

The total kinetic energy e_k of a system consisting of n separate masses $m_1, \ldots, m_n$ in a space is defined as

$$e_k = \sum_{j=1}^{n} \frac{1}{2} m_j v_j^2 \tag{3.1}$$

where v_j is the speed of m_j and $\vec{v}_j$ is the corresponding velocity vector with $v_j = \|\vec{v}_j\| = \sqrt{\langle \vec{v}_j, \vec{v}_j \rangle}$.

The coordinate system (x, y) used in the simulations is two-dimensional and defines the coordinate system's origin $(0, 0)$ as midship of the struck ship at the moment of collision. The struck ship, at that point, travels with speed v_2 in the positive x-direction; the striking ship moves towards the struck ship under an angle ϕ at speed v_1 (see also Figure 2.3). The corresponding velocity vectors are then:

$$\vec{v}_1 = (-v_1 \cos\phi, -v_1 \sin\phi) \tag{3.2}$$

$$\vec{v}_2 = (v_2, 0) \tag{3.3}$$

It is noteworthy that the y-components of the velocities are perpendicular to the struck ship's direction of motion, and that the x-components are tangential to it. Considering the ships as separate masses, total kinetic energy becomes

$$e_k = \frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 \tag{3.4}$$

This term can be decomposed into perpendicular and tangential components, $e_{k,p}$ and $e_{k,t}$, respectively:

$$e_{k,p} = \frac{1}{2} m_1 (v_1 \sin\phi)^2 \tag{3.5}$$

$$e_{k,t} = \frac{1}{2} m_1 (v_1 \cos\phi)^2 + \frac{1}{2} m_2 v_2^2 \tag{3.6}$$

It follows that $e_k = e_{k,p} + e_{k,t}$. However, this decomposition does not discriminate in relative direction of motion. If two ships collide at certain speeds and $\phi = 0°$, $e_{k,t}$ will have the same value as when they travel at the same speeds and $\phi = 180°$. In Figure 3.2, it can be seen that in the left situation, a lot less damage will be inflicted as opposed to the right situation because of the difference in tangential velocity, although $e_{k,t}$ as defined in Equation 3.6 stays the same. Therefore, a modified definition of tangential kinetic energy could be introduced:

$$e_{k,t} = \frac{1}{2} m_1 \kappa(\phi) (v_1 \cos\phi)^2 + \frac{1}{2} m_2 v_2^2 \tag{3.7}$$

where

$$\kappa(\phi) = \begin{cases} 1, & 0 < \phi \leq \frac{\pi}{2} \\ -1, & \frac{\pi}{2} < \phi \leq \pi \end{cases} \tag{3.8}$$

However, in that case $e_{k,p}$ and $e_{k,t}$ do not sum up to e_k when $\kappa(\phi) = -1$ and is thus not consistent with the kinetic energy formulation of a set of seperate bodies. From this argument, the notion arises that the difference in perpendicular and tangential velocities has to be taken into account. Consider the following:

$$e_k = \frac{1}{2} m_1 v_r^2 + \frac{1}{2} m_2 v_2^2 \tag{3.9}$$

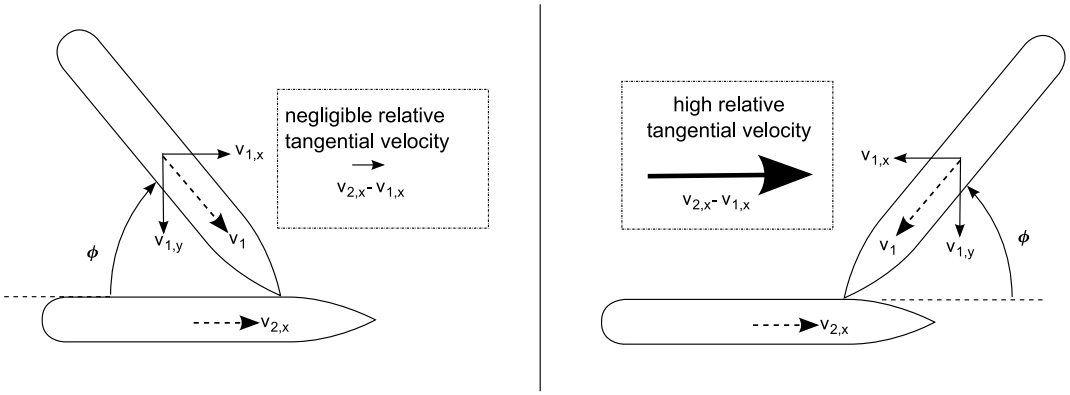


Figure 3.2: Tangential velocity difference

which might be decomposed into

$$e_{k,p} = \frac{1}{2} m_1 v_p^2 \tag{3.10}$$

$$e_{k,t} = \frac{1}{2} m_1 v_t^2 + \frac{1}{2} m_2 v_2^2, \tag{3.11}$$

where

$$\vec{v_r} = \vec{v_2} - \vec{v_1} = (v_2 + v_1 \cos\phi, v_1 \sin\phi) = (v_t, v_p)$$

is the velocity of the striking ship relative to the struck ship's velocity. However, consider again two ships travelling in the same direction with exactly the same speed. No collision damage will occur, but this decomposition will not accomodate that scenario.

Hence, it appears that interpreting the vessels as seperate bodies does not lead to a set of predictor variables with the desirable properties. To get a consistent decomposition of kinetic energy that holds up to the concepts mentioned at the beginning of this subsection, one should consider the two ships to represent a single mass at the exact moment of impact with a residual velocity that is the vector sum of the velocities of the individual vessels. Now imagine a measure of kinetic energy that represents the "collision kinetic energy", being the kinetic energy that can be released in a collision in perpendicular and tangential directions:

$$e_k = e_{k,p} + e_{k,t} = \frac{1}{2} m_{tot} v_r^2 \tag{3.12}$$

where

$$e_{k,p} = \frac{1}{2} m_{tot} v_p^2 \tag{3.13}$$

$$e_{k,t} = \frac{1}{2} m_{tot} v_t^2 \tag{3.14}$$

and $m_{tot} = m_1 + m_2$. It is important to mention that, using this kinetic energy model, two ships travelling in the same direction at the same speed will result in zero kinetic energy upon collision, regardless of their masses.

Location

The relative collision location l possibly has an influence on the ability to convert the perpendicular motion of the striking ship into rotational motion of the struck ship can be determined. If a collision occurs at the bow or stern, more kinetic energy is transformed into rotation of the struck ship around the vertical axis. When the collision instead occurs near midship, the struck ship is less able to transform perpendicular motion into rotation. A new variable l' is introduced that indicates how far a collision takes place from midship of the struck ship:

$$l' = \left| l - \frac{1}{2} \right| \tag{3.15}$$

Striking Ship Type

The striking ship type t determines the mass, dimensions and other parameters of the striking ship. t itself cannot be used as a predictor variable because it qualifies rather than quantifies a ship's characteristics ("type" cannot be measured whereas, for example, "mass" or "length" can). Since dimensions are directly related to mass [3], and since mass is already taken up in $e_{k,p}$ and $e_{k,t}$, the only variable that could further represent t is the bow half entrance angle η.

η affects the striking ship's ability to penetrate te struck ship. The sharper the angle, the higher the probability that the striking ship will penetrate the struck ship, and the further the striking ship will penetrate.

Combined Model Variable

In the combined collision models the single hull datasets (SH40, SH150) and double hull datasets (DH40, DH150) are combined into combined single hull and double hull datasets (SHCOM, DHCOM). These datasets are thus twice as long as the original ones and represent simulation data for a generic single hull or double hull ship. Because the variables in these sets do not present explicit information on the origin of the data -i.e., which dataset it belonged to originally- an additional variable will be added that improves the quality of the regression model. This variable, d, represents either the length or the width of the ship (depending on which dependent variable it is used on in regression, e.g. y_l or y_t). In a sense, it is an indicator variable, indicating ship type, but because it $d \in \mathbb{R}$ it can be used in regression among the other variables.

3.3 Transformation of Predictor Variables to CDF

Now there are four variables defining the input of a collision event for four types of tankers, and five variables for two combined tanker designs. Each set of predictor variables $(e^i_{k,p}, e^i_{k,t}, l'^i, \eta^i, d^i)$, for all $i \in \{1, \ldots, n\}$ can be seen as realizations of random variables $E_{k,p}, E_{k,t}, L', H$ and D. Their corresponding cumulative distribution functions (CDFs) are $F_{E_{k,p}}, F_{E_{k,t}}, F_{L'}, F_H$ and F_D. Instead of taking the predictor variables as they are, all realizations for each variable are transformed through their CDF values, resulting in the transformed predictor variables

$$\begin{aligned} x_{1,i} &= F_{E_{k,p}}(e^i_{k,p}), \\ x_{2,i} &= F_{E_{k,t}}(e^i_{k,t}), \\ x_{3,i} &= F_{L'}(l'^i), \\ x_{4,i} &= F_H(\eta^i), \\ x_{5,i} &= F_D(d^i), \\ & \quad \forall i \in \{1 \ldots, n\} \end{aligned} \tag{3.16}$$

The rationale behind this transformation step is as follows:

- The transformed variables are in the domain [0,1], increasing numerical stability in regression computations.
- The transformed variables are dimensionless, since a CDF typically represents the probability of an event. Any regression analysis performed on these variables will yield parameters that have the same dimension as the response variable.

Note that the CDFs for variables l' and η are the same in all collision models, even in the combined ones, but not $e_{k,p}$ and $e_{k,t}$ because the masses of the struck ships vary. Since d only plays a role in the combined models, it is not used in the other ones. Figure 3.3 gives an overview of the transformation steps converting the original variables to predictor variables to be used in the regression analysis.

3.3.1 CDFs of $E_{k,p}$, $E_{k,t}$

$E_{k,p}$ and $E_{k,t}$ are stochastic variables composed of other stochastic variables, as can be derived from Equations 3.13 and 3.14:

$$E_{k,p} = \frac{1}{2}(M_1 + m_2)V_1^2 \sin^2 \Phi \tag{3.17}$$

$$E_{k,t} = \frac{1}{2}(M_1 + m_2)(V_2 + V_1 \cos \Phi)^2 \tag{3.18}$$

Because of the complexity of these equations, it is difficult to find the exact distribution functions $F_{E_{k,p}}$ and $F_{E_{k,t}}$. An alternative would be to use the empirical CDFs

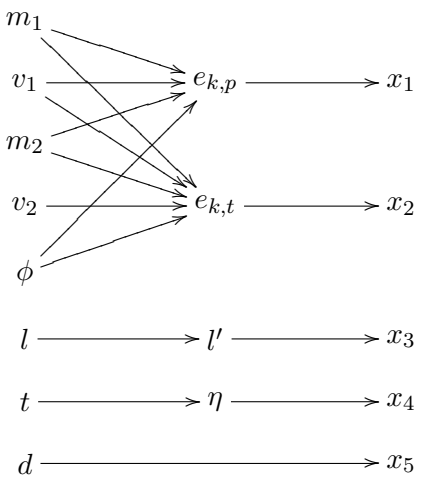


Figure 3.3: Transformation of input variables to predictor variables

of $E_{k,p}$ and $E_{k,t}$, which is found by looking at the distribution of the realizations of these random variables (see Appendix B). Since $n = 10,000$ and thus sufficiently large, the empirical CDFs for $E_{k,p}$ and $E_{k,t}$ would be excellent approximations for the real CDFs because of the strong limit properties of the empirical CDF.

Herein, however, also lies also a weak point: since there are $10,000$ realizations for $E_{k,p}$ and $E_{k,t}$, it would be cumbersome to implement their empirical CDF in the application of the outflow model: each time it is invoked, up to 10,000 values have to be looked up from a table containing the realizations, which will lengthen the run time of a application using the model significantly and makes the model higly unportable, i.e. these values have to be stored somewhere.

Therefore, it's better to find a parametric fit to the empirical CDF, which, in the case of a closed-form parametric fit, would require a calculation time that is magnitudes less than using empirical CDFs. A parametric CDF to fit a random variable X is denoted by $F_X(x|\alpha)$, where α is a set of parameters that define the function's characteristics. For $E_{k,p}$ and $E_{k,t}$, numerous options exist for a parametric distribution. The Weibull distribution (see Appendix B) does a good job, is only nonnegative, is closed-form and is shaped by two parameters instead of 10,000 realizations of random variables.

Fits for $F_{E_{k,p}}$ and $F_{E_{k,t}}$ were generated using Minitab: see Figure 3.4 for a comparison between the Weibull and empirical CDF of perpendicular kinetic energy in the SH40 case, and a probability plot that shows how well the data aligns with the fit. In Table 3.1 the coefficients for all Weibull fits are given.

The drawback to using the Weibull fit is that the p-value for the Anderson-Darling

		SH40	SH150	SHCOM	DH40	DH150	DHCOM
$E_{k,p}$	α	0.4699	0.4724	0.4515	0.4699	0.4724	0.4514
	β	320.3	1010	590.0	319.8	1010	589.4
$E_{k,t}$	α	0.4546	0.4567	0.4379	0.4546	0.4567	0.4378
	β	385.7	1217	709.9	385.1	1217	709.1

Table 3.1: Coefficients for Weibull fits, kinetic energy

test[1] is very low, which essentially means that the use of the parametric CDF as a fit for the empirical CDF has to be discarded. However, because the number of datapoints is so high, the margin of acceptance becomes extremely narrow and it is unlikely that any parametric fit would be accepted. For practical reasons, judging a fit by 'visual' goodness-of-fit trumps the statistical test (as discussed in Section 3.1.3). In that view, the Weibull distribution is accepted. Alternative parametric distributions, such as Gamma, Exponential (which is a special case of the Weibull family) and Logistic have significantly worse fits (see Figure 3.5).

3.3.2 CDF of L'

Given that $L \sim Beta(1.25, 1.45)$, Equation 3.19 returns the exact distribution of L' which was defined as $L' = |L - \frac{1}{2}|$. See figure 3.6 for a graph of $F_{L'}$.

$$\begin{aligned} F_{L'}(x) &= P(L' \leq x) \\ &= P(\left|L - \frac{1}{2}\right| \leq x) \\ &= P(-x \leq L - \frac{1}{2} \leq x) \\ &= P(\frac{1}{2} - x \leq L \leq \frac{1}{2} + x) \\ &= P(L \leq x + \frac{1}{2}) - P(L \leq -x + \frac{1}{2}) \\ &= F_L(x + \frac{1}{2}) - F_L(-x + \frac{1}{2}) \end{aligned} \tag{3.19}$$

3.3.3 CDF of H

Since H only takes on three possible values, namely 17, 20 and 38 degrees, the best transformation is the empirical CDF, which is given in Table 3.2 and Figure 3.3.3.

[1]The Anderson-Darling test puts up two hypotheses: one saying that the data follows the specified distribution (in this case Weibull), and one saying that it doesn't. A p-value below a certain level of signifcance, here 0.05, pleads for the latter hypothesis.

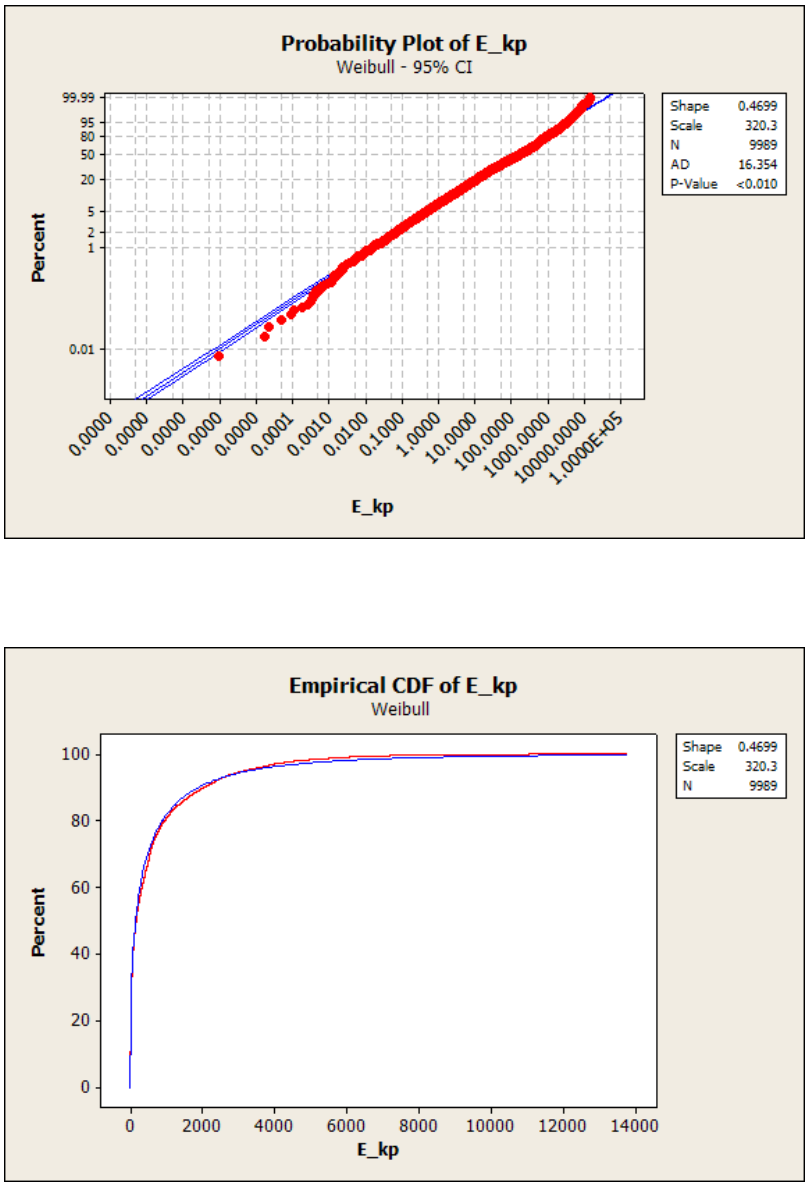


Figure 3.4: Probability plot & Weibull fit of empirical CDF, perpendicular kinetic energy, SH40 case

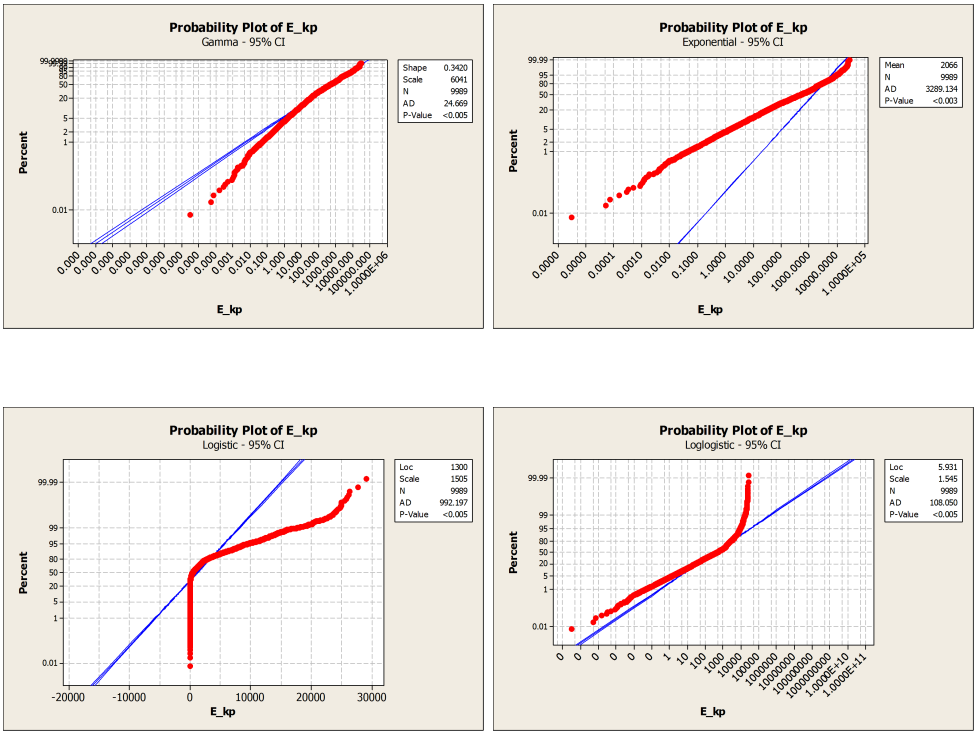


Figure 3.5: Probability plots of alternative parametric fits, perpendicular kinetic energy, SH40 case

