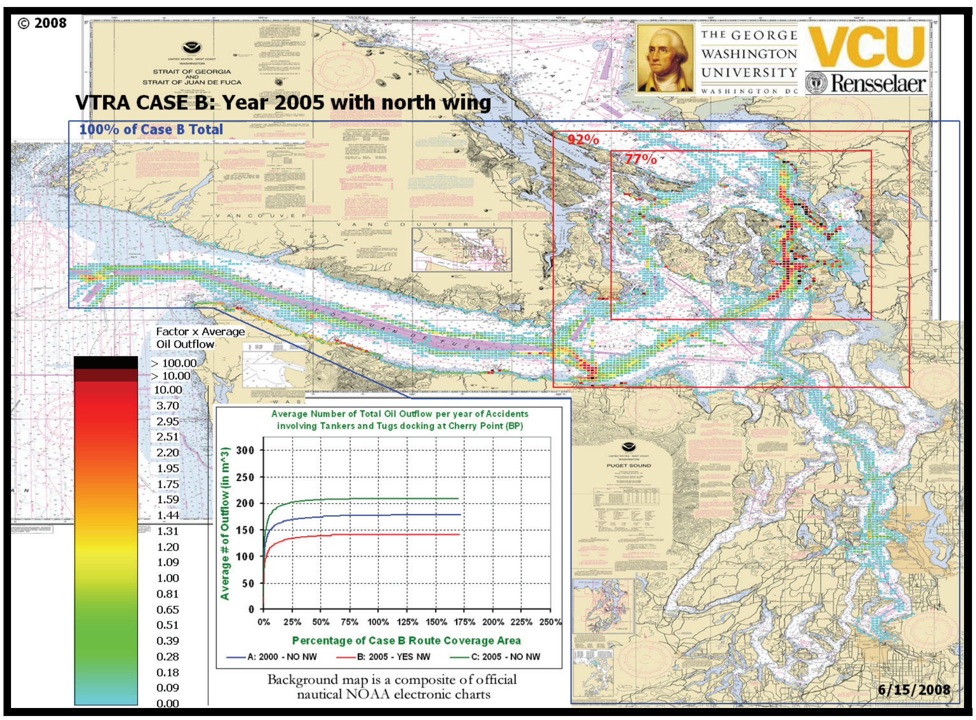
TECHNICAL APPENDIX D: EXPERT JUDGMENT ELICITATION



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 FIGURES (continued)

TABLE OF TABLES

D-1. Organizations that provided experts

Our model represents the chain of events that could potentially lead to an oil spill (see Figure D-1). This model and approach has been used in the Prince William Sound Risk Assessment, the Washington State Ferries Risk Assessment, and the Exposure Assessment of the San Francisco Bay ferries.

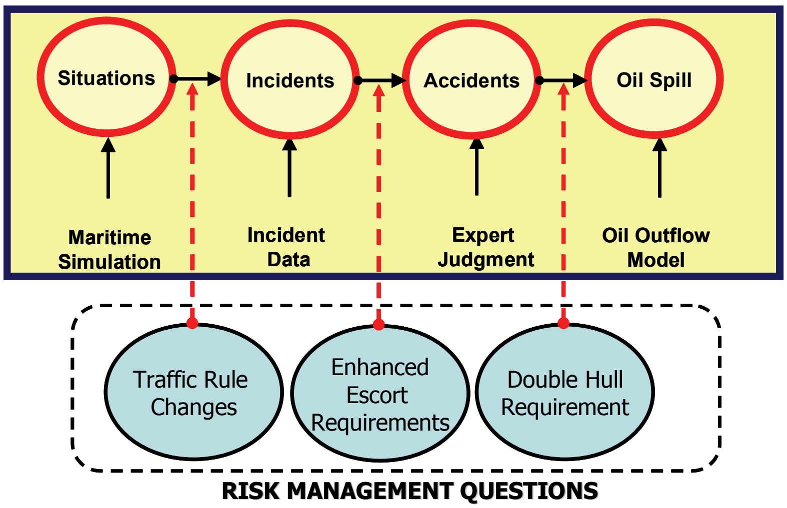


Figure D-1. Overview of a causal chain leading to an oil spill

It is based on the methodology developed for the dynamic risk simulation of tanker operations in Prince William Sound, Alaska (1995-96), for the Washington State Ferries (WSF) Risk Assessment (1998-1999) and for the San Francisco Bay Exposure Assessment (2002). The overall methodology is described in the following journal papers:

- J.R.W. Merrick, J.R. van Dorp, J.P. Blackford, G.L. Shaw, T.A. Mazzuchi and J.R. Harrald (2003). "A Traffic Density Analysis of Proposed Ferry Service Expansion in San Francisco Bay Using a Maritime Simulation Model", *Reliability Engineering and System Safety*, Vol. 81 (2): pp. 119-132.
- J.R.W. Merrick, J. R. van Dorp, T. Mazzuchi, J. Harrald, J. Spahn and M. Grabowski (2002). "The Prince William Sound Risk Assessment". *Interfaces*, Vol. 32 (6): pp.25-40.
- J.R. van Dorp J.R.W. Merrick , J.R. Harrald, T.A. Mazzuchi, and M. Grabowski (2001). "A Risk Management procedure for the Washington State Ferries", *Journal of Risk Analysis*, Vol. 21 (1): pp. 127-142
- **P. Szwed, J. R. van Dorp, J.R.W.Merrick, T.A. Mazzuchi and A. Singh (2006). "A Bayesian Paired Comparison Approach for Relative Accident Probability Assessment with Covariate Information", *European Journal of Operations Research*, Vol. 169 (1), pp. 157-177.**

The accident types included in this study are collisions between two vessels, groundings (both powered and drift), and allisions. However, as our maritime simulation counts the situations in which accidents could occur, it also records attributes that could affect the chance that the accident will occur; these include e.g. the proximity of other vessels, the types of the vessels, the location of the situation, and environmental variables, such as wind, current and visibility. The construction of this maritime simulation is described in Appendix C. We know how often accidents do occur from our analysis of incident and accident data. The accident and incident data collected for this particular project and its process is described in Appendix A. However, there is not enough data to say how each of these attributes affects the chances of an accident; accidents are rare! To determine this, we must turn to the experts (see the third event in Figure D-1) in maritime operations. Specifically, we must turn to experts who are primarily familiar with the sailing of tugs and tankers in the study area and preferably have long term sailing experience with either one or both of these vessel types. Experts were invited to and referred to the VTRA team through the United States Coast Guard and the Puget Sound Harbor Safety committee. The organizations that provided experts to construct our accident probability models are:

1. Puget Sound Pilots
2. ATC
3. US and Canadian Tug Companies operating in the VTRA study area:
   US-Based: Foss, Crowley, Olympic Tug and Barge (US),
   K-Sea, Sea Coast, Sause Bros.
   Canadian Based: Seaspan, Island Tug and Barge
4. The Washington State Ferries
5. Seattle sector US Coast guard VTS.

Expert judgment elicitation sessions were scheduled predominantly at the US Coast Guard VTS, sector Seattle in December 2006, February 2007, June 2007, August 2007, September 2007 and December 2007. The elicitation session with the ATC tanker captains and master was scheduled during an ATC conference in February 2007 in Portland,Oregon.

D-1.1. Questionnaires Developed

Table D-1 below summarizes the elicitation process that was followed in the overall expert judgment elicitation procedure. A total of 9 questionnaires were developed that were distributed to 38 experts over 7 separate elicitation sessions (2 elicitation sessions were held during February 2007) dispersed over a 1 year period. The combined numbers of years sailing experience of the experts who participated in the elicitation process of the VTRA study area exceeds 922 years. The number of years experience of the experts by questionnaire is further detailed in Table D-14. The last expert judgment elicitation session was held in December 2007 after which final results were analyzed and were prepared for integration into the maritime vessel traffic risk assessment simulation tool. The first expert judgment elicitation session was held in December 2006.

Table D-1. Overview of questionnaires developed for and expert experience during the VTRA expert judgment elicitations.