η	Count	$F_H(\eta)$
17	2440	0.2440
20	5323	0.7763
38	2236	1.0000

Table 3.2: Empirical CDF of H

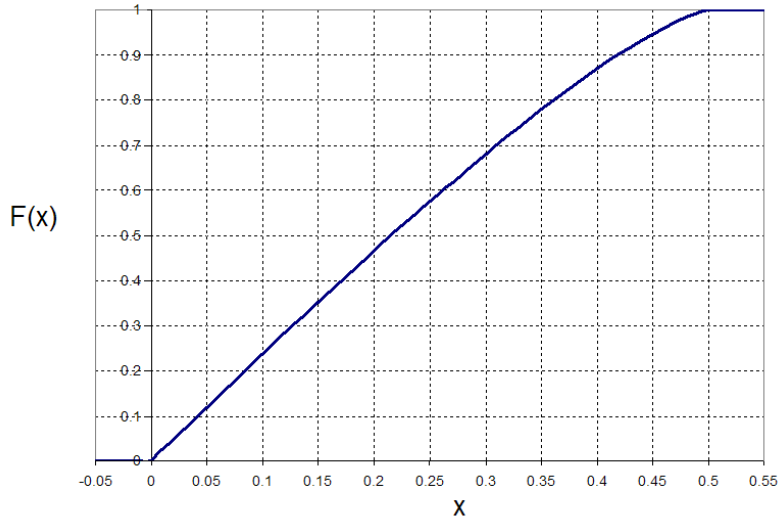


Figure 3.6: Cumulative distribution function for L'

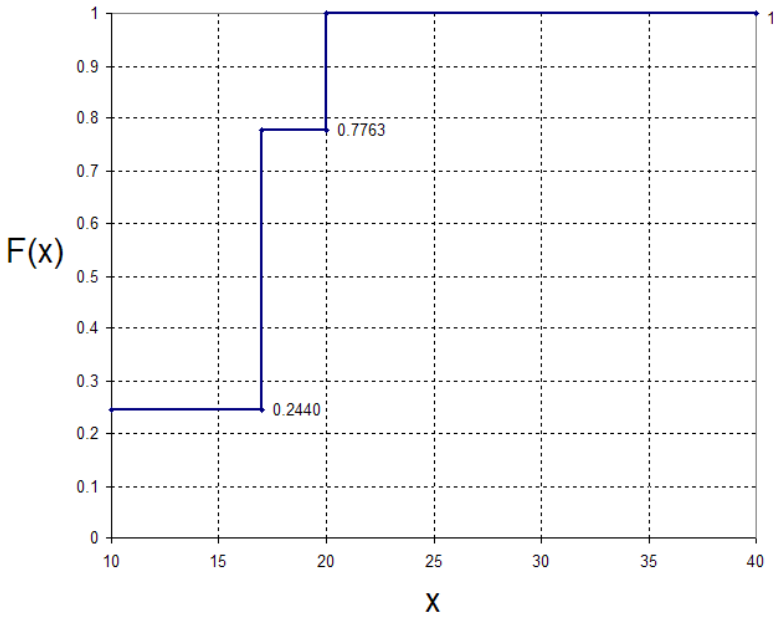


Figure 3.7: Empirical CDF of H

3.4 Damage Extent

Now that the predictor variables have been defined, it is time to look at the effect they have on damage extent. Damage extent is measured by two parameters: y_l and y_t, or damage length and maximum penetration (the damage is assumed to extend vertically along the entire depth of the ship). Assume that y_l^i and y_t^i are realizations of random variables Y_l and Y_t. Given a set of predictor variables $\mathbf{x}$, the goal is to give an estimate of Y_l and Y_t:

$$Y_l = h_l(\mathbf{x}) + R_l \tag{3.20}$$
$$Y_t = h_t(\mathbf{x}) + R_t, \tag{3.21}$$

where the functions h_l and h_t give a conditional expected value for Y_l and Y_t and R_l and R_t are random variables that give the variation in Y_l and Y_t that cannot be "explained" by $\mathbf{x}$. In linear regression, h_l and h_t are estimated by a set of coefficients $\beta = (\beta_0, \ldots, \beta_5)$:

$$E(Y_l|\mathbf{x}) = h_l(\mathbf{x}|\beta^l) = \beta_0^l + \beta_1^l x_1 + \ldots + \beta_5^l x_5 \tag{3.22}$$
$$E(Y_t|\mathbf{x}) = h_t(\mathbf{x}|\beta^t) = \beta_0^t + \beta_1^t x_1 + \ldots + \beta_5^t x_5 \tag{3.23}$$

Regression analysis on the datasets $\{(\mathbf{x}_1, y_l^1), \ldots, (\mathbf{x}_n, y_l^n)\}$ and $\{(\mathbf{x}_1, y_t^1), \ldots, (\mathbf{x}_n, y_t^n)\}$ yields the models

$$\hat{h}_l(\mathbf{x}) = h_l(\mathbf{x}|\hat{\beta}^l) \tag{3.24}$$
$$\hat{h}_t(\mathbf{x}) = h_l(\mathbf{x}|\hat{\beta}^t) \tag{3.25}$$

where $\hat{\beta}^l$ and $\hat{\beta}^t$ are found by minimizing the sum of squared residuals over β^l and β^t:

$$\min_{\beta^l} \sum_{i=1}^{n} (y_i^l - h_l(\mathbf{x}_i|\beta^l))^2$$
$$\min_{\beta^t} \sum_{i=1}^{n} (y_i^t - h_t(\mathbf{x}_i|\beta^t))^2$$

(See Appendix A for a concise discussion about linear regression.)

Linear regression for Y_l and Y_t might not be adequate, because this would assume the fitting of a flat slope through the data whereas the data shows a more curved behaviour. For example, take the SH150 case. From Figure 3.8, it can be seen that there is a strong nonlinear relationship between y_t and x_1 (the CDF of perpendicular kinetic energy): instead of a straight line, a nonlinear curve would describe this relation more accurately. Therefore linear regression is expanded to polynomial linear regression to accomodate for curve fitting: besides $x_1, \ldots, x_5$, their powers

(up to a certain order) are introduced as predictor variables, giving an extended set of variables

$$\mathbf{x} = \begin{matrix} (x_1,\ldots,x_5, \\ x_1^2,\ldots,x_5^2, \\ \vdots \\ x_1^p,\ldots,x_5^p) \end{matrix} \tag{3.26}$$

Where p is the polynomial order. Note that polynomial linear regression is the same as linear regression: the solution is linear in the coefficients β^l and β^t of h_l and h_t, respectively. Polynomials were chosen because of their flexibility as a nonlinear function and because they are easy to integrate into linear regression.

What has to be noted is that the variation in y_t is small for low values of x_1 and large for high values of x_1, which is an undesirable effect. However, when transforming y_t by taking the natural logarithm, residual variation is much more constant: see Figure 3.9. Other transformations, such as taking the root, are possible as well. The natural logarithm is chosen typically to remove heteroscedasticity in residual performance which it achieved in this case; moreover, it gives reasonable regression fits.

Because $\ln(0)$ does not exist, all zero values of y_l and y_t are removed from the dataset. From now on in this section, the datasets $(\mathbf{x}_j, y_l^j)$ and $(\mathbf{x}_j, y_t^j)$ are used, with

$$j \in J \subset \{1,\ldots,n\}$$

such that $y_l^j > 0$ and $y_t^j > 0$ for all $j \in J$.

In Minitab, the linear regression for $\ln y_l$ and $\ln y_t$ is performed in three steps:

- First, a stepwise regression algorithm sequentially adds and deletes variables until a suitable set of predictor variables is obtained. The algorithm inserts variables based on a statistical significance test that requires an assumption of normality of the residuals. This technique is commonly applied even though the algorithm does not test for normality of residuals.
- After a set of candidate variables have been determined by the stepwise regression a best subset regression is performed on this set of variables. A best subset regression algorithm determines which superfluous variables can be removed from the previously obtained set without compromizing its quality, resulting in a best subset of variables. The removal of variables from subsets is heuristically determined by looking at each subset's Mallows' Cp-value, which indicates possible overfitting of a regression model. Mallow's Cp allows the residual distribution to be nonnormal for this method to work. (Alternatively, it would be possible to remove variables based on significance testing, but this

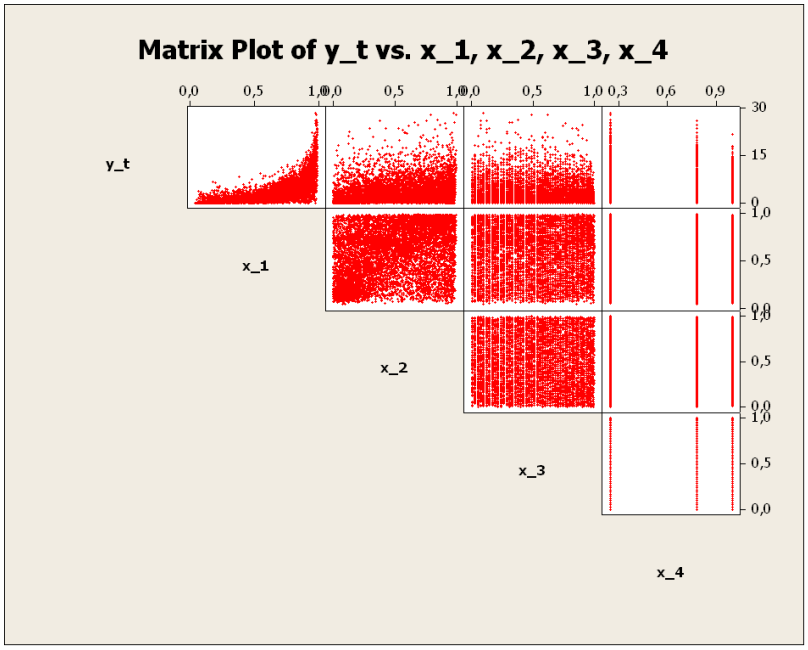


Figure 3.8: Matrix plot of y_t against $\mathbf{x}$, SH150 case

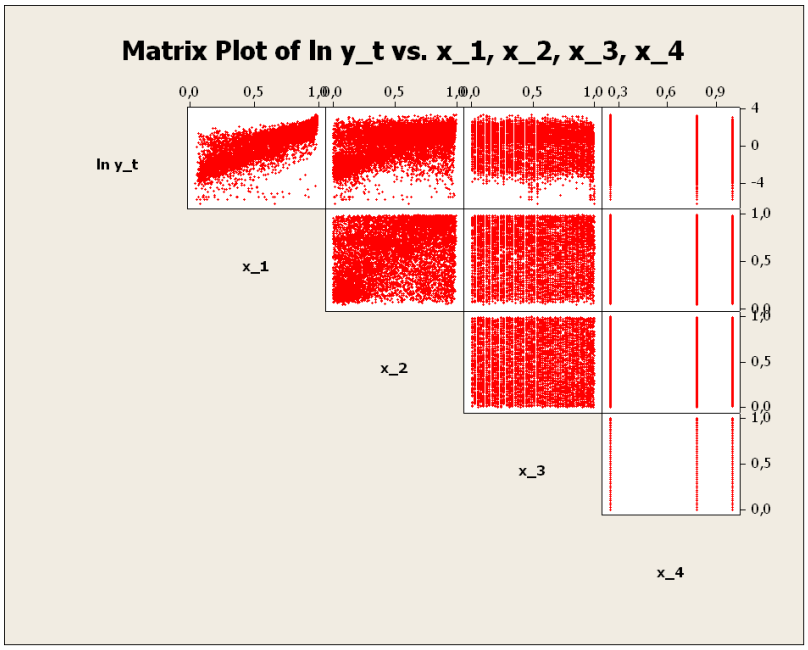


Figure 3.9: Matrix plot of $\ln y_t$ against $\mathbf{x}$, SH150 case

assumes normality of residuals.) A widely accepted approach is that subsets with N variables are suitable for regression when $N < Cp < 2N$, of which the subset with lowest number of variables is chosen [2].

- Third, linear regression analysis is done using this reduced best subset of variables, resulting in coefficients $\hat{\beta}$. Now, the multiple polynomial functions $\hat{h}_l$ and $\hat{h}_t$ express the expected value of $\ln Y_l$ and $\ln Y_t$ conditioned on the set of input variables $\mathbf{x}$:

$$\begin{aligned} h_l(\mathbf{x}|\hat{\beta}^l) &= \hat{\beta}^l_0 + \hat{\beta}^l_{1,1}x_1 + \ldots + \hat{\beta}^l_{1,5}x_5 + \\ &\quad \hat{\beta}^l_{2,1}x_1^2 + \ldots + \hat{\beta}^l_{2,5}x_5^2 + \\ &\quad \ldots + \\ &\quad \hat{\beta}^l_{p,1}x_1^p + \ldots + \hat{\beta}^l_{p,5}x_5^p \end{aligned} \tag{3.27}$$

In this study, $p = 5$ was chosen. The set of coefficients $\hat{\beta}^l$ and $\hat{\beta}^t$ for h_l and h_t, resulting from the regression analysis, can be found in Tables D.1 and D.3.

Correlation between Predictor Variables

When variables are correlated, some problems might appear that affect the overall robustness of a regression analysis. But even with very strong correlation (or multicollinearity) between predictor variables, the predictive value of the regression model may still be good as long as predictions are based on combinations of these variables [13]. The correlation matrix between $x_1, \ldots, x_4$ is as follows in the SH150-case:

	x_1	x_2	x_3
x_2	0.30		
x_3	0.01	−0.01	
x_4	0.02	−0.03	−0.02

There is only some positive correlation between x_1 and x_2 (as could be expected, since they are the CDFs of perpendicular and kinetic energy, which share some common variables such as speed and mass). Therefore, one should be cautious when using the coefficient estimates to explain the individual effects that their corresponding variables have on damage extent.

Since powers of the predictor variables have been used as variables in the polynomial linear regression, there is inevitable correlation between higher and lower powers. This is only problematic for x_5, which only takes 2 values: the CDF values of ship length (or width) distribution. x_5^2 is exactly collinear with x_5 and leads to a division by zero somewhere in the regression analysis. Minitab resolves these issues by means of notification during the regression process.

3.4.1 Fitting Residual Distribution

Now that the conditional expected value of $\ln Y_l$ and $\ln Y_t$ is known, the set of residuals can be used to model the randomness of the data. The residuals r_l and r_t are defined as

$$r_l^j = h_l(\mathbf{x}_j) - \ln y_l^j, \quad \forall j \in J \tag{3.28}$$
$$r_t^j = h_l(\mathbf{x}_j) - \ln y_t^j, \quad \forall j \in J \tag{3.29}$$

These sets can be seen as realizations of random variables R_l and R_t, respectively. These variables are typically assumed to have a Normal distribution with mean 0; this, however, is not a requirement of least squares estimation; in this case even, a Normal distribution would not fit as can be seen from the resisual plots and histograms in Figures 3.10 and 3.11. To this end, an alternative parametric distribution is introduced: the Generalized Trapezoidal distribution (see Appendix B.4). This distribution is fitted to the empirical CDFs of R_l and R_t by means of least squares. Because the distribution function is nonlinear in its coefficients, the least squares fit is approximated numerically. These coefficients are displayed in Table D.2 and D.4 in the Appendix.

The upper bound for the support of these distributions were found by determining the highest possible value of $\ln y_l$ and $\ln y_t$, which are restricted by respectively the length and width of the tanker types involved. Since $\ln y_l$ and $\ln y_t$ have no lower bound, the lower bounds for the GT distribution were determined by taking the difference between the highest and lowest residual value found and substracting this from the lowest residual value.

The quality of the fit can be measured by looking at the plot of the empirical CDF against the fitted CDF (see Figure 3.12). When this plot is close enough to the centerline (going from $(0, 0)$ to $(1, 1)$ in the graph) then the fit is a good representation of the actual CDF of the random variable.

As can be observed, this is a very close fit; all other plots are similarly close to the centerline.

3.5 Probability of Rupture

The next step is to relate this damage extent to the outflow volume, or rather the occurrence of outflow. It is assumed that zero outflow ($z = 0$) implies no rupture. Since occurrance of rupture this is a binary event (it either happens or it doesn't) the model should yield a measure of how likely rupture occurs, i.e. a probability of rupture. Binary regression analysis on the dataset $(y_l^i, y_t^i, z_i), i \in 1, \ldots, n$ will yield an expected probability of rupture conditioned on damage extent.

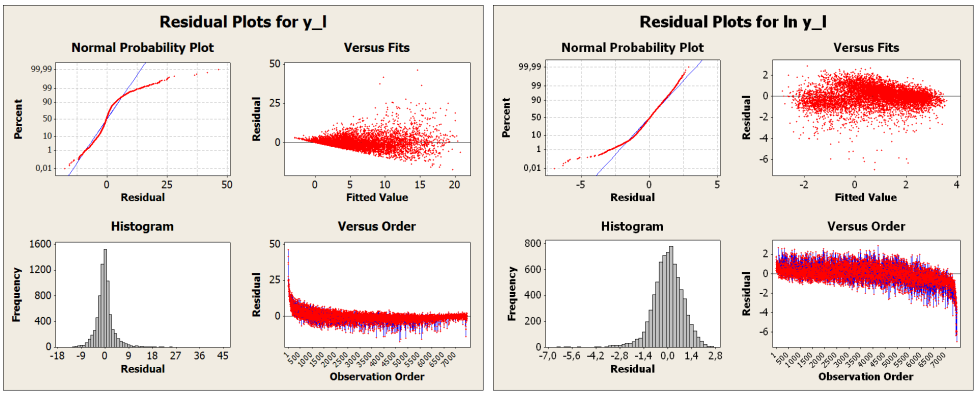


Figure 3.10: Residual plots for y_l resp. $\ln y_l$, SH150 case

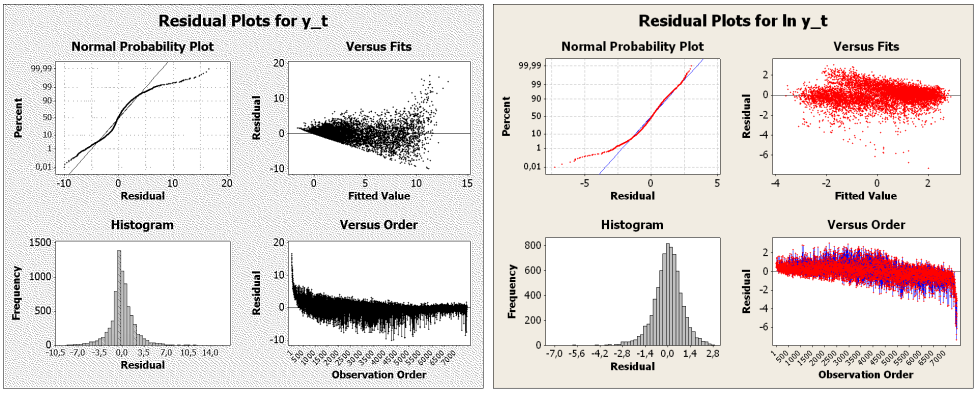


Figure 3.11: Residual plots for y_t resp. $\ln y_t$, SH150 case

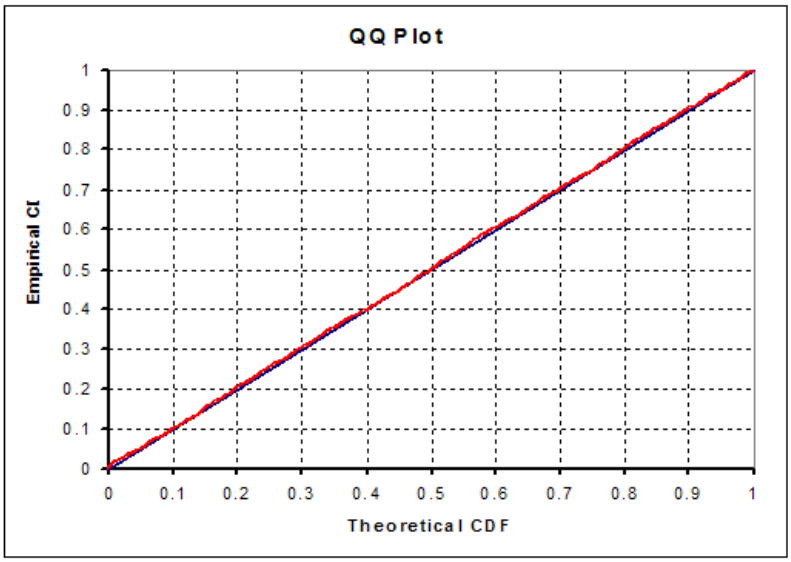


Figure 3.12: QQ-plot for the fit of residuals of $\ln y_t$, SH150 case

3.5.1 Binary Logistic Regression

Suppose the random variable Z expresses the outflow volume in a collision scenario. The following variable is introduced:

$$Z' = 1_{(0,\infty)}(Z) = \begin{cases} 1, & Z > 0 \\ 0, & Z = 0 \end{cases} \tag{3.30}$$

In other words, if outflow occurs, $Z' = 1$, otherwise $Z' = 0$. Again, by assumption, $Z' = 1$ means that rupture occurs. A binary logistic regression analysis (see Appendix Chapter A) can now be done on this variable against variables y_l and y_t. This analysis leads to coefficients that will be used in calculating the probability of rupture (which is the expected value of rupture occurrence $E(Z')$) in the outflow model.

However, since that calculation step comes after calculating damage extent (step 1), and since in the outflow model step 1 yields $\ln y_l$ and $\ln y_t$, the binary logistic regression will be done using the natural logarithms of damage length and maximum penetration.

Note that the logarithms of observed datapoints are used, not expected values calculated in Step 1 of the collision model. This results in a more accurate analysis in the sense that an estimation error in the first step (polynomial linear regression) does not propagate into the binary logistic regression.

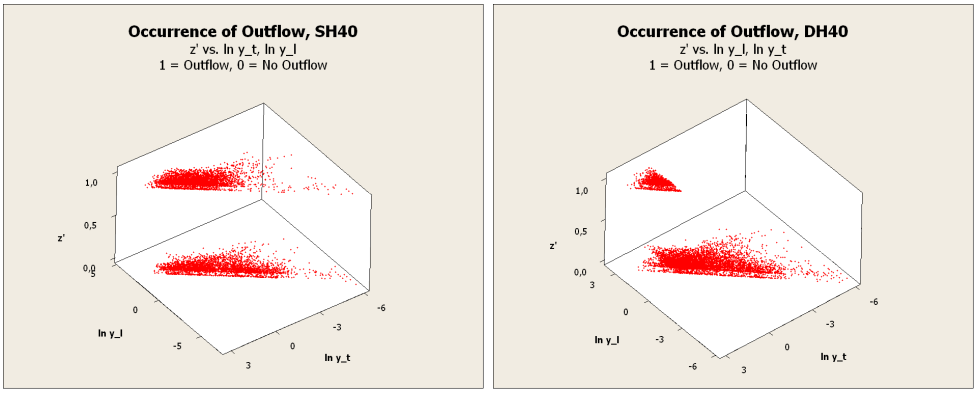


Figure 3.13: Scatterplot of z' against $\ln y_l$ and $\ln y_t$, SH40 case (left) and DH40 case (right)

The regression model is expressed as follows:

$$\begin{aligned} E[Z'|\ln y_l, \ln y_t] &= \pi(\ln y_l, \ln y_t|\beta) \\ &= \frac{\exp(\beta_0 + \beta_l \ln y_l + \beta_t \ln y_t)}{1 + \exp(\beta_0 + \beta_l \ln y_l + \beta_t \ln y_t)} \end{aligned} \tag{3.31}$$

It would have been possible to do binary logistic regression of Z' against predictor variables $x_1, \ldots, x_5$, i.e. the transformed variables used in determining $\ln y_l$ and $\ln y_t$ in the previous section. However, this would mean reusing the same data again and discard the information present in y_l^i and y_t^i.

In Figure 3.13, occurrence of outflow (z') is plotted against $\ln y_l$ and $\ln y_t$ for SH40 and DH40 tanker types, respectively. Note that in the single hull case, outflow occurs when damage extent is less severe than in the double hull case. From these figures it can be observed that $\ln y_l$ and $\ln y_t$ are interdependent. This means that any significance test on either one of these variables will be highly influenced by this interdependency, and thus no results from these tests may be used to discard either $\ln y_l$ or $\ln y_t$ from the binary logistic regression model.

The logistic function was chosen because it supports the behavior present in the data. Its range is between 0 and 1, which is essential because it represents a probability, and is monotonic (changing a predictor variable in a certain direction will either increase or decrease the logistic function), This fits the data as the number of outflow occurrences does not decrease when $\ln y_l$ or $\ln y_t$ go up.

3.5.2 Validity of Binary Logistic Model

QQ-plot

Is the binary logistic regression analysis worth the effort–does it provide enough information given the outflow data? Or would it be easier and simpler to determine the occurrence of outflow (0 or 1) by chance? In other words, it has to be determined if the resulting binary logistic model is different from a purely random model, i.e. a model where an alternative oil outflow variable Z'_{RND} is Bernouilli distributed with parameter p:

$$P(\{Z'_{RND} = 1\}) = p \tag{3.32}$$
$$P(\{Z'_{RND} = 0\}) = 1 - p, \tag{3.33}$$

where

$$p = \frac{\#\text{ outflow events}}{\#\text{ events}} \tag{3.34}$$

This hypothesis is tested by looking at the residuals of the expected probabilities with the outflow data versus the residuals of the expected probabilities with the randomly generated data. Two sets of residuals are determined from the binary logistic regression above, $\{r_{OUT,i}\}$ and $\{r_{RND,i}\}$:

$$r_{OUT,i} = z'_i - \hat{\pi}(\mathbf{x}_i), \quad i \in \{1, \ldots, n\} \tag{3.35}$$
$$r_{RND,i} = z'_{RND,i} - \hat{\pi}(\mathbf{x}_i), \quad i \in \{1, \ldots, n\} \tag{3.36}$$

Now, consider the empirical cumulative distribution functions of both residuals:

$$F_{OUT}(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,x]}(r_{OUT,i}) \tag{3.37}$$

$$F_{RND}(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,x]}(r_{RND,i}) \tag{3.38}$$

Both CDFs are set out against each other in a so-called QQ-plot (see Figure 3.14). If the plot does not diverge significantly from the centerline, one may conclude that the regression model concurs with both the available outflow data as with a randomly generated set of outflows. In other words, the BLR model then gives little information on whether the predictor variables, such as perpendicular kinetic energy, are significant in determining oil outflow. It would then be perfectly valid to determine the occurrence of outflow by chance. As can be seen from the figure, this is not the case.

It is quite possible that this methodology could be developed into a formal statistical hypothesis test, i.e. how close would the QQ-plot have to be to the centerline where one would say that the model doesn't distinguish between "real" data and randomly generated data?

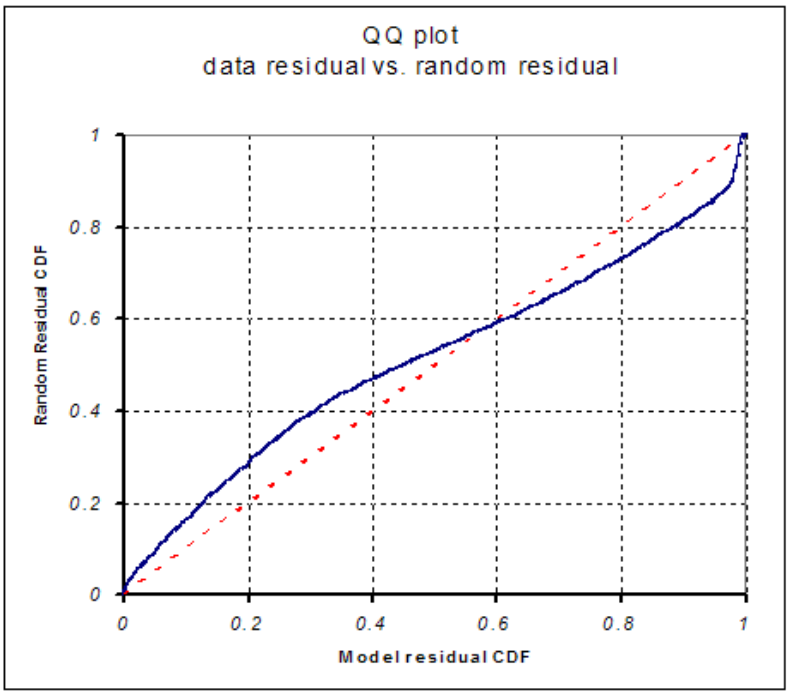


Figure 3.14: QQ-plot of probability residuals, SH150 case, collisions

Point-Biserial Correlation Coefficient

For now, the formal statistical model used to determine if the model should be rejected is the point biserial correlation coefficient r_{pb} using "real" occurence of outflow data and randomly generated data. r_{pb} determines correlation between a continuously measured variable (expectation of outflow Z', as calculated in the binary logistic regression) and a dichotomous variable (the actual occurence of outflow values z'):

$$r_{pb} = \frac{M_1 - M_0}{s_n}\sqrt{\frac{n_1 n_0}{n^2}}, \tag{3.39}$$

where

$$s_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z'_i - \bar{z}')^2}, \tag{3.40}$$

is the standard deviation of z', n_1 and n_0 are the number of occurrences of 1 and 0 in z', respectively, and M_1, M_0 are the mean values of Z' conditioned on the value of z' (either 1 or 0, respectively).

The statistic for assessing the significance of r_{pb} is

$$t = r_{pb}\sqrt{\frac{n_1 + n_0 - 2}{1 - r_{pb}^2}}.$$

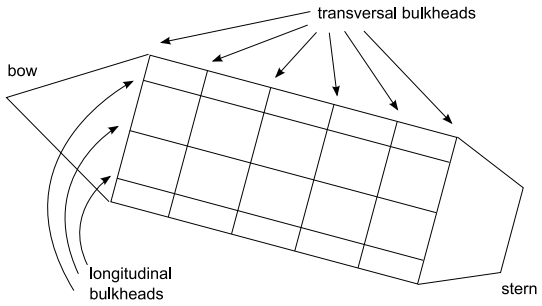


Figure 3.15: Bulkhead placement

If $P(T > t) < \alpha$, where T follows an unpaired Student's t-distribution with n_1+n_0-2 degrees of freedom, then the null hypothesis is rejected, i.e. the binary logistic model should be accepted.

The same thing can be done with random data: z' is then replaced by z'_{RND} which is generated in the same way as with the QQ-plot methodology.

The p-values for these tests (random an non-random) can be found in Table D.6.

3.6 Outflow Volume

Based on damage length, maximum penetration and collision location, the last section of the model involves calculating the oil outflow volume given that penetration has occurred and damage length and maximum penetration have been calculated.

3.6.1 Determining Damaged Area

As opposed to the original simulation, the model makes the assumption that the damaged area is a rectangular volume. Its longitudinal and transversal dimensions determined respectively by damage length (y_l) and maximum penetration (y_t). It is also assumed that damage occurs over the entire vertical extent of the ship, so this has no influence in the outflow volume. Furthermore, each compartment that coincides with the damaged area is assumed to lose all its oil. This differs from the original simulations, where the damaged area is not necessarily rectangular (see Figure 2.4).

For all four struck ship models, compartment configurations are available in the form of transverse and longitudinal bulkhead coordinates and compartment volumes. A schematic of one of these configurations is given below in Figure 3.15. Table C.9 in the Appendix gives the bulkhead coordinates.

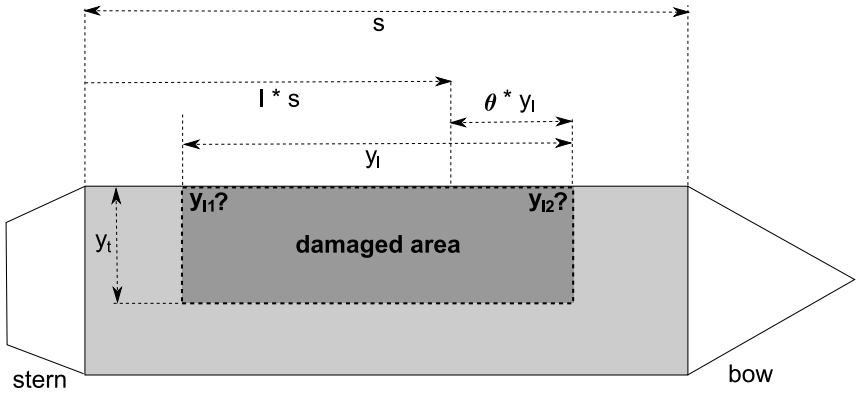


Figure 3.16: Collision location (l) and damage length (y_l) are known, start and end position (y_{l1}, y_{l2}) are unknown.

Determining Longitudinal Bounds

In each accident scenario, the longitudinal position of the damaged area is determined by the relative collision location l. However, neither a starting coordinate nor ending coordinate are present in the output data. Therefore these coordinates y_{l1}, y_{l2} have to be calculated by using ship length s, damage length y_l and a weight θ (see also Figure 3.16):

$$y_{l1} = (1-\theta)y_l + (1-l)s, \tag{3.41}$$

$$y_{l2} = -\theta y_l + (1-l)s, \tag{3.42}$$

$$\theta \in [0,1]$$

y_{l1} and y_{l2} are measured from the forward perpendicular because all bulkhead locations are given from this point as well. If $\theta = 0$, then all longitudinal damage is behind the collision location as measured from the forward point. If $\theta = \frac{1}{2}$, then the collision location is in the middle of the longitudinal damage. If $\theta = 0$, then all longitudinal damage is in front of the collision location.

By taking original datapoints (l_i, y_l^i, y_t^i), and calculating y_{l1} and y_{l2} for each i using a particular θ, one can also calculate which compartments have been breached and hence the total oil outflow $\tilde{z}_i$. If this outflow differs from the outflow value in the original data (z_i), then the model is incorrect. Since the assumption holds that no outflow implies no rupture, only cases where positive outflow occurs are taken into account.

Counting the fraction q of correct cases for all datapoints is a metric for assessing the quality of θ. Additionally, the average absolute error of outflow $\frac{1}{n}\sum|\tilde{z}_i - z_i|$ and conditional average absolute error of outflow can be assessed to this end.

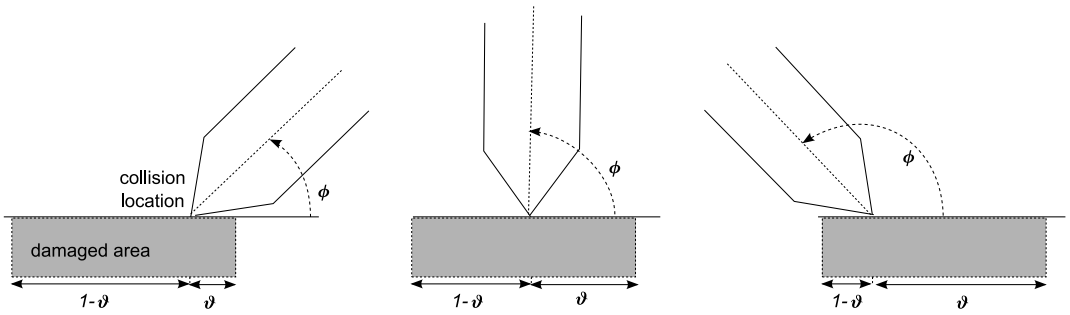


Figure 3.17: Determining position of damage location

The former measures the average error over all assessed cases, even if $|\tilde{z}_i - z_i| = 0$. The latter conditions on cases where $|\tilde{z}_i - z_i| > 0$. The goal is to find a suitable model for θ, and then optimize that model by maximizing q.

One can imagine a simple model:

$$\theta = \frac{1}{2} \tag{3.43}$$

i.e. in any situation, collision location will lengthwise always be in the middle of the longitudinal damage. However, when the collision angle is very oblique, the striking ship will probably cause the most longitudinal damage on one side of the collision location. Therefore the following model for is introduced as a function of collision angle ϕ (in degrees):

$$\theta = \frac{\phi}{180} \tag{3.44}$$

In short, if ϕ is near 0 degrees, longitudinal damage extends backwards of the collision location; if $\phi = 90$, the collision location is in the middle of longitudinal damage; if ϕ is near 180, then longitudinal damage extends forward of the collision location. In Figure 3.17 some examples are shown to clarify this model.

The proposed function is linear in ϕ, but an S-shape could be more appropriate as one would think that collision location stays close to one end of the longitudinal damage when $\phi < 90$ and close to the other end when $\phi \geq 90$. Therefore one might introduce an extra parameter n that describes this nonlinear behaviour:

$$\theta(\phi; n) = \begin{cases} 0, & \phi = 0 \\ \frac{1}{2}(\frac{\phi}{90})^n, & 0 < \phi < 90 \\ 1 - \frac{1}{2}(\frac{180-\phi}{90})^n, & 90 \leq \phi < 180 \\ 1, & \phi = 180 \end{cases} \tag{3.45}$$

Note that this model includes the previous models. If $n = 0$, then $\theta = \frac{1}{2}$. If $n = 1$, then $\theta = \frac{\phi}{180}$. For $n < 0$, θ will have a very unusual if not unrealistic profile, so this

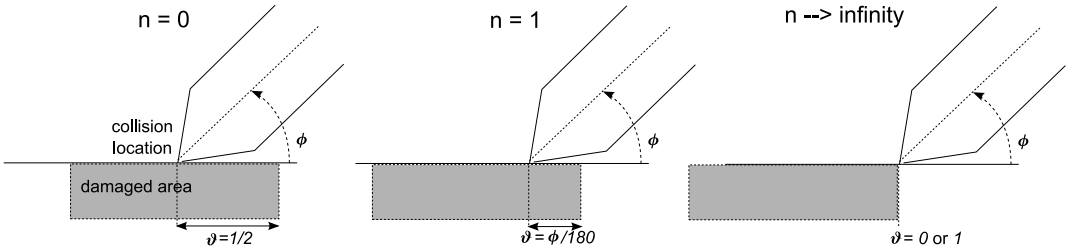


Figure 3.18: Determining position of damage location with added parameter

possibility is discarded. If $n \to \infty$, then

$$\lim_{n\to\infty} \theta(\phi; n) = \begin{cases} 0, & 0 \leq \phi < 90 \\ 1, & 90 \leq \phi \leq 180 \end{cases} \tag{3.46}$$

Some profiles of θ for different values of n are shown in Figure 3.18.