9 QUESTIONNAIRES	38 EXPERTS - Numbers indicate years sailing experience in VTRA Study area	CUMULATIVE EXPERIENCE (YRS)	7 SESSIONS
Bradley-Terry Pair Wise Comparison Location Questionnaire	7 PILOTS (42,34,32,25,16,16) 6 TUG OPERATORS (39, 30, 30, 30, 15, 12) 4 FERRY OPERATORS (31, 30, 25, 8) 2 PORT CAPTAINS (27, 25) 1 VTS WATCH (25)	186 156 94 52 25	Dec-06 Feb-07
Bradley-Terry Pair Wise Comparison Traffic Scenario Questionnaire	7 PILOTS (42,34,32,25,16,16) 6 TUG OPERATORS (39, 30, 30, 30, 15, 12) 4 FERRY OPERATORS (31, 30, 25, 8) 2 PORT CAPTAINS (27, 25) 1 VTS WATCH (25)	186 156 94 52 25	Dec-06 Feb-07
Bradley-Terry Pair Wise Comparison 1st Traffic Type Questionnaire	7 PILOTS (42,34,32,25,16,16) 6 TUG OPERATORS (39, 30, 30, 30, 15, 12) 4 FERRY OPERATORS (31, 30, 25, 8) 2 PORT CAPTAINS (27, 25) 1 VTS WATCH (25)	186 156 94 52 25	Dec-06 Feb-07
Bradley-Terry Pair Wise Comparison 2nd Traffic Type Questionnaire	6 PILOTS (35, 34, 24, 22, >20, >20) 5 TUG OPERATORS (53, 32, 38, 20, 18) 2 PORT CAPTAINS (32, 30)	> 155 151 62	Apr-07 Aug-07 Sep-07
Bradley-Terry Pair Wise Comparison Tug Barge Questionnaire	7 TUG OPERATORS (53, 21, 20, 32, 30, 28, 18) 2 PORT CAPTAINS (32, 30)	202 52	Aug-07 Sep-07 Dec-07
Tanker Pair Wise Situation Collision Accident Probability Questionnaires Given Propulsion Failure	6 PILOTS (35, 34, 24, 22, >20, >20) 5 TANKER OPERATORS (21, 20, 21, 18, 16)	> 155 96	Feb-07 Apr-07
Tanker Pair Wise Situation Collision Accident Probability Questionnaires Given Steering Failure, Given Navigational Aid Failure Given Human Error Given Near By Vessel Failure	6 PILOTS (35, 34, 24, 22, >20, >20) 5 TANKER OPERATORS (21, 20, 21, 18, 16)	> 155 96	Feb-07 Apr-07
Tug Pair Wise Situation Accident Probability Questionnaires Given Propulsion Failure	7 TUG OPERATORS (53, 21, 20, 32 30, 28, 18) 2 PORT CAPTAINS (32, 30)	202 52	Aug-07 Sep-07 Dec-07
Tug Pair Wise Situation Collision Accident Probability Questionnaires Given Steering Failure, Given Navigational Aid Failure Given Human Error Given Near By Vessel Failure	7 TUG OPERATORS (53, 21, 20, 32 30, 28, 18) 2 PORT CAPTAINS (32, 30)	202 52	Aug-07 Sep-07 Dec-07

We were extremely fortunate that in November 2006 the Puget Sound Harbor Safety committee agreed to provide us a platform to present interim results of the VTRA study and ask for feedback from the Puget Sound maritime community. This platform and the close relationship between the Puget Sound maritime community were instrumental in obtaining access to experts and the expert participation that we received. We were able to hold our first expert judgment elicitation session one month after the introduction to the Puget Sound Harbor Safety committee. Invitations to the expert judgement elicitation sessions were sent out initially by the US Coast Guard and later on by the Puget Sound Harbor Safety committee. None of the experts personally benefited from participating in the expert judgment elicitation. They donated their time for the enhancement of the safety levels in their maritime domain and they should be commended for it. Each expert judgment elicitation session consisted of a morning and afternoon session.

D-2. Overview of expert judgment technique

Of the four papers listed in the introduction, the fourth one Szwed et. (2006) (indicated in bold above) describes in detail how we estimate the parameters in our accident probability models using the expert judgment. For convenience it is included as a sub-appendix to this appendix. Below, we shall provide an overview of the specific implementation of this technique in this particular project.