Finally, one could argue that relative tangential velocity v_t plays a role in determining where longitudinal damage occurs relative to the collision location. If $v_{1,x}$ and $v_{2,x}$ are the x-components of the striking and struck ships' velocities, respectively, then $v_t = v_{1,x} - v_{2,x}$. If the striking ships moves faster than the struck ship in the direction of the struck ship, then $v_t \geq 0$; if the striking ship moves slower in that direction, then $v_t \leq 0$. The direction of v_t should be a factor in the location of longitudinal damage. So, to integrate relative velocity into θ, the following model is proposed:

$$\theta(\phi, v_t; m, n) = \begin{cases} 0, & \phi = 0 \\ (\frac{1}{2}(\frac{\phi}{90})^n)^{\exp(mv_t)}, & 0 < \phi < 90 \\ (1 - \frac{1}{2}(\frac{180-\phi}{90})^n)^{\exp(mv_t)}, & 90 \leq \phi < 180 \\ 1, & \phi = 180 \end{cases} \tag{3.47}$$

m determines how much influence v_t has on θ. The use of the exponential allows for positive and negative values of v_t. Note that if $m = 0$ then θ is the same as in Equation 3.45. If $m \neq 0$, then v_t influences θ because this assumes that if the striking ship moves faster than the struck ship, longitudinal damage is oriented forward; otherwise it is oriented backwards. In Figure 3.19, the function $\theta(\phi, v_t; 1, 1)$ is plotted to give an impression of this model.

The idea is now to find optimal values $\hat{m}$ and $\hat{n}$ for each ship design, i.e. values that result in the highest fraction of correct outflow predictions q.

This maximization method is not easily solvable by general methods (the goal function invokes an algorithm to count the number of damaged compartments). Also, q is not continuous. Therefore a "brute force" approach was chosen to find a local maximum $\hat{m}, \hat{n}$ by taking a grid containing evenly spread values for m and n spread

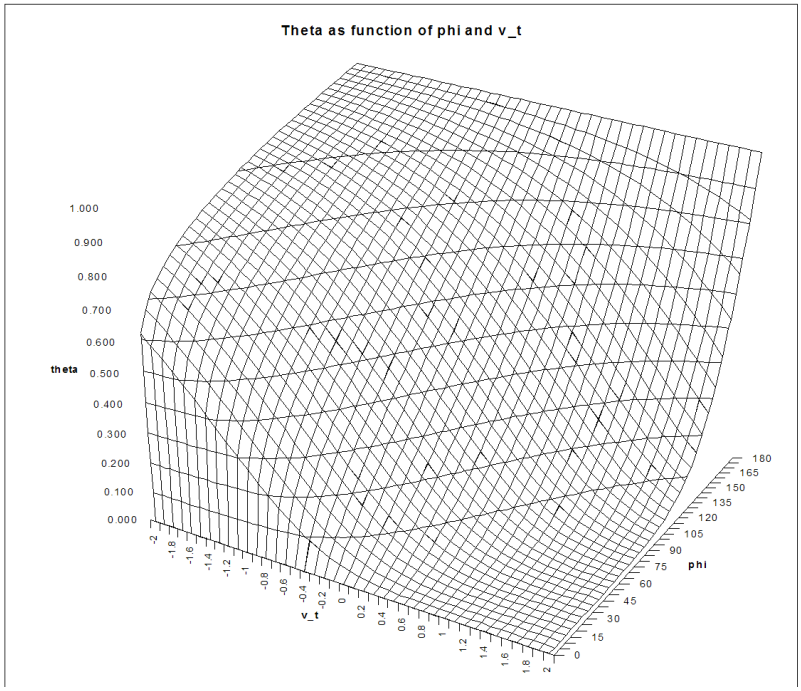


Figure 3.19: θ under different angles and relative tangential velocities

out over heuristically determined intervals and counting the corresponding value of q. After calulating these values, the values of m and n for which q was the highest were used as midpoints of a narrower grid. This was repeated down to 3 significant digits, beyond which it was deemed unlikely that any increase in significant digits would lead to a higher maximum of q. The maximum values of q are given in Table D.7.

3.7 Results

Damage Extent

Tables D.1 and D.3 show that the fits calculated for estimating the expected value of $\ln y_l$ and $\ln y_t$ have R^2-values between 68% and 75%. Interpreting these values as a qualitative metric to explain variation in the response variable, this result means that damage extent can be explained reasonably well by the input variables. The smaller vessels give slightly better R^2-values than the larger ones.

For $\ln y_l$, overall, x_1 and x_2 (representing kinetic energy) seem to account mostly for this explanation when looking at the coefficients (note that these variables are correlated). This fits with the idea that longitudinal damage extent is largely caused by the released amount of energy in the tangential direction. However, x_3 and x_4 also come into play depending on ship type. A few selected graphs are displayed in Figure 3.20 to show the difference between the effects of the variables on single hull and double hull damage (in the combined cases).

For $\ln y_t$, x_1 and x_2 are again dominant in causing transversal damage. x_3 (absolute collision location relative from the center) is also a major factor but only for the SH models. x_4 (bow angle) has little influence overall on the transversal damage extent. Again, this is a reasonably adequate argument for the notion that transversal damage is caused mostly by the energy release in the struck ship's perpendicular direction.

A switch in polarity and increase of magnitude of consecutive coefficients (for example $\beta_{3,1}, \beta_{3,2}, \ldots, \beta_{3,5}$ in Table D.1) can be observed.

Especially for the DHCOM model and, to a lesser extent for SHCOM, the added variable used to differentiate between the small ship dataset and the large ship's one seems not very significant for either $\ln y_l$ or $\ln y_t$.

Probability of Rupture

Table D.5 presents the coefficients that determine the probability of rupture $E(Z')$ given $\ln y_l$ an $\ln y_t$. Striking is the fact that the coefficient for transversal damage (β_t) is far bigger than β_l in the DH models, and the reverse is true for the SH models although to a far lesser extent; its coefficients are smaller (see also Figure 3.21).

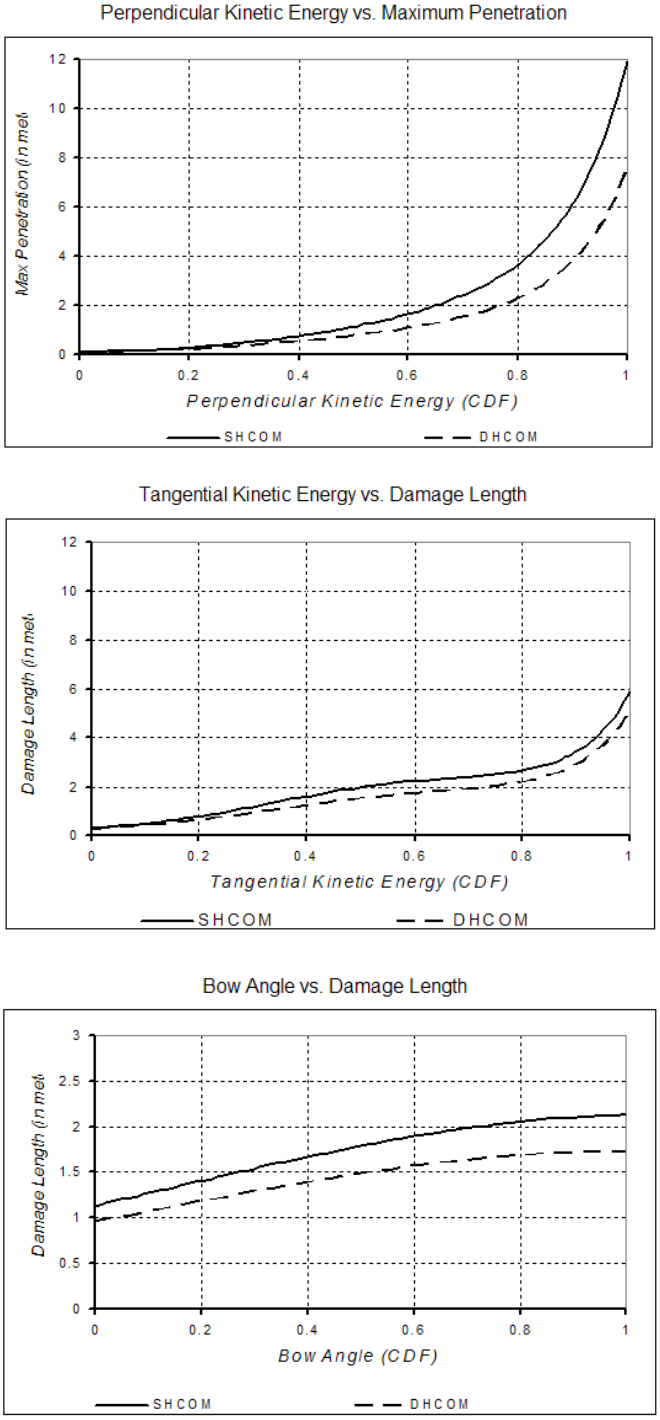


Figure 3.20: Effects of predictor variables on damage extent for a large ship using combined models

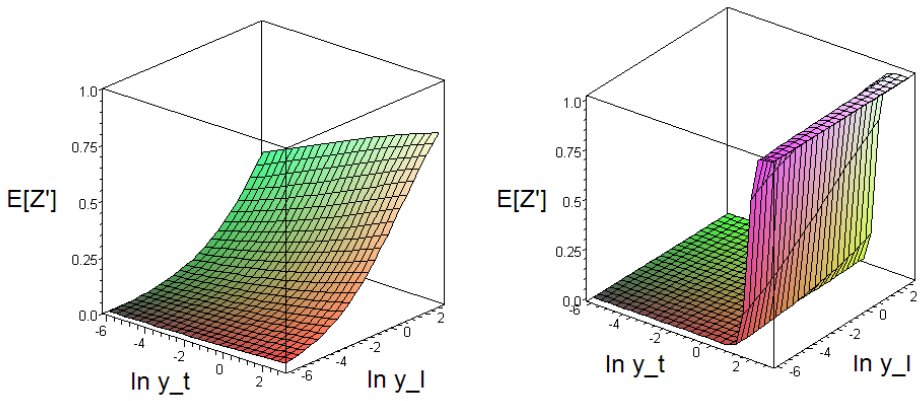


Figure 3.21: Expected probability of rupture as function of $\ln y_l$ and $\ln y_t$, SH150 vs. DH150

Also, in the latter, the intercept (β_0) is closer to 0.

These observations make clear that, in this model, probability of rupture in double hull ships is mainly due to transversal damage and that this probability does not start to become siginificantly large until a certain level of longitudinal damage is sustained; beyond this threshold, however, rupture becomes a near certainty. For single hull ships, probability of rupture increases more gradually and becomes quite large for modest damage extents.

The goodness-of-fit test values given in the table are mostly 0, meaning that — strictly speaking— their corresponding fits should be rejected based on the tests. As mentioned before, because of the large sample size, it is highly unlikely that any test would accept these fits. The QQ-plot (see Figure 3.14 of the data residual vs. random residual fits of the regression model show that the regression analysis matters in determining probability of rupture. The point biserial correlation coefficient, comparing the model with the data, gives significantly high values in all cases (between 0.5 and 0.8), thereby rejecting the null hypothesis. Moreover, testing with random data leads to a failed rejection of the null hypothesis.

Outflow Volume given Damage Extent and Rupture

By optimizing coefficients of a function that gives longitudinal damage location in relation to collision angle and relative tangential velocity, correct outflow volumes can be calculated with 95%—98% accuracy (see Table D.7). On average, this gives an outflow error between 88 and 417 m^3.

The calculation method of start- and endpoints for longitudinal damage might be improved upon by finding a more principled optimization algorithm. Also, for very low and very high values of ϕ the model might not be accurate.

Chapter 4

Grounding Model

Since the grounding model follows the same principles as the collision model, it is divided into three consecutive stages as well: based on the grounding input and output variables presented in Chapter 2, the model is supposed to

1. calculate the damage extent to the struck ship given the scenario input variables;
2. calculate the probability of rupture given damage extent;
3. calculate the oil spill volume given rupture.

This model is represented schematically in Figure 4.1.

The damaged area determines which compartments are ruptured. When the damage area overlaps a compartment it is assumed again that all its cargo is lost. Note that this methodology differs from the grounding simulation methodology [21] which this model is based on, because the latter invokes hydrostatic balance equations to determine final outflow volume. Another difference with the collision model is that no detailed analysis can be performed in determining damage locations, since the grounding simulation study does not provide bulkhead locations describing compartment locations.

In total, six different grounding models will be developed: four models based on individual tanker types and two combined models that are each based on simulation data from two tanker types.

4.1 Defining Predictor Variables

Kinetic Energy

Again, the grounding input variables can be transformed into predictor variables. Just as with collisions, kinetic energy is a desired variable to include in the grounding

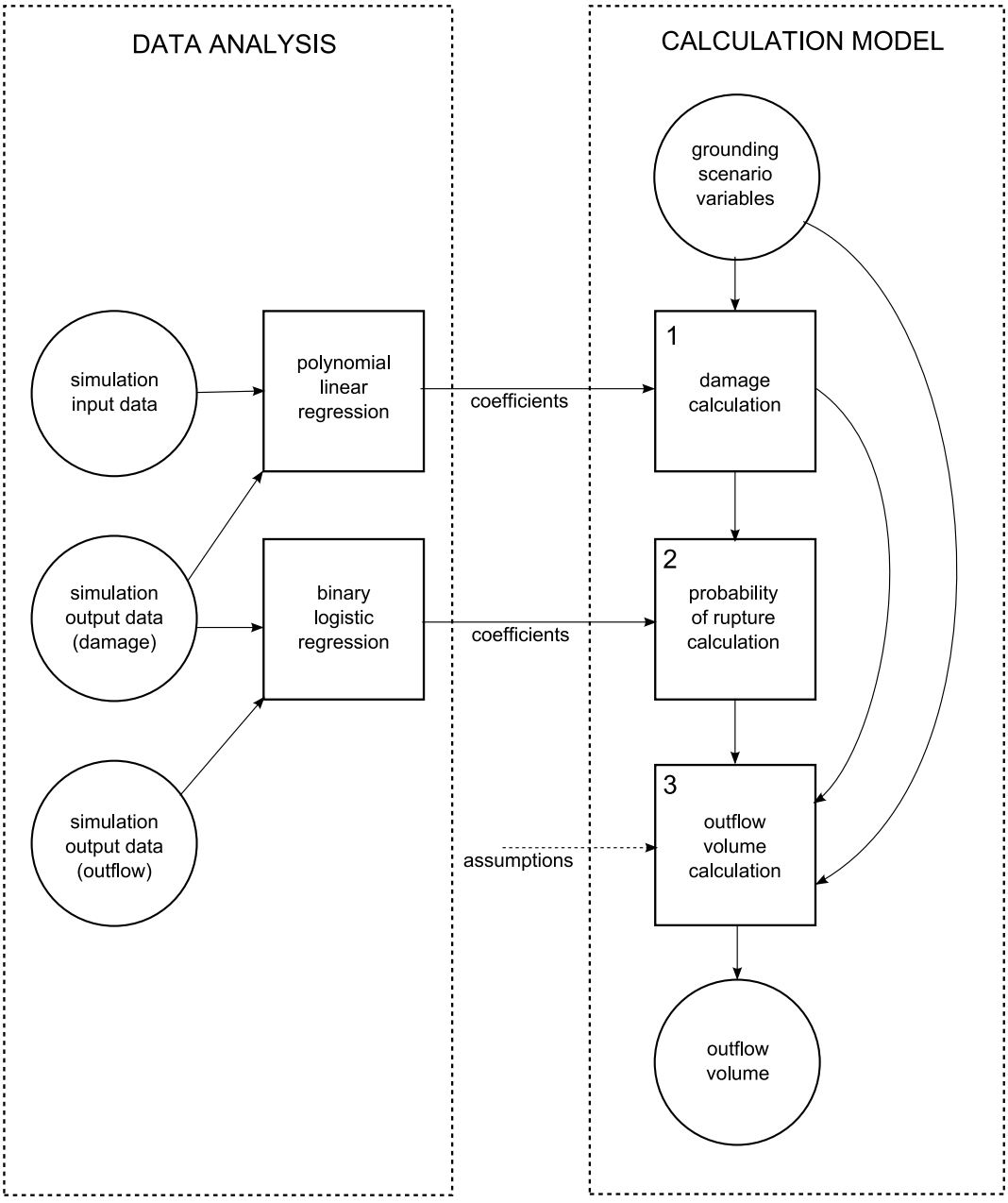


Figure 4.1: Grounding model schematic

model. Since groundings are head-on, and includes only one moving object, kinetic energy is defined as

$$e_k = \frac{1}{2}mv^2 \tag{4.1}$$

where m is the ship's mass and v its speed.

Obstruction Variables

Obstruction apex (o_a), obstruction depth (o_d), obstruction tip radius (o_r) and rock eccentricity (c) are straightforward variables and could have a strong influence on damage size. o_a, o_d and o_r describe the obstruction geometry and thus have a direct relationship with damage, whereas c describes how well a tanker can convert the tanker's longitudinal motion into other degrees of freedom. If $c = 0$, the rock tip is located at the centerline of the ship, making it difficult for the forward motion to change into a yawing or rolling motion. However, if $c = 1$ the rock tip is at either port or bow and leaves some leverage for the tanker to turn, thereby reducing forward speed and thus kinetic energy.

Other Variables

Since it is assumed that a breached compartment loses all its cargo, variables such as minimum outflow percentage ν and ballast tank capture b have no influence on the total amount of outflow. Furthermore, inert tank pressure p is unlikely to influence outflow since its maximum value (1000 mm water gauge) corresponds to approximately 0.1 atmosphere. This pressure refers to the inert gas that is added to the air in cargo compartments to prevent accidental combustion. Overpressure in the compartments might increase grounding damage and thus influence the probability of outflow or the size of the damage area, but since the tanks are assumed 98% full, the case can be made that the volume of air is too small to be of any influence; p should not make any difference to this argument. Finally, tidal variance τ is used in hydrostatic balance equations which is ignored in this study's grounding models. Hence, ν, b, p and τ will not be used as predictor variables in the model.

4.1.1 Transformation of Predictor Variables

As with the collision model, the predictor variables are transformed over their cumulative distribution functions. In some cases, these CDFs are known exactly: in other cases, a parametric distribution has to be fitted.