The aim of our expert judgment elicitation technique is to be able to estimate the conditional probability of an accident given that a particular incident has occurred in a particular scenario on the water. This incident can either be a propulsion failure, a steering failure, a navigational aid failure, a human error or an event of a vessel nearby. We refer to the later incident as a NBV failure (NBV=Near By Vessel). Scenario on the water are summarized by a set of attributes and these sets of attributes are stored in a database using the maritime simulation and may be described by a vector $\underline{X}$. We shall refer to the elements of the vector $\underline{X}$ as accident attributes in the sense that the value of such an attribute may adversely affect the accident probability given that a particular incident has occurred. At what level these attributes affect an accident probability may very well depend on the incident type as well. We capture this multitude of effects via our expert judgment approach. Below we shall discuss in more detail our expert inducement procedure for our accident probability models. Separate accident probability models are constructed for tankers and tugs.

Our tanker and tug collision probability models follow the set-up in Szwed et al. (2006):

$$Pr(Collision|Incident,\, \underline{X}) = P_0 exp\left\{\underline{\beta}^T \underline{X}\right\}. \qquad \text{(D-1)}$$

Whereas in the Prince William Sound Risk Assessment (see, e.g. Merrick et. al (2002)), we used a similar formulation as (D-1) for groundings and allision accident probability models, we have enhanced the accident probability models in this project for groundings and allisions to allow for explicit representation of "a time to shore" variable t that is now also recorded in our maritime simulation. In the Prince William Sound Risk Assessment this time component was only taken into account implicitly through the attribute "Location". This, however, would not allow for modeling of a difference in convergence of the waterway within a particular location. The expressions for our accident probability models for grounding and allision are as follows:

$$Pr(Grounding|Incident, \underline{X}, t) = P_0 exp\Big\{ - \alpha_0[1 + \underline{\gamma}^T(\underline{1} - \underline{X})] \times t\Big\} \quad \text{(D-2)}$$

$$Pr(Allision|Incident, \underline{X}, t) = P_0 exp\Big\{ - \delta_0[1 + \underline{\kappa}^T(\underline{1} - \underline{X})] \times t\Big\} \quad \text{(D-3)}$$

The parameter vectors $\underline{\beta}, \underline{\gamma}, \underline{\kappa}$ describe the effect that a particular element in the attribute vector $\underline{X}$ has on the accident probability. The parameters P_0, α_0 and δ_0 are used for calibrating our maritime simulation model to the accident data that has been collected for the VTRA study area. The data collection procedure and process is described in detailed in Appendix A. Before we can estimate these parameters, however, we need to establish a measurement of scale for the various accident attributes $\underline{X}$. In the next section we shall discuss the scale development for the both the tanker and tug accident attributes.

D-2.1. Attribute scale development

Table D-2 summarizes the accident attributes $\underline{X}$ for tankers. The discretization column in Table D-2 gives the number of levels that a particular attribute may have. For example we have considered nine different locations in the accident probability model. The designations for the specific locations are specified in Table D-3. Table D-3 describes all the different levels of the various accident attributes listed in Table D-2 and that we have accounted for in our accident probability models. Tables D-4 and D-5 provide similar information for the tug accident probability models.

From Table D-2 it immediately follows that in our model the maximum number of possible situations that a tanker could encounter equals 2,156,544 or over 2 million different situations. Likewise, from Table D-3 we have modeled potentially 5,031,936 or over 5 million different situations for tugs. Needless to say, it is impossible to estimate the accident probability for each situation individuality and hence we have to resort to theoretical

Table D-2. Accident attributes for tanker accident probability models

	TANKER DESCRIPTION	DISCRETIZATION
1	Location	9
2	Direction	2
3	Cargo	2
4	Escorts	3
5	Tethering	2
	INTERACTING VESSEL	**DISCRETIZATION**
6	Vessel Type	13
7	Traffic Scenario	4
8	Traffic Proximity	2
	WATERWAY CONDITIONS	**DISCRETIZATION**
9	Visibility	2
10	Wind Direction	2
11	Wind Speed	4
12	Current	2
13	Current Direction	3

Table D-3. Levels of accident attributes for tanker accident probability models

LOCATION	DIRECTION	CARGO	ESCORTS	TETHERED
Cherry Point Area	Inbound	Unladen	2 Escorts	tethered
Puget Sound South	Outbound	Laden	1 Escort	untethered
Strait of Juan de Fuca East			No Escorts	
Strait of Juan de Fuca West				
Puget Sound North				
Saddle Bag Area				
Rosario Strait				
Haro Strait\Boundary Pass				
Guemes Channel				

VESSEL TYPE	TRAFFIC PROXIMITY	TRAFFIC SCENARIO
Tug without Barge	1 to 5 miles	Crossing Astern
Tug ATB's or ITB's	Less than 1 mile	Meeting
Tug Pushing Ahead		Overtaking
Container		Crossing the Bow
Tanker		
Bulk carrier		
Freighter		
Passenger vessel		
Service vessel		
Public vessel		
Fishing Vessel		
Tug Towing Astern		
Recreational Vessel		

VISIBILITY	WD	WIND SPEED	CURRENT	CUR_DIR
More than 0.5 mile	Along Vessel	Less than 10 knots	Almost Slack	Along Vessel - Opposite
Less than 0.5 mile	Abeam Vessel	20 knots	Max Eb or Max Flood	Along Vessel - Same Dir.
		30 knots		Abeam Vessel
		More than 40 knots		

Table D-4. Accident attributes for tug accident probability models

	TUG DESCRIPTION	DISCRETIZATION
1	Location	9
2	Direction	2
3	Cargo	7
4	Hook-up	4
	INTERACTING VESSEL	**DISCRETIZATION**
5	Vessel Type	13
6	Traffic Scenario	4
7	Traffic Proximity	2
	WATERWAY CONDITIONS	**DISCRETIZATION**
8	Visibility	2
9	Wind Direction	2
10	Wind Speed	4
11	Current	2
12	Current Direction	3

Table D-5. Levels of Accident attributes for tug accident probability models

LOCATION	DIRECTION	CARGO	HOOKUP
Cherry Point Area	Inbound	No Barge	No Barge
Puget Sound South	Outbound	Unladen Barge	ATB or ITB
Strait of Juan de Fuca East		Laden Container Barge	Pushing Ahead
Strait of Juan de Fuca West		Laden Bulk Cargo Barge	Towing Astern
Puget Sound North		Laden Derrick/Crane Barge	
Saddle Bag Area		Laden Oil Barge	
Rosario Strait		Log Tow	
Haro Strait\Boundary Pass			
Guemes Channel			

VESSEL TYPE	TRAFFIC PROXIMITY	TRAFFIC SCENARIO
Tug without Barge	1 to 5 miles	Crossing Astern
Tug ATB's or ITB's	Less than 1 mile	Meeting
Tug Pushing Ahead		Overtaking
Container		Crossing the Bow
Tanker		
Bulk carrier		
Freighter		
Passenger vessel		
Service vessel		
Public vessel		
Fishing Vessel		
Tug Towing Astern		
Recreational Vessel		

VISIBILITY	WD	WIND SPEED	CURRENT	CUR_DIR
More than 0.5 mile	Along Vessel	Less than 10 knots	Almost Slack	Along Vessel - Opposite
Less than 0.5 mile	Abeam Vessel	20 knots	Max Eb or Max Flood	Along Vessel - Same Dir.
		30 knots		Abeam Vessel
		More than 40 knots		

probability models that capture the effect on an accident probability from attribute to attribute and the effect within an attribute from level to level. The expressions for these accident probability models are given by equations (D-1, D-2, D-3) above. The first element X_1 representing "Location", the second element X_2 representing "Direction", etc. (see Tables D-2 and D-4).

The first step in creating our quantitative accident probability is to develop a measurement scale for each individual accident attribute. For some this is relatively straightforward. For example, in case of tankers X_9 represents visibility and we assign a value of 1 to "Less than 0.5 mile" and a value 0 to "More than 0.5 mile". Hence, the scale is ordered in such a manner that worse levels in an accident attribute attain a higher value. Creating such a scale for other attributes is less straightforward. For example consider the vessel type attribute in Table D-3. First of all, we have 13 levels for the vessel type attribute and while we have ordered the vessel types from best to worst in both Tables D-4 and Tables D-5, it is not all obvious if going from a "container vessel" to a "tanker" in this scale is as bad as going from a "tug towing astern" to a "recreational vessel".

An important class of elicitation techniques are the so-called the psychological scaling models that use the concept of paired comparisons. Origins of this class can be traced back to Thurstone's (1927a,b) pioneering work where Weber's and Fechner's law were used to quantify the intensity of psychophysical stimuli using a discriminative process. An extension of this concept found application in the field of consumer research (see, Bradley (1953)). An examination of the Bradley- Terry model is provided by Cooke (1991), among other numerous sources. We used the Bradley- Terry paired comparison method to develop attribute level measurement scales for the following attributes: Location, Vessel Type, Traffic Scenario, Cargo (for Tugs) and Hookup.