Kinetic Energy

Because the struck ship's mass m is a constant (four different masses are used for the four different ship types), kinetic energy is proportional to velocity squared:

$e_k = \frac{1}{2}mv^2$. The probability distribution of v is known from Table 2.9. From this, the probability distribution of the kinetic energy random variable E_k kan be derived:

$$\begin{aligned} P(E_k \leq x) &= P(\frac{1}{2}mV^2 \leq x) \\ &= P(V \leq \sqrt{\frac{2x}{m}}) \\ &= \begin{cases} 0, & \sqrt{\frac{2x}{m}} \leq 0 \\ \frac{1}{20}\sqrt{\frac{2x}{m}}, & 0 < \sqrt{\frac{2x}{m}} \leq 5 \\ \frac{1}{4} + \frac{3}{20}\ (\sqrt{\frac{2x}{m}} - 5), & 5 < \sqrt{\frac{2x}{m}} \leq 8 \\ \frac{7}{10} + \frac{2}{175}\ (\sqrt{\frac{2x}{m}} - 8), & 8 < \sqrt{\frac{2x}{m}} \leq 15 \\ \frac{39}{50} + \frac{1}{5}(\sqrt{\frac{2x}{m}} - 15), & 15 < \sqrt{\frac{2x}{m}} \leq 16 \\ \frac{49}{50} + \frac{1}{200}(\sqrt{\frac{2x}{m}} - 16), & 16 < \sqrt{\frac{2x}{m}} \leq 20 \\ 1, & \sqrt{\frac{2x}{m}} > 20 \end{cases} \end{aligned} \tag{4.2}$$

This distribution is used only for the SH40, SH150, DH40 and DH150 models. For the combined models (SHCOM and DHCOM), combining the kinetic energy dataset gives a different probability distribution:

$$P(E_k \leq x) = \frac{1}{2}[P(E_{k_1} \leq x) + P(E_{k_2} \leq x)] \tag{4.3}$$

Where E_{k_1} represents the kinetic energy of the smaller ship (SH40 or DH40) and E_{k_2} the one belonging to the larger ship (SH150 or DH150), both following a distribution as in Equation 4.2. The probabilities are weighted equally because the datasets are equally large.

Obstruction apex

A parametric distribution is fitted to the realizations of O_a because it is a truncated Normal distribution with unknown mean and variance. A generalized power distribution (see Appendix B.3) was chosen because it has a closed-form mathematical expression and is very flexible for a distribution that has bounded support. The coefficients of the fit are described in Table 4.1.

The fit is chosen by means of the least squares sum method, with n the same on each side to ensure the fitted probability distribution function is continuous. a and b were fixed, leaving α, m and n the coefficients to be determined. See Figure 4.2 for a QQ-plot that compares the fit with the cumulative CDF of O_a.

The parameters for the GP distribution of O_a are listed in Table 4.1.

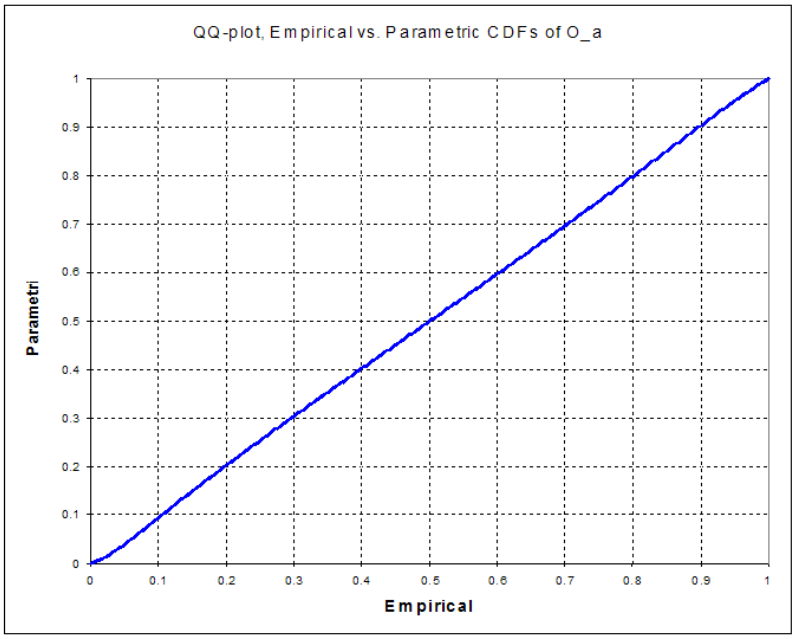


Figure 4.2: QQ-plot, Empirical vs. Parametric CDF, O_a

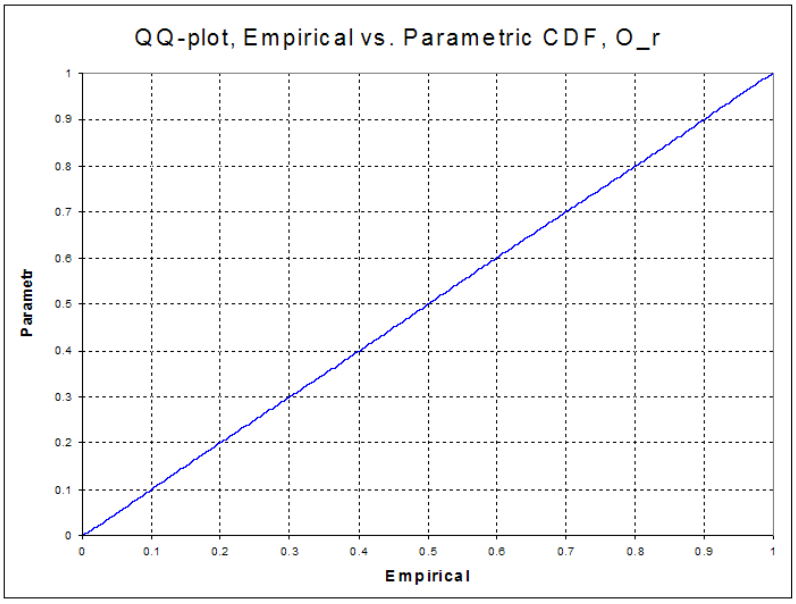


Figure 4.3: QQ-plot, Empirical vs. Parametric CDF, O_r

Coefficient	Value
a	15
m	19.557
b	50
α	1.186
n	4.018

Table 4.1: Coefficients for GP distribution of O_a

Coefficient	Value
a	0
m	5
b	10
α	1.507
n	2.379

Table 4.2: Coefficients for GP distribution of O_r

Obstruction Tip Radius

For O_r the generalized power distribution was selected for fitting since the original distribution is a truncated Normal as well. Since the probability distribution is symmetric around the mean 5, the fit is optimized by means of the least squares sum method with fixed mean and unknown variance. See Figure 4.3 for a QQ-plot that compares the fit with the cumulative CDF of O_r. The parameters for the GP distribution are listed in Table 4.2.

Obstruction Depth

By analyzing the grounding data, it is clear that obstruction depth O_d has CDF

$$P(O_d \leq x) \quad = \quad F_{O_d}(x) = \frac{1}{400}x^2, \quad x \in [0, 20] \tag{4.4}$$

(see Figure 4.5), which is validated by plotting this CDF against the empirical CDF obtained from the Data (Figure 4.4).

Rock Eccentricity

Rock eccentricity C is distributed uniformly on the interval $[0, 1]$.

So now each set of predictor variables $(e_k^i, o_d^i, o_a^i, o_r^i, c^i, d^i)$ for all $i \in \{1, \ldots, n\}$ can be seen as realizations of the aforementioned random variables E_k, O_d, O_a, O_r, C and D. Their corresponding CDFs are F_{E_k}, F_{O_d}, F_{O_a}, F_{O_r}, F_C and F_D which are given. The realizations are transformed through their corresponding CDF functions,

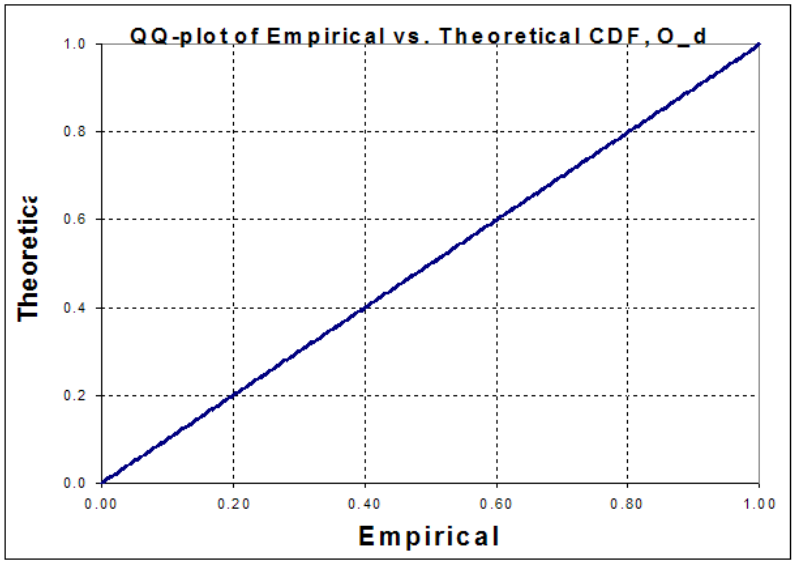


Figure 4.4: QQ-plot, Empirical vs. Theoretical CDF, O_d

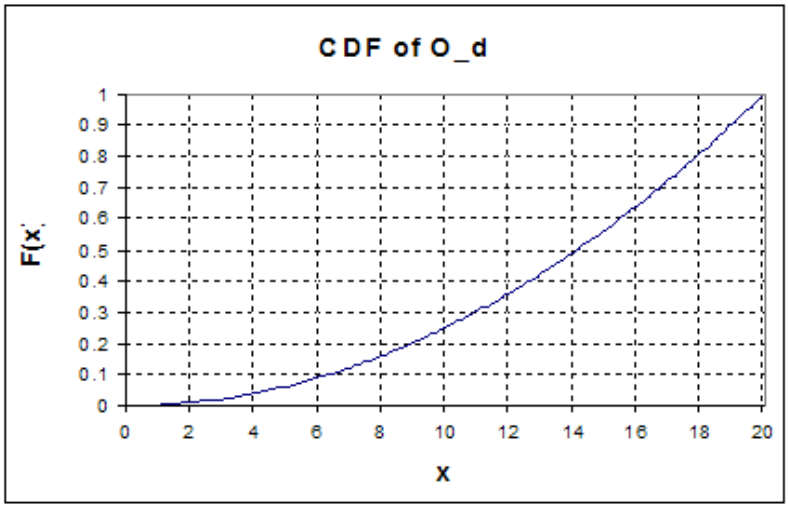


Figure 4.5: Obstruction depth distribution: fit vs. data

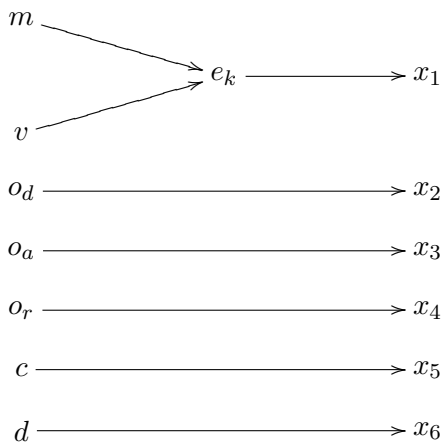


Figure 4.6: Transformation of input variables to predictor variables

resulting in the following transformed predictor variables:

$$\begin{aligned} x_{1,i} &= F_{E_k}(e_k^i) \\ x_{2,i} &= F_{O_d}(o_d^i) \\ x_{3,i} &= F_{O_a}(o_a^i) \\ x_{4,i} &= F_{O_r}(o_r^i) \\ x_{5,i} &= F_C(c^i) \\ x_{6,i} &= F_D(d^i) \\ & \forall i \in \{1 \ldots, n\} \end{aligned} \tag{4.5}$$

An overview of the transformation steps from input variables to predictor variables is given in Figure 4.6.

4.2 Damage Extent

The damage extent given input variables is determined by polynomial linear regression on the available datasets, just the same as in the collision model. Assuming that y_l^i and y_t^i are realizations of random variables Y_l and Y_t, polynomial linear regression determines the expected values of these variables conditioned on input variables $\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,6})$. Again, the logarithm of damage extent variables (y_l and y_t) is taken to ensure the correct application of linear regression. Since obstruction elevation y_v is directly related to obstruction depth o_d, there is no need to do linear regression on this variable.

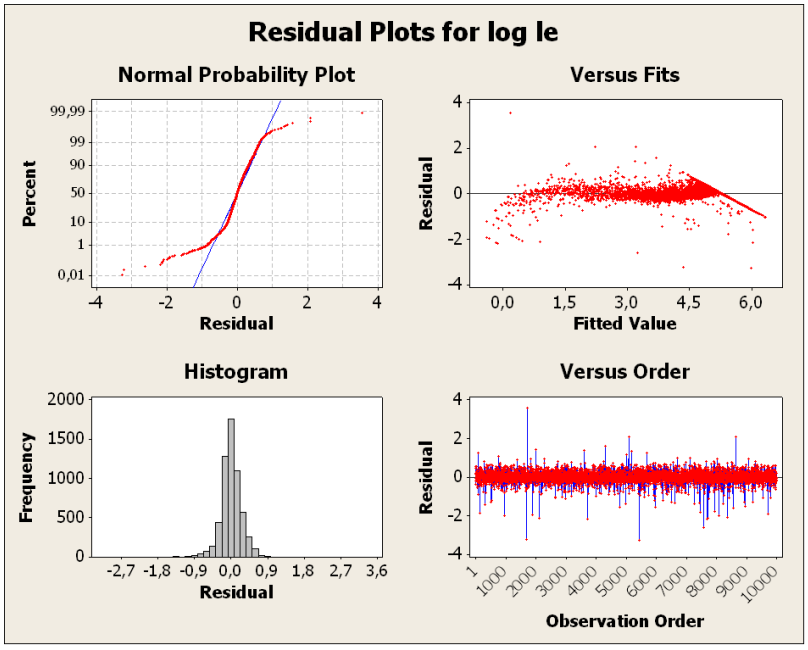


Figure 4.7: Residual plots for $\ln y_l$, SH150 case

$\ln Y_l$ is given as follows:

$$\begin{aligned} \ln Y_l &= h_l(\mathbf{x}|\beta^l) + R_l \\ &= \beta_0^l + \beta_{1,1}^l x_1 + \ldots + \beta_{1,6}^l x_6 \\ &= + \ldots \\ &= + \beta_{p,1}^l x_1^p + \ldots + \beta_{p,6}^l x_6^p \end{aligned} \tag{4.6}$$

($\ln Y_t$ is expressed analogously.) For this linear regression, $p = 5$ was chosen with the same procedure for selecting variables as in the collision model. The coefficients found by minimizing the sum of squares $\hat{\beta}^l$ and $\hat{\beta}^t$ can be found in Appendix E. Figures 4.7 and 4.8 show the residual plots.

4.2.1 Fitting Residual Distribution

This analysis is exactly the same as in the collision chapter. Residuals are treated as realizations of random variables R_l and R_t. The distributions of these variables are approached by the cumulative CDFs determined by the realizations, which in turn are fitted by a generalized trapezoidal distribution using a least squares method. The coefficients of this distribution are found in Tables E.3 and E.4 for $\ln y_l$ and $\ln y_t$, respectively. The QQ-plot of the empirical vs. the GT distributions of the residual R_t is plotted in 4.9.

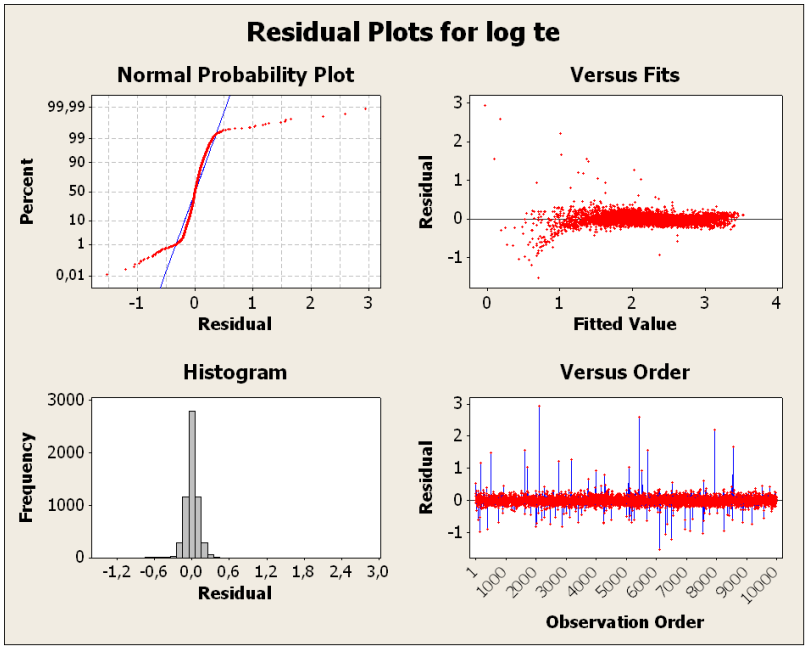


Figure 4.8: Residual plots for $\ln y_t$, SH150 case

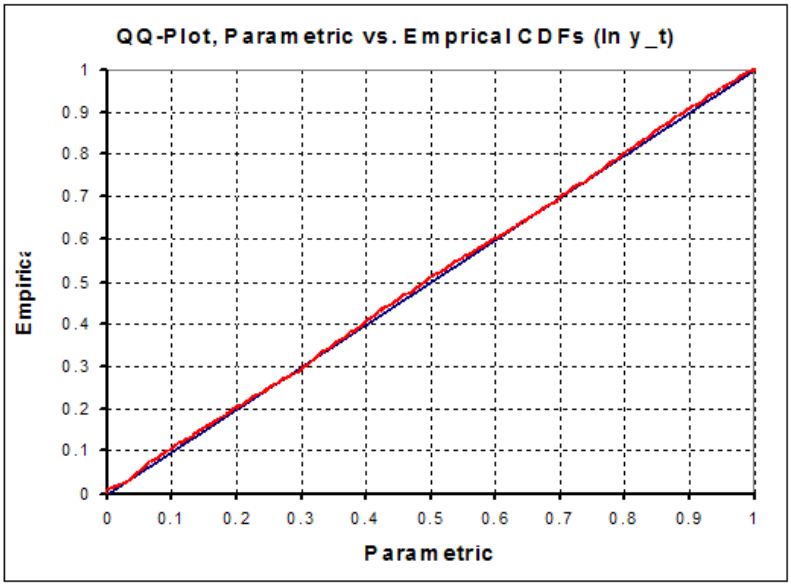


Figure 4.9: QQ-plot of empirical vs. parametric CDFs of r_t, SH150 case

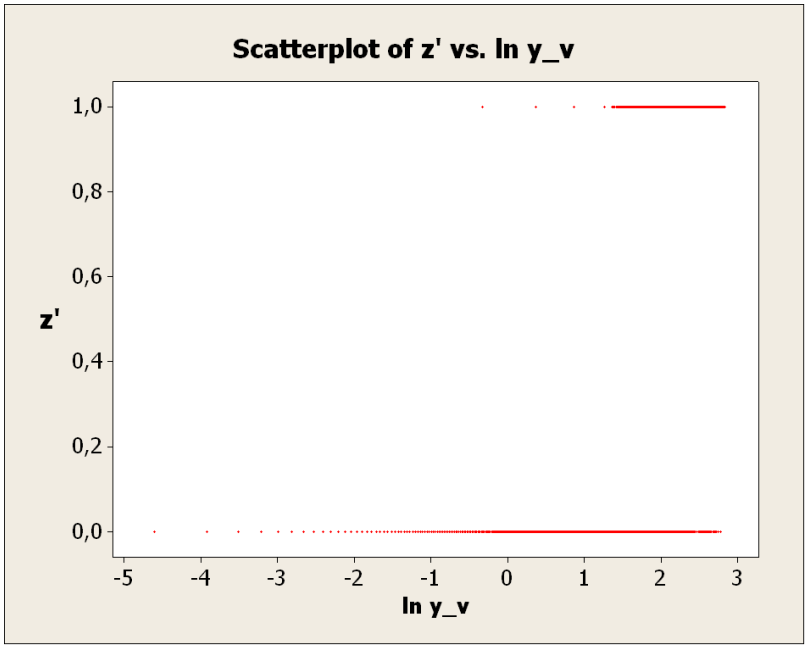


Figure 4.10: z' vs. $\ln y_v$, DH150 case

4.3 Probability of Rupture

The probability of rupture given grounding damage is determined by binary logistic regression of the occurrence of outflow, just the same as in the collision model. Again, the assumption goes that no outflow means no rupture. However, from the three variables that determine grounding damage - y_l, y_t and y_v - only y_v has positive values when $z' = 0$, i.e. when there is no outflow. This means that when transforming these variables by taking the natural logarithm, zero values of y_l, y_t cannot be used and leaves only those cases where outflow occurs. But binary logistic regression requires that all possible values of z' are present in the data, making regression on z' by y_l and y_t impossible. Therefore, binary logistic regression is carried out with only one predictor variable, $\ln y_v$, resulting in the following model:

$$E(Z'|\ln y_v) = \frac{\exp(\beta_0 + \beta_v y_v)}{1 + \exp(\beta_0 + \beta_v y_v)} \tag{4.7}$$

In Figure 4.10 the occurrence of outflow z' is plotted against $\ln y_v$. Results are given in Table E.5.