Figure D-2 and D-3 provide an example explanation used in one of our paired comparison questionnaires to established a scale for the traffic scenario attribute. As part of our Institutional Review Board procedure regarding research involving human subjects it is a requirement that the expert remains anonymous. However, the experts were asked to provide their job title and number of years of sailing experience (see Figure D-1) in the VTRA area (although they were not forced to provide this information to participate in the survey). It was explained to the experts that every effort will be made to keep their provided information confidential. There were instructed that if any of the questions they were asked as part of this study made them feel uncomfortable they could refuse to answer that question.

Explanation of an Example Question

Your responses to this questionnaire will be used to develop an accident probability model to be used in conjunction with a Maritime Simulation for the Vessel Traffic Risk Assessment (VTRA) Study. The results of this questionnaire will allow a preliminary ordering of different traffic scenarios from a collision perspective. You will be asked to compare traffic scenarios through pair wise comparison. Before you compare two traffic scenarios, you are asked to imagine a vessel type common to both these traffic scenarios under particular weather conditions. For the comparison, this should preferably be a worst case collision scenario.

The responses you provide will be anonymous but we ask you to indicate your experience level with the marine environment in the VTRA study area above by stating your affiliation and the number of years you have been exposed to the VTRA study area.

Job Title:
Number of Years Experience:

Figure D-2. Example introduction of a paired comparison questionnaire for accident attribute scale development

An example question is as follows. Let the two traffic scenarios you are comparing be: Overtaking and Meeting situations. Next, you are asked assuming **<u>a common vessel type for both traffic scenarios</u>** to indicate the traffic scenario for which you would be more concerned for a collision to occur and you are asked to indicate your answer in the following format.

Overtaking	<--	=	-->	Meeting	?

If you are **equally** concerned about an overtaking and meeting situation you answer:

Overtaking	<--	X	-->	Meeting	?

If you are **more** concerned about an overtaking than a meeting situation you answer:

Overtaking	X--	=	-->	Meeting	?

If you are **less** concerned about an overtaking than a meeting situation you answer:

Overtaking	<--	=	X	Meeting	?

If you cannot answer this question, you answer:

Overtaking	<--	=	-->	Meeting	X

Figure D-3. Example explanation of a paired comparison question in a paired comparison questionnaire for accident attribute scale development

WHICH ONE CONCERNS YOU MORE?

Meeting	<--	=	-->	Crossing the Bow	?
Overtaking	<--	=	-->	Crossing the Bow	?
Overtaking	<--	=	-->	Crossing astern	?
Overtaking	<--	=	-->	Meeting	?
Crossing astern	<--	=	-->	Crossing the Bow	?
Crossing astern	<--	=	-->	Meeting	?

Figure D-4. Example explanation of a paired comparison question in a paired comparison questionnaire for accident attribute scale development

They were allowed to take a break at any time during the study. They could stop their participation in this study at any time. It was explained to the experts that they will not benefit directly from their participation in the study, but rather that the benefits that might result from this study are to science, humankind and a scientific and impartial assessment of oil spill risk due to potential increased vessel traffic at Cherry Point, WA. If results of this research study are reported in journals or at scientific meetings, the people who participated in this study will not be named or identified.

Figure D-3 provides the format of the explanation of a paired comparison question, whereas Figure D-4 list all the paired comparison questions for the traffic scenario questionnaire. Since we are comparing pair wise four traffic scenarios we have a total of $\binom{4}{2} = 6$ questions. The Bradley-Terry paired comparison technique allows for testing the consistency of an expert. An expert commits what is called "a circular triad" if the expert responds $A > B$, $B > C$, but $C > A$.

The more circular triads are present within his/her expert judgment the less consistent the expert. Of course, the question arises how many circular triads would be too many. This is (naturally) also a function of the number of pairwise comparison question he/she is asked to answer. The Bradley Terry methods considers an expert consistent if his/her number of committed circular triads compares favorable to a hypothetical expert responding at random. This is conducted via a statistical hypothesis test. If an expert had less than a 5% chance of having the number of circular triads if the expert had responded at random, the expert was deemed consistent. Otherwise, his/her responses were not considered in the analysis. Besides allowing for testing the inconsistency within an individual's expert judgment, the Bradley Terry method allows for testing agreement amongst the expert judgments. This is achieved by measuring the association of the various rankings from the individual experts through what is called a "measure of concordance". Higher values of this measure indicate a higher level agreement. A statistical test is formulated that evaluates a threshold such that there would be less than a 5% chance of achieving this measurement of concordance assuming all the expert rankings were independently generated (and thus not exhibiting agreement).