The significance of this model against a purely random model is measured again by looking at the departure of the residuals of this model with the current dataset against the residuals of this model with a Bernouilli generated dataset (which generates 1's with probability p and 0's with probability $1 - p$, p being the frequency of

outflow occurrence).

For formal significance testing, the point-biserial correlation coefficients are tested in the same way as in collisions, for the real data and randomly generated data. Results of these tests are in Table E.6.

4.4 Outflow Volume

Both start and end locations (y_{l1} and y_{l2}) appear in the original dataset for longitudinal damage extent. However, because $y_{l1} = 0$ in an overwhelming amount - above 98.5% and 94% in SH and DH cases, respectively - it is assumed in the modelling of oil outflow that $y_{l1} = 0$.

There is no data available on innermost and outermost edges of transversal damage extent y_{t1} and y_{t2} but it is assumed that these factors are determined as

$$y_{t1} = (\frac{1}{2} + c) \cdot s_b - \frac{1}{2} y_t \tag{4.8}$$

$$y_{t2} = (\frac{1}{2} + c) \cdot s_b + \frac{1}{2} y_t \tag{4.9}$$

Unlike in the collision model, there are no bulkhead locations given for the ship types in groundings so there is no way to validate these assumptions directly. When the grounding bulkhead locations are set to be the same as with collisions (as in Table C.9) there is a poor match with the real data w.r.t. which compartments are damaged.

Furthermore, setting the damaged area equal to a rectangular volume with dimensions y_l, y_t and y_v at the determined coordinates, all compartments coinciding with this volume will be assumed ruptured and all oil from these compartments is assumed lost.

4.5 Results

4.5.1 Damage Extent

By looking at the coefficients in Table E.1, $\ln y_l$ is by far the most dependent on kinetic energy (x_1) in the polynomial linear regression model. Obstruction depth (x_2) and tip radius (x_4) to a much lesser extent with some minor significance to rock eccentricity in the DH models. The R^2-values are high: around 93% for all SH models, and above 87% for the DH40 and DH150 models. Only the combined DH model performs less according to this metric, but is still reasonably good at 79%.

The regression results for $\ln y_t$ (see Table E.2) are even better in this view: all six models have R^2-values ranging between 90% and 94%. x_4 is the most influential variable, followed by x_2. A simple explanation for this is the fact that

- a bigger tip radius makes a bigger hole;
- because its shape is broader at the base, the rock will create more transversal damage if its tip is at lower depth;
- a higher apex angle means a broader cone base and thus creates a bigger hole.

Finally it should be noted that in the combined models for $\ln y_l$, the added variable (x_6) doesn't play a big role and shows a negative relationship. In $\ln y_t$, this variable is more substantial.

In Figure 4.12 some graphs are plotted between predictor variables and response variables y_l (longitudinal damage extent) and y_t (transversal damage extent) where, for each graph, all other variables are fixed at 0.5. It appears that damage extents are smaller for the SH40 case than for the SH150 case. It can be seen that

Tip radius has a negative influence on damage length; this is because the force exerted on the ship is greater when tip radius is larger. Note that longitudinal damage goes down when the kinetic energy CDF increases in the last few percentiles. This is not plausible and could be attributable to artifacting of the polynomial function.

4.5.2 Probability of Rupture

From Table E.5, it seems that the double hull ships are more resistant to rupture (the lower values for β_0 mean that the probability of rupture is near zero even for a relatively high $\ln y_v$). Probability of rupture goes up fast after a certain threshold has been reached (higher values for β_v). A plot of all logistic fits are given in 4.11. At least one of the goodness-of-fit tests for each binary logistic model give a p-value of 1 (see Table E.5), with the DH40 model scoring a p-value over 0.05 in all three tests. The point biserial correlation coefficient gives significantly high values in all cases (over 0.58), thereby rejecting the null hypothesis. Moreover, testing with random data leads to a failed rejection of the null hypothesis.

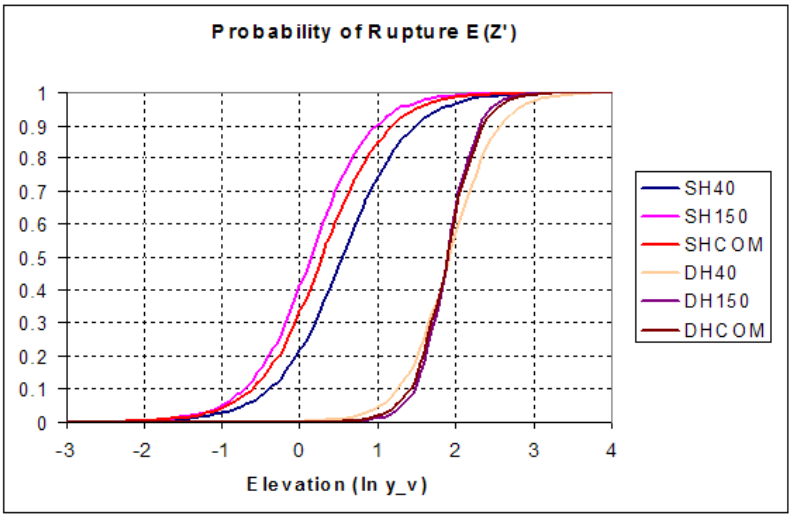


Figure 4.11: $E(Z')$ as function of $\ln y_v$

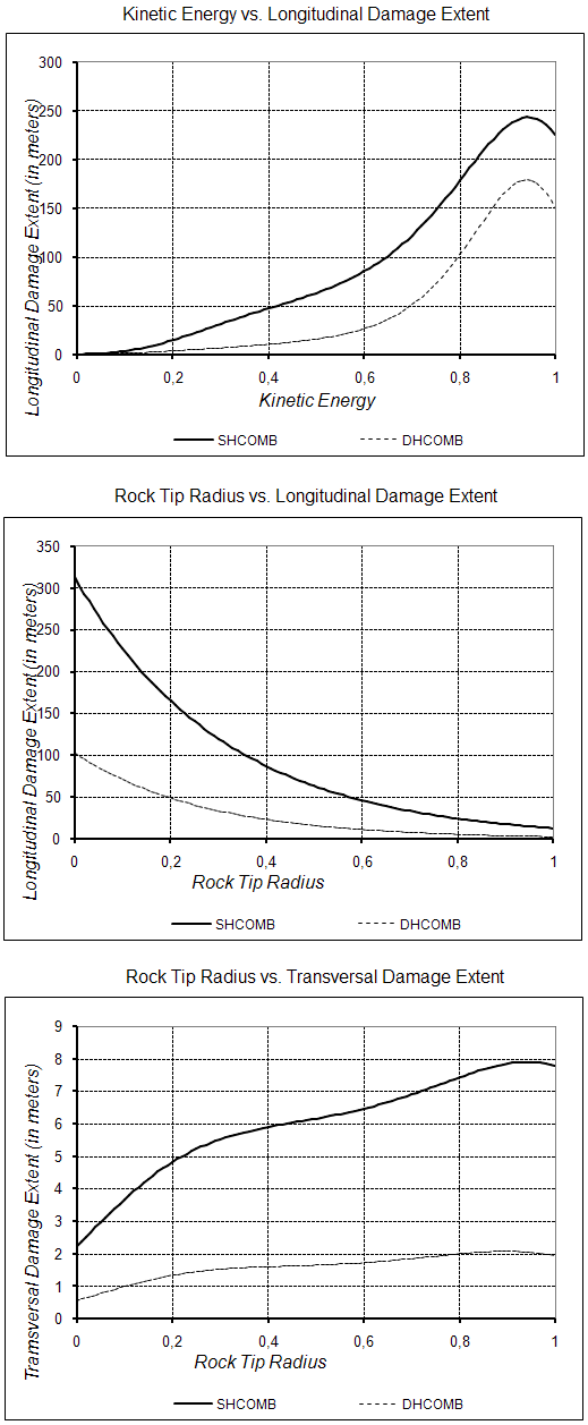


Figure 4.12: Effects of predictor variables on damage extent for a large ship using combined models

Chapter 5

Calculation Examples

Now that the outflow models have been discussed, a collision example and a grounding example are given to suggest how these models should be applied.

5.1 Struck Ship Configuration

To keep things simple, a single hull and double hull design are used in the examples in this chapter, each sharing the same input variables. The struck ship parameters that need to be configured are:

- Displacement
- Dimensions (length, breadth, depth)
- Bulkhead locations (longitudinal and transversal)
- Compartment volumes

For both collisions and groundings, a struck ship is chosen with 175,000 metric tonnes displacement.

- The dimensions, bulkhead locations and compartment volumes for the collision struck ships are determined according to the configurations of the SH150 and DH150 tankers as specified in the collision section of Chapter 2.
- For groundings, the dimensions and compartment volumes are the same as in the grounding section of Chapter 2; the bulkhead locations will be the same as the collision struck ships.

The outflow models of choice will be the combined single hull (SHCOM) and combined double hull (DHCOM). Because the struck ship dimensions are the same as the large ships specified in Chapter 2, the dimensional variable d is set to 1 in all models.

Variable	Value	Unit
v_1	12	knots
m_1	50	$\times$ 1000 metric tons
v_2	5	knots
ϕ	45	degrees
l	0.7	-
η	25	degrees

Table 5.1: Collision example variables

5.2 Collision Example

5.2.1 Input Variables

In a collision scenario, aside from the struck tanker's parameters, six input variables are needed to calculate expected damage size and expected probability of rupture. In Table 5.1, five arbitrary input variables are given. These fall within the bounds given by the probability distributions in Chapter 3.

Note that to obtain the collision models in Chapter 3, the variable t was involved in determining bow angle η. In this section η is arbitrarily chosen directly instead. This factually introduces a new striking ship type and shows the flexibility of the collision model.

5.2.2 Transformations

Now, calculate $e_{k,p}$ and $e_{k,t}$ as in Equation 3.13:

$$\begin{aligned} e_{k,p} &= \frac{1}{2}(m_1+m_2)(v_1\sin\phi)^2 \\ &= \frac{1}{2}(50+175)(12\cdot\frac{1}{2}\sqrt{2})^2 \\ &= 8100 \qquad (5.1) \\ e_{k,t} &= \frac{1}{2}(m_1+m_2)(v_2+v_1\cos\phi)^2] \\ &= \frac{1}{2}(50+175)(5+12\cdot\frac{1}{2}\sqrt{2})^2 \\ &= 20458 \qquad (5.2) \end{aligned}$$

Calculate l':

$$l' = |l-\frac{1}{2}| = |0.7-\frac{1}{2}| = 0.2 \qquad (5.3)$$

Transforming these through CDFs from Chapter 3 gives the set of input variables $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$:

$$x_1 = F_{E_{k,p}}(e_{k,p}) = 1 - \exp(-\frac{e_{k,p}}{\beta})^{\alpha} \quad (5.4)$$

$$x_2 = F_{E_{k,t}}(e_{k,t}) = 1 - \exp(-\frac{e_{k,t}}{\beta})^{\alpha} \quad (5.5)$$

$$x_3 = F_{L'}(l') = Beta(l' + \frac{1}{2}|1.25, 1.45) - Beta(-l' + \frac{1}{2}|1.25, 1.45) \quad (5.6)$$

$$x_4 = F_H(\eta) = 1 \quad (5.7)$$

$$x_5 = F_D d = 1 \quad (5.8)$$

Because the transformation parameters for $e_{k,p}$ and $e_{k,t}$ are almost the same for single hull and double hull models, the transformations have (almost) the same values:

	Single Hull	Double Hull
x_1	0.962	0.962
x_2	0.987	0.987
x_3	0.465	0.465
x_4	1	1
x_5	1	1

5.2.3 Step One: Damage Extent

Given the input variables $\mathbf{x}$, one can now get the expected logarithm of damage length ($\ln y_l$), the expected logarithm of maximum penetration ($\ln y_t$) and their associated random error terms r_l and r_t:

$$\ln y_l = h_l(\mathbf{x}|\hat{\beta}^l) + r_l \quad (5.9)$$

$$\ln y_t = h_t(\mathbf{x}|\hat{\beta}^t) + r_t, \quad (5.10)$$

or, taking the exponential,

$$y_l = \exp(h_l(\mathbf{x}|\hat{\beta}^l) + r_l) \quad (5.11)$$

$$y_t = \exp(h_t(\mathbf{x}|\hat{\beta}^t) + r_t) \quad (5.12)$$

where h_l and h_t are functions given in Equation 3.27; r_l and r_t are the corresponding error terms and generated from random variables R_l and R_t. For simplicity, the random terms are ignored in this calculation. The coefficients $\hat{\beta}^l$ and $\hat{\beta}^t$ can be found in Tables D.1 and D.3. Calculating results in the following values:
Or:

	Single Hull	Double Hull
$\ln y_l$	3.376	3.084
$\ln y_t$	2.289	1.915

	Single Hull	Double Hull
y_l	29.249	21.854
y_t	9.863	6.789

5.2.4 Step Two: Probability of Rupture

Next, $\ln y_l and \ln y_t$ are put into the probability function $\pi(\ln y_l, \ln y_t|\hat{\beta})$, where coefficients $\hat{\beta}$ can be found in Table D.5.

$$\pi(\ln y_l, \ln y_t|\hat{\beta}) = \frac{\exp(\beta_0 + \beta_l \ln y_l + \beta_t \ln y_t)}{1 + \exp(\beta_0 + \beta_l \ln y_l + \beta_t \ln y_t)} \tag{5.13}$$

This is the probability of rupture. The results are:

	Single Hull	Double Hull
π	0.822	0.976

5.2.5 Step Three: Outflow Volume

With the probability of rupture $\pi = P(Z' = 1)$, the actual occurrence of rupture can be determined by "flipping a coin" (i.e. sampling a Bernouilli distributed random variable with parameter π). Suppose that the outcome is zero: then no rupture occurs and thus no outflow. In the other case, the longitudinal coordinates of the damaged area have to be determined.
Take m, n from D.7:

	Single Hull	Double Hull
m	0.112	0.091
m	5.91	5.62

Then $v_t = v_2 + v_1 \cos(\phi)$ and θ can be calculated (see Equation 3.47):

$$\theta(\phi, v_t; m, n) = (\frac{1}{2}(\frac{\phi}{90})^n)^{\exp(mv_t)} \tag{5.14}$$

This results in:

	Single Hull	Double Hull
θ	≈ 0	≈ 0

Determine damaged compartments from y_{l1}, y_{l2} and y_t using ship length s:

$$y_{l1} = (1-\theta)y_l + (1-l)s \quad (5.15)$$
$$y_{l2} = -\theta y_l + (1-l)s \quad (5.16)$$

Which leads to

	Single Hull	Double Hull
y_{l1}	79.89	78.30
y_{l2}	109.14	100.15

Now, for the single hull tanker, the bulkheads have to be looked up from the second column of Table C.9 that bound these locations: these are bulkheads 2 and 4 (which are 53.9 resp. 137.1 meters away from the FP). From this it can be seen that the longitudinal damage runs across the 3rd and 4th compartment as counted from the FP. (The first compartment is in between the FP and the first bulkhead.) Since $y_t = 9.863$ meters, the transversal damage extends only into the outermost compartments. Thus, the 3^{rd} and 4^{rd} outer compartments have been ruptured. Looking at Table C.2, the 3^{rd} contains $15311m^3$ of oil; the other zero. Hence the total outflow volume z for the single hull tanker equals 15311 m^3.

In the double hull case, the bulkhead locations are looked up from the 4^{th} column of Table C.9. This shows that the longitudinal damage is contained by bulkheads 3 and 4. Since $y_t = 6.789$, transversal damage reaches 2 compartments inward from the outer hull. Thus, one outer and one inner compartment in the the 4^{th} row from the front are ruptured. Since the outer one is a ballast tank (compartment volume is 0) only the inner compartment spills oil, which amounts to 14651 m^3.

5.3 Grounding Example

5.3.1 Input Variables

In Table 5.2, some possible values of grounding input variables are given.

The only predictor variable that has to be calculated is e_k:

$$e_k = \frac{1}{2}mv^2 = \frac{1}{2} \cdot 175 \cdot 8.1^2 = 5741 \quad (5.17)$$

Variable	Value	Unit
v	8.1	knots
o_d	15	meters
o_a	42	degrees
o_r	6.7	meters
c	0.61	-

Table 5.2: Grounding example variables

5.3.2 Transformations

$\mathbf{x}$ is determined through transformating the input variables through their CDF values:

$$x_1 = F_{E_k}(e_k) = \frac{7}{10} + \frac{2}{175}(\sqrt{\frac{2e_k}{m}} - 8) = 0.7001 \quad (5.18)$$

$$x_2 = F_{O_d}(o_d) = \frac{1}{400}o_d^2 = \frac{1}{400}225 = \frac{9}{16} \quad (5.19)$$

$$x_3 = F_{O_a}(o_a) = 0.843 \quad (5.20)$$

$$x_4 = F_{O_r}(o_r) = 0.737 \quad (5.21)$$

$$x_5 = F_C(c) = c = 0.61 \quad (5.22)$$

5.3.3 Step One: Damage Extent

y_l and y_t are determined using the polynomial linear regression model, whose coefficients β^l and β^t can be found in Tables E.1 and E.2, respectively.

$$\ln y_l = h_l(\mathbf{x}|\hat{\beta}^l) + r_l \quad (5.23)$$

$$\ln y_t = h_t(\mathbf{x}|\hat{\beta}^t) + r_t \quad (5.24)$$

r_l and r_t are the corresponding error terms and generated from random variables R_l and R_t. Again, for simplicity the random terms are ignored. The coefficients $\hat{\beta}^l$ and $\hat{\beta}^t$ can be found in Tables D.1 and D.3, resulting in:

	Single Hull	Double Hull
$\ln y_l$	4.602	3.740
$\ln y_t$	2.462	1.755

Or:

	Single Hull	Double Hull
y_l	99.63	42.10
y_t	11.73	5.781

Using the ship depth $s_d = 16.76$ for both ships, one can calculate $y_v = \max(0, s_d - o_d) = 1.76$.

5.3.4 Step Two: Probability of Rupture

Next, put $\ln y_v$ into the binary logistic model $\pi(\ln y_v|\hat{\beta})$, where coefficients $\hat{\beta}$ can be found in Table E.5.

$$\pi(\ln y_v|\hat{\beta}) = \frac{\exp(\beta_0 + \beta_v \ln y_v)}{1 + \exp(\beta_0 + \beta_v \ln y_v)} \tag{5.25}$$

This results in the following probabilities

	Single Hull	Double Hull
π	0.665	0.002

5.3.5 Step Three: Outflow Volume

Given rupture, it is assumed that damage starts at the front of the ship. Also, rock eccentricity c is assumed to be in the middle of transversal damage extent. So,

$$y_{l1} = 0 \tag{5.26}$$
$$y_{l2} = y_l \tag{5.27}$$
$$y_{t1} = \frac{1}{2}(1+c)\cdot s_b - \frac{1}{2}y_t \tag{5.28}$$
$$y_{t2} = \frac{1}{2}(1+c)\cdot s_b + \frac{1}{2}y_t \tag{5.29}$$

This results in

	Single Hull	Double Hull
y_{l1}	0	0
y_{l2}	99.63	42.10
y_{t1}	34.38	37.36
y_{t2}	46.12	43.14

Where $s_b = 50.0$ is the ship's breadth in meters. Since y_{t2} is larger than the ship's breadth, it is reset at 50.

Using these coordinates, ruptured compartments can be determined using the bulkhead locations in Table C.9.

In the single hull case, longitudinally, the first four compartments as seen from the FP are damaged; transversally, the center and side compartments. The corresponding cargo volumes are presented in Table C.6, and thus the total outflow volume can be

calculated:

$$\begin{aligned} z &= 3,951,288 + 2,911,920 + 4,793,184 + 0 \\ &+ 4,792,392 + 3,402,960 + 4,192,584 + 0 \\ &= 24,044,328 \text{ gallons,} \end{aligned}$$

corresponding to $91,018\ m^3$.

In the double hull case, longitudinally, the first two compartments as seen from the FP are damaged; transversally, the center compartments. The corresponding cargo volumes are presented in Table C.8, and thus the total outflow volume can be calculated:

$$z = 2,593,272 + 3,254,064 = 5,847,336 \text{ gallons,}$$

corresponding to $22,135\ m^3$.

5.4 Conclusions

Comparing the example results, it should be noted that the double hull ships incur less damage extent given the same input variables: particularly in the grounding examples, the damaged area is more than four times smaller in the double hull case. This results an outflow volume four times smaller than in the single hull case. The difference in the collision examples is much less striking.

Chapter 6

Conclusions and Recommendations

In this report, twelve accidental outflow models have been presented: six collision models and six grounding models. These models determine the amount of oil that flows from an oil tanker in case it is struck by another ship or runs aground on a rocky pinnacle. Based on simulation data, these models have the ability to calculate fairly accurately the extent of collision or grounding damage, the probability of rupture and oil spill volume and the damage location given a set of accident variables. Uncertainties in outcomes of damage extent have been accurately modeled by fitting residuals to a parametric distribution.

Each of these models can be quickly and easily implemented in large scale system simulations of tanker movements because they involve formulas using only elementary functions and include an overseeable amount of parameters and coefficients. In short, they combine the power of physical simulations with the simplicity of explicit functions.

Moreover, these models improve significantly upon the previous IMO model since

- they are based on a large dataset obtained by physically meaningful simulations, rather than a model with simplistic assumptions based on a small historic dataset;
- they allow for size-dependent damage extent and probability of rupture assessments, whereas the old model gave damage and probability independently of ship size;
- damage extent parameters are dependent on scenario input variables as opposed to independently distributed;
- damage extent parameters take into account the physical characteristics of the ship designs and accident scenarios, such as speed, mass, collision angle etc.

6.1 Collision Model Results

- Kinetic energy is mostly responsible for damage extent;
- The regression model for damage extent fits reasonably well with data, giving R^2-values of 68%-75%;
- Single hull ships incur more damage overall than double hull designs;
- The regression model for probability of rupture shows higher rupture resistance for double hull tankers;
- Probability of rupture is strongly influenced by maximum penetration for double hull designs, whereas damage length is mostly responsible for rupture in single hull ones;
- Probability of rupture shows significant correlation with outflow occurrence in data;
- Damage location and outflow calculation model gives 95%-98% accuracy of outflow volume given rupture and damage extent.

6.2 Grounding Model Results

- Kinetic energy is mostly responsible for longitudinal damage;
- A large obstruction tip radius reduces longitudinal damage;
- Variables that describe rock geometry have the overhand in predicting transversal damage;
- The regression models for damage extent fits very well to the data, with 10 out of 12 giving R^2-values over 90%;
- The rupture probability model shows higher rupture resistance for double hull tankers, given obstruction elevation;
- Probability of rupture shows significant correlation with outflow occurrence in data.

6.3 General Remarks

A number of aspects should be considered in light of this research.

- The actual shape of the damaged area in collisions and groundings cannot be determined from the data: the models are only based on simplified measurements. They assume the damaged area to be a rectangular block, which holds the maximum damage volume possible.

- Given rupture, all compartments coinciding with the damaged area are assumed ruptured, whereas it might be possible that rupture takes place in a fraction of that area.
- All oil in a ruptured compartment is assumed lost, which is —in the case of grounding— a worst case simplification.
- The event that no outflow occurs is assumed to imply that there is no hull rupture, since no information is provided that would allow one to conclude otherwise.

6.4 Recommendations for Further Research

Below are some issues that may be considered topics of further research.

Using the large data set and great number of predictor variables available, reasonable to good fit performance was achieved for both polynomial linear and binary logistic regressions. As with any regression technique and especially due to the large number of predictor variables, other combinations of independent variables (taking advantage of e.g. interaction terms) could potentially lead to even better performance in terms of fit. A preliminary investigation of the use of interaction terms only showed a marginal improvement, while not reducing the number of variables.

The outflow models are based on statistical analysis, where output data is compared to input data. These factors mostly concern the 'outside' aspects of the struck tanker: no consideration is given to the influence of the ship's inner conditions, such as number of bulkheads etc. on damage size or probability of rupture—they only matter in determining the outflow volume. Improvements could be made in this, but it should be noted that the model in its current form is already both simple and effective; therefore any inclusion of mentioned internal aspects should only marginally increase the model's complexity. It then has to be tested how effective this inclusion is.

Bibliography

[1] Bedford, T., and Cooke, R. *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge University Press, Cambridge, U.K., 2001.

[2] Bookrags. Mallows' cp summary. http://www.bookrags.com/wiki/Mallows

[3] Brown, A. *Alternative Tanker Designs, Collision Analysis*. NRC Marine Board Committee on Evaluating Double-Hull Tanker Design Alternatives, 2001.

[4] Brown, A., and Amrozowicz, M. Tanker environmental risk - putting the pieces together. *Joint SNAME/SNAJ Conference on Designs and Methodologies for Collision and Grounding Protection of Ships* (1996).

[5] Friis-Hansen, P., and Simonsen, B. Gracat: software for grounding and collision risk analysis. *Marine Structures 15* (2002), 383–401.

[6] Herbert Engineering Corp. Oil outflow analysis for a series of double hull tankers, 1998. Report No. 9749-1 Rev. A.

[7] Hosmer, D., and Lemeshow, S. *Applied Logistic Regression. Second Edition*. Wiley-Interscience, 2000.

[8] Huijer, K. Trends in oil spills from tanker ships, 1995-2004. London.

[9] International Maritime Organization. Interim guidelines for the approval of alternative methods of design and construction of oil tankers under regulation 13f(5) of annex i of marpol73/78, 1995.

[10] Jongbloed, G., and Groeneboom, P. *Voortgezette Statistiek*. Technische Universiteit Delft, 1999.

[11] Kotz, S., and van Dorp, J. *Beyond Beta. Other Continuous Families of Distributions with Bounded Support and Applications*. World Scientific Publishing, Singapore, 2004.

[12] Marine Board Commission on Engineering and Technical Systems. Tanker spills: Prevention by design, 1991.

[13] Meko, D. Applied time series analysis, lecture notes 11: Multiple linear regression, 2007. http://www.ltrr.arizona.edu/ dmeko/geos585a.html#cLesson11.

[14] Merrick, J., et al. The prince william sound risk assessment. *Interfaces 32*, 6 (2002), 25–40.

[15] Minorsky, V. An analysis of ship collisions with reference to protection of nuclear power plants. *Journal of Ship Research 3*, 1 (1959).

[16] Rawson, C., Crake, K., and Brown, A. Assessing the environmental performance of tankers in accidental grounding and collision. *SNAME Transactions 106* (1998), 41–58.

[17] Rodrigue, J.-P., Comtois, C., and Slack, B. *The Geography of Transport Systems*. Routledge, New York, 2006.

[18] Schmeiser, B. Advanced input modelling for simulation experimentation. *Proceedings of the 1999 Winter Simulation Conference* (1999).

[19] Simonsen, B., and Hansen, P. Theoretical and statistical analysis of ship grounding accidents. *Journal of Offshore Mechanics and Arctic Engineering 122* (2000), 200–207.

[20] The National Academies. *Special Report 259. Environmental Performance of Tanker Designs in Collision and Grounding*. The National Academies Press, 2001.

[21] Tikka, K. *Alternative Tanker Designs, Grounding Analysis*. NRC Marine Board Committee on Evaluating Double-Hull Tanker Design Alternatives, 2001.

[22] UNCTAD. Review of maritime transport 2007. New York / Geneva, 2007.

[23] van der Laan, M. *Environmental Tanker Design*. Delft University of Technology, 1997.

[24] van Dorp, J., and Kotz, S. Generalized trapezoidal distributions. *Metrika* (2003).

Appendix A

Regression

A.1 Binary Logistic Regression

This section discusses binary logistic regression as described by Hosmer and Lemeshow Chapters 1 and 2 [7]. Given is a binary random variable Y. In a regression analysis, the expected value of Y (the response variable) is related to a function of a set of predictor variables $\mathbf{x} = (x_1, \ldots, x_m)$, which in turn is based on a sample set $(\mathbf{x}_i, y_i), \quad i \in \{1, \ldots, n\}$.

In a binary logistic regression, this function is the logistic function π and represents the expected value of Y conditioned on $\mathbf{x}$. Notation:

$$E(Y|\mathbf{x}) \quad = \quad \pi(\mathbf{x}|\beta) \tag{A.1}$$

Where π is defined as

$$\pi(\mathbf{x}|\beta) \quad = \quad \frac{e^{g(\mathbf{x}|\beta)}}{1 + e^{g(\mathbf{x}|\beta)}}. \tag{A.2}$$

With

$$g(\mathbf{x}|\beta) \quad = \quad \beta_0 + \beta_1 x_1 + \ldots + \beta_m x_m \tag{A.3}$$

$\beta = (\beta_0, \ldots, \beta_m)$ is a set of coefficients that defines the shape of g and thus π. Binary linear regression determines an optimal set of coefficients $\hat{\beta}$, i.e. coefficients that result in the 'most accurate' fit of π against the variables.

A.1.1 Fitting the Logistic Regression Model

Given n realizations of independent, identically distributed sets of variables

$$(\mathbf{X}_i, Y_i), \quad i \in \{1, \ldots, n\} \tag{A.4}$$

Now, the coefficients β are fitted from the dataset of scenarios $\mathbf{x}_i$ by means of the maximum likelihood estimation. Consider the set $(\mathbf{x}_i, y_i)$ of observed data, where y_i

is the dependent variable corresponding to independent variables $\mathbf{x}_i$.

The maximum likelihood method yields values for the unknown coefficients β which maximize the probability of obtaining the observed set of data. This is done by constructing a likelihood function l, which expresses the probability of observed data as a function of β.

Since, by definition,

$$E(Y|\mathbf{x}) = 0 \cdot P(Y = 0|\mathbf{x}) + 1 \cdot P(Y = 1|\mathbf{x}) \quad (A.5)$$
$$= P(Y = 1|\mathbf{x}) \quad (A.6)$$

for any $\mathbf{x}$, it follows that $P(Y = 1|\mathbf{x}) = \pi(\mathbf{x}|\beta)$ and $P(Y = 0|\mathbf{x}) = 1 - P(Y = 1|\mathbf{x}) = 1 - \pi(\mathbf{x}|\beta)$. Then, one may express the contribution for the pair $(\mathbf{x}_i, y_i)$ to the likelihood function as

$$\pi(\mathbf{x}_i|\beta)^{y_i} \left[1 - \pi(\mathbf{x}_i|\beta)\right]^{1-y_i}. \quad (A.7)$$

As the observations are assumed independent, the likelihood function is obtained as the product of these contributions:

$$l(\beta) = \prod_{i=1}^{n} \pi(\mathbf{x}_i|\beta)^{y_i} [1 - \pi(\mathbf{x}_i|\beta)]^{1-y_i}. \quad (A.8)$$

Now, β is estimated as the value which maximizes the right hand side of A.8, also referred to as $\hat{\beta}$. The loglikelihood is defined as follows:

$$L(\beta) = \ln[l(\beta)] \quad (A.9)$$
$$= \sum_{i=1}^{n} y_i \ln[\pi(\mathbf{x}_i)] + (1 - y_i) \ln[1 - \pi(\mathbf{x}_i)] \quad (A.10)$$

Because l and L have a maximum at the same value(s) of β, It becomes relatively straightforward to find β by maximizing L (as opposed to l), which in turn is done by partially differentiating $L(\beta)$ to $\beta_0, \ldots, \beta_m$ and equating the resulting expressions to 0:

$$\frac{\partial L}{\partial \beta_0} = 0 \quad (A.11)$$
$$\frac{\partial L}{\partial \beta_j} = 0, \quad j \in \{1, \ldots, m\} \quad (A.12)$$

These are the likelihood equations; solving them for $\beta_0, \ldots, \beta_m$ will result in the *maximum likelihood estimate* $\hat{\beta}$. However, these equations are nonlinear and the workings of the required solving method go beyond the scope of this report. The statistical software package Minitab 15 is capable of performing this method and was used in this report.

The maximum likelihood estimate of $\pi(\mathbf{x}|\beta)$, which is $\pi(\mathbf{x}|\hat{\beta})$, is denoted as $\hat{\pi}(\mathbf{x})$ and represents a "best" estimate of the probability that outflow occurs, given a scenario $\mathbf{x} = (x_1, \ldots, x_m)$. Thus, $\hat{\pi}(\mathbf{x})$ is the *probability* of the event $Y = 1$ happening based on binary logistic regression.

A.2 Linear Regression

The method of linear regression as described here was based on Chapter 3 of [10]. Given a set of scenarios $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^p$ and outcomes $y_1, y_2, \ldots, y_n \in \mathbb{R}$ realizations of random variables $Y_1, Y_2, \ldots, Y_n$. Then a linear regression model expresses the relationship between Y_i and $\mathbf{x}_i$ as follows:

$$\begin{aligned} Y_i &= h(\mathbf{x}_i|\beta) + R_i && \text{(A.13)} \\ &= \beta_0 + \beta_1 x_{i,1} + \ldots \beta_p x_{i,p}, \quad \forall i \in \{1, \ldots, n\} && \text{(A.14)} \end{aligned}$$

Where $R_1, R_2, \ldots, R_n$ are assumed to be uncorrelated random variables with mean zero and finite variance. Y_i is the response variable and x_i is the vector containing predictor variables. h is the function that needs to be determined by changing the coefficients in vector $\beta = (\beta_0, \beta_1, \ldots, \beta_p) \in \mathbb{R}^{p+1}$.

Based on a sample $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\} \in \mathbb{R}^p \times \mathbb{R}$, an estimate of β can be found. A systematic method to do this is the least squares method, whereby a least squares estimate $\hat{\beta}$ is found by minimizing the sum of squares of the residuals over β:

$$\begin{aligned} S(\hat{\beta}) &= \min_{\beta} S(\beta) && \text{(A.15)} \\ &= \min_{\beta} \sum_{i=1}^{n} (y_i - h(\mathbf{x}_i|\beta))^2 && \text{(A.16)} \end{aligned}$$

Minima of S are found by determining the partial derivatives of S to β, equating these derivatives to 0 and solving these equations for β, resulting in the linear regression estimator $\hat{\beta}$. If S is convex, then $\hat{\beta}$ is a global minimum.

$$\begin{aligned} \frac{\partial S}{\partial \beta_j} &= 2 \sum_{i=1}^{n} \epsilon_i \frac{\partial \epsilon_i}{\partial \beta_j} && \text{(A.17)} \\ &= -2 \sum_{i=1}^{n} (y_i - f(\mathbf{x}_i, \beta)) \frac{\partial f(\mathbf{x}_i, \beta)}{\partial \beta_j}, \quad , \forall j \in \{0, \ldots, p\} && \text{(A.18)} \end{aligned}$$

Appendix B

Probability Distributions

B.1 Empirical Distribution Function

For a random variable X the cumulative distribution function F is defined as $F(x) = \mathbb{P}(\{X \leq x\})$. The empirical cumulative distribution function F_n of a sample of n i.i.d. random variables $X_1, \ldots, X_n \sim X$ is defined as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{(-\infty,x]}(X_i) \tag{B.1}$$

,
Where $1_{(-\infty,x]}(y) = 1$ if $y \leq x$, and $1_{(-\infty,x]}(y) = 0$ otherwise. The empirical CDF has the property that $F_n(x) \to F(x)$ almost surely for a fixed x by the strong law of large numbers.

B.2 Typical Distributions

Beta Distribution

The Beta probability distribution function is given as

$$f(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) + \Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{B.2}$$

Where Γ is the Gamma function and α, β are the function's parameters.

Normal Distribution

The Normal probability distribution function is given as

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) \tag{B.3}$$

Where μ and σ^2 are the mean and variance and determine location and scale of the distribution, respectively.

Weibull Distribtion

The Weibull distribution has the following cumulative distribution function:

$$F(x|\alpha,\beta) \quad = \quad 1 - e^{(-\frac{x}{\beta})^{\alpha}} \tag{B.4}$$

Where $\alpha \geq 0$ is the shape parameter and $\beta \geq 0$ is the scale parameter. Note that for $\alpha = 1$, the Weibull distribution is equivalent to the Exponential distribution with parameter β.

B.3 Generalized Power Distribution

For $0 < m < 1$ and $0 \leq x \leq 1$, the Generalized Power Distribution [11] is defined as follows:

$$f(x|\alpha,m,n) \quad = \quad \begin{cases} p(\frac{x}{m}|\alpha,n), & 0 \leq x \leq m \\ p(\frac{1-x}{1-m}|\alpha,n), & m < x < 1 \end{cases} \tag{B.5}$$

where

$$p(x|\alpha,n) = \alpha + n(1-\alpha)x^{n-1}$$

and, for $0 \leq x \leq 1$,

$$\begin{cases} 0 \leq \alpha \leq \frac{n}{n-1}, & n > 1 \\ 0 \leq \alpha \leq 1, & 0 < n \leq 1. \end{cases}$$

If x is on an interval $[a,b]$, then it should be scaled by transforming it to a variable y on the interval $[0,1]$:

$$y \quad = \quad \frac{x-a}{b-a} \tag{B.6}$$

Thus,

$$f(y|\alpha,m,n) = f(\frac{x-a}{b-a}|\alpha,m,n)$$

.

B.4 Generalized Trapezoidal Distribution

Suppose X is a random variable on the bounded support $[a,b]$. If X follows the Generalized Trapezoidal distribution [24], its probability distribution function is defined as follows:

$$f(x|a,b,c,d,n_1,n_3,\alpha) = \begin{cases} 0, & x < a \\ \frac{2\alpha n_1 n_3}{2\alpha(b-a)n_3+(\alpha+1)(c-b)n_1 n_3+2(d-c)n_1}(\frac{x-a}{b-a})^{n_1-1}, & a \leq x < b \\ \frac{2n_1 n_3}{2\alpha(b-a)n_3+(\alpha+1)(c-b)n_1 n_3+2(d-c)n_1}((\alpha-1)\frac{c-x}{c-b}+1), & b \leq x < c \\ \frac{2n_1 n_3}{2\alpha(b-a)n_3+(\alpha+1)(c-b)n_1 n_3+2(d-c)n_1}(\frac{d-x}{d-c})^{n_3-1}, & c \leq x < d \\ 0, & x \geq d \end{cases} \tag{B.7}$$

Provided that $n_1 > 0, n_3 > 0, \alpha > 0$ and $a < b < c < d$.

Appendix C

Tanker Data

Compartment location	Port	Center	Starboard
Bow	1865.4	3641.1	1865.4
	2640.8	0.0	2640.8
	2673.5	3646.1	2673.5
	0.0	3644.4	0.0
	2668.0	3643.6	2668.0
Stern	2529.1	3642.2	2529.1

Table C.1: Tanker compartment volumes (m^3), SH40, collisions

Compartment location	Port	Center	Starboard
Bow	13102.9	17779.5	13102.9
	0.0	21566.6	0.0
	15311.4	21563.4	15311.4
	0.0	18864.3	0.0
	8364.9	19658.5	8364.9
Stern	3820.4	19658.5	3820.4

Table C.2: Tanker compartment volumes (m^3), SH150, collisions

Compartment location	Port	Port Center	Starboard Center	Starboard
	0.0	2269.7	2267.7	0.0
	0.0	2825.3	2825.3	0.0
	0.0	2845.9	2845.9	0.0
	0.0	2845.9	2844.9	0.0
	0.0	0.0	0.0	0.0
	0.0	2276.5	2276.5	0.0
	0.0	2845.9	2844.9	0.0
	0.0	2845.9	2845.9	0.0
Stern	0.0	2669.5	2671.5	0.0

Table C.3: Tanker compartment volumes (m^3), DH40, collisions

Compartment location	Port	Port Center	Starboard Center	Starboard
Bow	0.0	11694.3	11694.3	0.0
	0.0	14674.2	14674.2	0.0
	0.0	14650.4	14650.4	0.0
	0.0	14651.2	14651.2	0.0
	0.0	14650.8	14650.8	0.0
	0.0	13861.9	13861.9	0.0
Stern	0.0	5514.7	5514.7	0.0

Table C.4: Tanker compartment volumes (m^3), DH150, collisions

Compartment location	Port	Center	Starboard
Bow	413,688	792,528	413,688
	585,552	0	585,552
	592,944	808,632	592,944
	0	808,104	0
	591,624	808,104	591,624
Stern	560,736	783,816	560,736

Table C.5: Tanker compartment volumes (gallons), SH40, groundings

Compartment location			
	Port	Center	Starboard
Bow	2,911,920	3,951,288	2,911,920
	0	4,793,184	0
	3,402,960	4,792,392	3,402,960
	0	4,192,584	0
	1,859,088	4,368,936	1,859,088
Stern	849,024	0	849,024

Table C.6: Tanker compartment volumes (gallons), SH150, groundings

Compartment location		Port	Starboard	
	Port	Center	Center	Starboard
Bow	0	505,560	505,560	0
	0	626,472	626,472	0
	0	629,376	629,376	0
	0	630,168	630,168	0
	0	503,712	503,712	0
	0	630,168	630,168	0
	0	628,320	628,320	0
Stern	0	590,832	590,832	0

Table C.7: Tanker compartment volumes (gallons), DH40, groundings

Compartment location		Port	Starboard	
	Port	Center	Center	Starboard
Bow	0	2,593,272	2,593,272	0
	0	3,254,064	3,254,064	0
	0	3,248,784	3,248,784	0
	0	3,249,048	3,249,048	0
	0	3,249,048	3,249,048	0
	0	3,074,016	3,074,016	0
Stern	0	1,083,192	1,083,192	0

Table C.8: Tanker compartment volumes (gallons), DH150, groundings

Transversal bulkheads (Location from FP (m))			
SH40	SH150	DH40	DH150
14.63	12.3	16.46	12.3
37.948	53.9	33.99	43.5
61.265	95.5	51.51	74.7
84.582	137.1	69.04	105.9
107.899	173.5	86.56	137.1
131.216	199.5	90.07	168.3
154.534	214.3	104.09	199.5
		121.62	214.3
		139.14	
		156.67	

Longitudinal bulkheads (Location from port bow (m))			
SH40	SH150	DH40	DH150
8.23	14.8	2.438	3.34
19.202	35.2	14.63	25
		26.822	46.66

Table C.9: Bulkhead locations

Appendix D

Collision Model Results

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
number of data points	7467	7473	14940	7454	7466	14920
R^2-value	70.9%	68.1%	68.9%	71.5%	69.9%	70.6%
Mallows C_p-value	19.0	19.8	13.1	14.2	24.0	16.0
Coefficients						
β_0	-2.914	-2.661	-2.982	-2.931	-2.786	-2.632
$\beta_{1,1}$	3.078	-1.215	2.246	2.128	2.047	-0.117
$\beta_{2,1}$	5.550	5.303	5.231	6.180	4.692	4.670
$\beta_{3,1}$	0.031	-2.493	-3.369	0.708	-3.224	-1.973
$\beta_{4,1}$	0.546	1.613	1.188	0.655	1.429	1.155
$\beta_{5,1}$	-	-	0.223	-	-	0.052
$\beta_{1,2}$	-	10.181	0.687	0.598	-	5.792
$\beta_{2,2}$	-	-	-	-5.563	-	-
$\beta_{3,2}$	-	20.261	25.010	-	24.187	16.819
$\beta_{4,2}$	-	-0.931	-0.560	-	-0.784	-0.566
$\beta_{5,2}$	-	-	-	-	-	-
$\beta_{1,3}$	-	-8.145	-	-	-	-
$\beta_{2,3}$	-11.982	-6.405	-6.750	-	-5.410	-5.756
$\beta_{3,3}$	-	-68.750	-75.742	-13.309	-69.908	-53.668
$\beta_{4,3}$	-	-	-	-0.158	-	-
$\beta_{5,3}$	-	-	-	-	-	-
$\beta_{1,4}$	-2.924	-	-	-	-	-10.900
$\beta_{2,4}$	9.403	-	-	-	-	-
$\beta_{3,4}$	-	94.811	96.400	27.442	85.081	69.372
$\beta_{4,4}$	-	-	-	-	-	-
$\beta_{5,4}$	-	-	-	-	-	-
$\beta_{1,5}$	2.823	2.008	-	-	0.542	7.798
$\beta_{2,5}$	-	4.134	4.529	2.291	3.724	4.031
$\beta_{3,5}$	-0.480	-44.783	-43.224	-15.354	-36.872	-31.216
$\beta_{4,5}$	-	-	-	-	-	-
$\beta_{5,5}$	-	-	-	-	-	-

Table D.1: Polynomial linear regression coefficients for $\ln y_l$, collisions

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
α	1	1	1	1	1	1
A	-17.266	-16.802	-17.261	-15.478	-15.402	-15.851
B	-0.153	-0.362	-0.278	-0.191	-0.312	-0.254
C	0.217	0.426	0.352	0.254	0.425	0.356
D	5.304	5.585	5.585	5.250	5.585	5.585
N1	35.833	26.036	30.196	31.101	26.547	29.222
N3	10.299	8.089	9.221	9.995	10.133	10.471

Table D.2: Parameters of GT distributions, R_l, collisions

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
number of data points	7470	7478	14948	7455	7467	14922
R^2-value	73.8%	70.4%	71.4%	74.6%	72.6%	73.5%
Mallows Cp-value	14.0	18.2	15.0	12.8	20.1	20.6
Coefficients						
β_0	-3.730	-3.507	-3.977	-3.655	-3.629	-3.681
$\beta_{1,1}$	8.661	4.492	6.767	6.527	6.793	6.650
$\beta_{2,1}$	5.439	3.479	4.828	4.585	2.790	3.985
$\beta_{3,1}$	-4.126	1.357	-3.234	-0.321	0.308	0.427
$\beta_{4,1}$	0.010	0.378	1.267	0.030	0.289	0.051
$\beta_{5,1}$	-	-	0.227	-	-	0.044
$\beta_{1,2}$	-6.939	-	-3.339	-3.250	-4.298	-3.758
$\beta_{2,2}$	-7.083	-	-5.251	-5.971	-	-4.329
$\beta_{3,2}$	28.940	-6.123	23.896	5.613	-	-
$\beta_{4,2}$	-	-	-1.313	-	-	-
$\beta_{5,2}$	-	-	-	-	-	-
$\beta_{1,3}$	-	-	-	-	-	-
$\beta_{2,3}$	-	-5.602	-	-	-4.492	-
$\beta_{3,3}$	-80.644	-	-72.669	-25.920	-6.807	-9.296
$\beta_{4,3}$	-	-	-	-	-	-
$\beta_{5,3}$	-	-	-	-	-	-
$\beta_{1,4}$	3.268	-	-	-	-	-
$\beta_{2,4}$	3.229	-	-	2.848	-	-
$\beta_{3,4}$	96.373	19.916	93.704	40.495	16.125	20.693
$\beta_{4,4}$	-	-0.585	-	-0.345	-0.531	-
$\beta_{5,4}$	-	-	-	-	-	-
$\beta_{1,5}$	-	0.243	1.534	1.462	2.212	1.828
$\beta_{2,5}$	-	3.841	2.074	-	3.285	1.872
$\beta_{3,5}$	-41.499	-15.976	-42.700	-20.501	-10.209	-12.407
$\beta_{4,5}$	-0.263	-	-	-	-	-0.354
$\beta_{5,5}$	-	-	-	-	-	-

Table D.3: Polynomial linear regression coefficients for $\ln y_t$, collisions

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
α	1	1	1	1	1	1
a	-15.282	-17.654	-17.346	-16.113	-14.270	-16.355
b	0.056	-0.207	-0.099	0.030	-0.256	-0.110
c	0.192	0.355	0.287	0.182	0.372	0.304
d	3.312	3.912	3.912	3.376	3.912	3.912
n_1	29.369	25.266	27.822	30.668	25.556	29.228
n_3	7.299	5.577	6.580	7.161	7.128	7.761

Table D.4: Parameters of GT distributions, R_t, collisions

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
No. Cases	7440	7430	14811	7423	7436	14788
Coefficients						
β_0	-0.229	-0.864	-0.511	-7.026	-10.823	-7.142
β_t	0.162	0.164	0.158	5.943	7.330	5.443
β_l	0.536	0.514	0.498	0.257	0.283	0.143
MLR	-4534	-4367	-9065	-1114	-796	-2190
Pearson Test	0	0	0	0	0	0
Deviance Test	0	0	0	1	1	1
Hosmer-Lemeshow Test	0	0	0	0	0	0

Table D.5: Binary logistic regression coefficients, collisions

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
No. Cases	7440	7430	14811	7423	7436	14788
r_{bp} (data)	0.40	0.43	0.41	0.85	0.86	0.82
p-value (data)	0	0	0	0	0	0
r_{bp} (random)	-0.01	-0.02	-0.00	0.00	0.01	0.01
p-value (random)	0.50	0.17	0.78	0.80	0.36	0.14

Table D.6: Binary logistic regression point-biserial correlation tests, collisions

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
No. of cases	4045	3183	7228	1404	1026	2430
% correct predictions	97.11%	97.86%	97.40%	94.87%	96.78%	95.60%
m	0.112	0.098	0.112	0.061	0.091	0.091
n	5.90	6.20	5.91	4.59	5.60	5.62
avg. absolute error (m^3)	88	289	189	134	417	255
conditional average absolute error (m^3)	3045	13513	7248	2609	12950	5800

Table D.7: Damage location coefficients

Appendix E

Grounding Model Results

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
number of data points	1806	5899	7705	609	2673	3282
R^2-value	93.3%	93.3%	93.2%	87.0%	90.8%	79.4%
Mallows Cp-value	21.8	23.4	30.8	18.2	15	21.7
Coefficients						
β_0	-2.866	-1.327	-1.403	-3.925	-2.403	-0.592
$\beta_{1,1}$	41.818	41.940	30.664	50.806	41.949	16.217
$\beta_{2,1}$	3.398	1.141	4.703	6.133	3.761	4.394
$\beta_{3,1}$	0.102	-0.044	0.085	-0.326	-0.150	-0.136
$\beta_{4,1}$	-4.750	-2.277	-3.194	-5.365	-3.027	-3.708
$\beta_{5,1}$	-0.406	-0.226	0.085	1.298	-0.610	1.175
$\beta_{6,1}$	-	-	-0.146	-	-	-0.320
$\beta_{1,2}$	-104.639	-116.403	-74.472	-139.873	-106.135	-25.308
$\beta_{2,2}$	-	-	-12.152	-20.431	-4.750	-8.377
$\beta_{3,2}$	-	-	-	-	-	-
$\beta_{4,2}$	11.369	4.509	7.174	8.726	4.519	6.078
$\beta_{5,2}$	-	-1.842	-2.851	-1.951	6.895	-
$\beta_{6,2}$	-	-	-	-	-	-
$\beta_{1,3}$	96.878	140.345	85.822	168.169	98.551	-
$\beta_{2,3}$	-5.096	0.286	14.138	20.867	-	4.459
$\beta_{3,3}$	-	-	-	-	-	-
$\beta_{4,3}$	-12.822	-4.769	-8.234	-5.621	-2.568	-3.568
$\beta_{5,3}$	1.362	4.033	5.109	-	-11.504	-3.524
$\beta_{6,3}$	-	-	-	-	-	-
$\beta_{1,4}$	-	-59.455	-35.523	-70.533	-	47.300
$\beta_{2,4}$	-	-	-	-	-	-
$\beta_{3,4}$	0.206	0.104	-	-	-	-
$\beta_{4,4}$	5.047	1.853	3.453	-	-	-
$\beta_{5,4}$	-	-	-	-	-	-
$\beta_{6,4}$	-	-	-	-	-	-
$\beta_{1,5}$	-26.548	-	-	-	-26.883	-32.250
$\beta_{2,5}$	-	-	-5.977	-	-	-
$\beta_{3,5}$	-	-	-	-	-	-
$\beta_{4,5}$	-	-	-	-	-	-
$\beta_{5,5}$	-2.361	-	5.330	-	5.330	2.444
$\beta_{6,5}$	-	-	-	-	-	-

Table E.1: Polynomial linear regression coefficients for $\ln y_l$, groundings

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
number of data points	2720	5904	8624	644	2724	3368
R^2-value	90.0%	93.6%	91.6%	92.5%	93.7%	92.7%
Mallows Cp-value	21.7	23.1	21.2	18	25.2	33.7
Coefficients						
β_0	1.473	2.049	1.112	1.229	1.769	1.095
$\beta_{1,1}$	0.065	0.111	0.096	0.170	0.095	0.142
$\beta_{2,1}$	-5.088	-4.060	-4.251	-0.775	-3.258	-4.002
$\beta_{3,1}$	0.720	1.239	0.740	0.008	0.767	0.782
$\beta_{4,1}$	7.520	5.857	6.397	9.308	6.709	7.575
$\beta_{5,1}$	-0.148	-0.186	-0.002	-0.825	0.103	-0.488
$\beta_{6,1}$	-	-	1.004	-	-	0.692
$\beta_{1,2}$	-	-0.093	-	-	-	-
$\beta_{2,2}$	25.437	12.507	8.287	-14.912	6.663	7.987
$\beta_{3,2}$	-2.210	-2.624	-1.025	-	-1.153	-1.315
$\beta_{4,2}$	-19.182	-14.714	-16.229	-28.430	-17.836	-20.824
$\beta_{5,2}$	0.175	-	-	2.922	-	1.652
$\beta_{6,2}$	-	-	-	-	-	-
$\beta_{1,3}$	-	-	-	-	-	-
$\beta_{2,3}$	-104.542	-17.161	-	-	-	-
$\beta_{3,3}$	2.893	2.865	0.593	-	0.668	0.801
$\beta_{4,3}$	22.161	16.975	18.870	37.140	21.247	25.400
$\beta_{5,3}$	-	1.303	-	-	-	-
$\beta_{6,3}$	-	-	-	-	-	-
$\beta_{1,4}$	-	-	-	-	-	-
$\beta_{2,4}$	187.918	-	-23.858	72.857	-37.722	-37.565
$\beta_{3,4}$	-1.291	-1.078	-	-	-	-
$\beta_{4,4}$	-9.019	-6.974	-7.793	-16.758	-8.939	-10.917
$\beta_{5,4}$	-	-1.734	-	-11.801	-2.828	-6.571
$\beta_{6,4}$	-	-	-	-	-	-
$\beta_{1,5}$	-	-	-0.058	-	-	-
$\beta_{2,5}$	-106.772	7.761	19.393	-62.279	37.674	36.762
$\beta_{3,5}$	-	-	-	-	-	-
$\beta_{4,5}$	-	-	-	-	-	-
$\beta_{5,5}$	-0.753	-	-0.591	9.846	2.639	5.347
$\beta_{6,5}$	-	-	-	-	-	-

Table E.2: Polynomial linear regression coefficients for $\ln y_t$, groundings

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
α	1	1	1	1	1	1
a	-10.994	-10.103	-13.431	-9.172	-13.119	-13.650
b	0.006	-0.090	-0.106	-0.037	0.033	0.105
c	0.006	0.049	0.105	0.157	0.033	0.296
d	5.304	5.585	5.585	5.250	5.585	5.585
n_1	64.487	61.720	71.475	33.447	67.252	34.301
n_3	30.138	27.469	27.964	20.649	30.055	23.793

Table E.3: Parameters of GT distributions, R_l, groundings

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
α	1	1	1	1	1	1
a	-6.125	-5.983	-6.776	-5.504	-7.561	-7.433
b	0.012	-0.026	0.014	-0.047	-0.006	0.012
c	0.012	0.013	0.014	0.041	-0.006	0.012
d	3.312	3.912	3.912	3.376	3.912	3.912
n_1	61.834	85.192	65.877	69.486	126.004	88.989
n_3	36.823	49.919	43.771	37.596	57.874	52.855

Table E.4: Parameters of GT distributions, R_t, groundings

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
No. Cases	2812	7035	9847	3116	7323	10439
Coefficients						
β_0	-1.274	-0.348	-0.694	-6.431	-9.818	-8.648
β_v	2.339	2.590	2.438	3.356	5.204	4.597
MLR	-1044	-1365	-2518	-984	-1981	-3003
Pearson Test	0	0	0	1	0	0
Deviance Test	1	1	1	0.831	1	1
Hosmer-Lemeshow Test	0	0	0	0.18	0	0

Table E.5: Binary logistic regression coefficients, groundings

	SH40	SH150	SHCOM	DH40	DH150	DHCOM
No. Cases						
r_{pb} (data)	0.71	0.78	0.76	0.58	0.80	0.76
p-value (data)	0	0	0	0	0	0
r_{pb} (random)	-0.02	0.00	-0.00	0.00	0.01	-0.01
p-value (random)	0.32	0.82	0.78	0.81	0.48	0.40

Table E.6: Binary logistic regression point-biserial correlation tests, groundings