Figure D-5 provides the resulting scales from the Bradley Terry analysis for the four different Traffic Scenarios resulting from the responses of 13 consistent experts (with agreement amongst these experts). From Figure D-5 it follows that in terms of level of concern a "crossing the bow" situation is about 6.6 times worse than a "meeting" situation and an "overtaking" situation is about twice worse. Moreover, a "crossing astern" situation is approximately 7.7 times better than a meeting situation (in level of concern) making it about 52 times better than a "crossing the bow" situation.

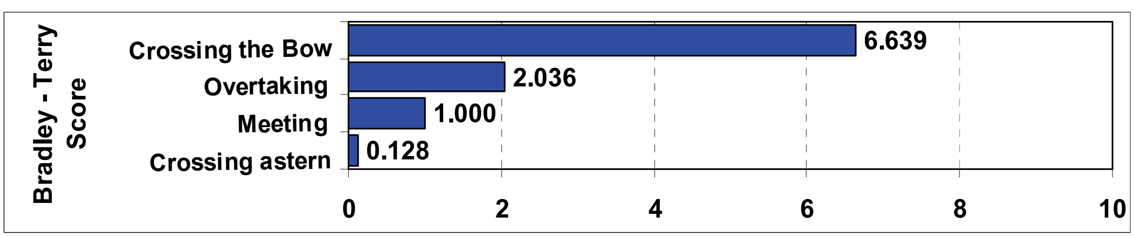


Figure D-5. Attribute scale for traffic scenario using tanker and tug operator responses

Figure D-6 provides the resulting scales from the Bradley Terry analysis for the nine different locations that we considered in the expert judgment elicitations. These scores followed from 8 consistent experts (with agreement amongst the experts). The definition of these nine different locations were provided to experts prior to the elicitation (see Figure D-

7). From Figure D-6 it follows that "Guemes Channel" is considered to be about 11 times worse in level of concern than the "Strait of Juan de Fuca East". Furthermore, it appears that "Haro-Strait\Boundary Pass" is similar in level of concern than "Rosario Strait" and the same applies to the grouping "Puget Sound North", "Stait of Juan de Fuca West", "Strait of Juan de Fuca East" and "Puget Sound South". The "Sadde Bag" area falls somewhere in between "Rosario Stait" and "Puget Sound North". Finally, the "Cherry Point Area" is about 2.7 times better in level of concern than the "Strait of Juan de Fuca East".

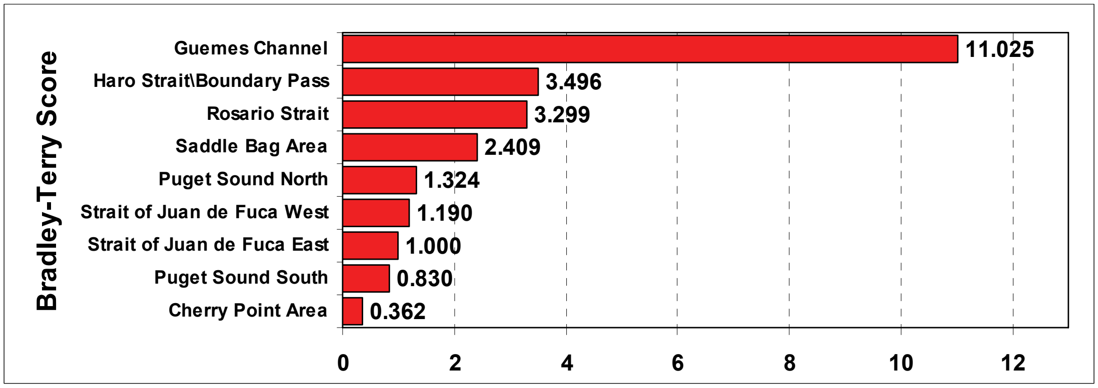


Figure D-6. Attribute scale for Locations using tanker and tug operator responses

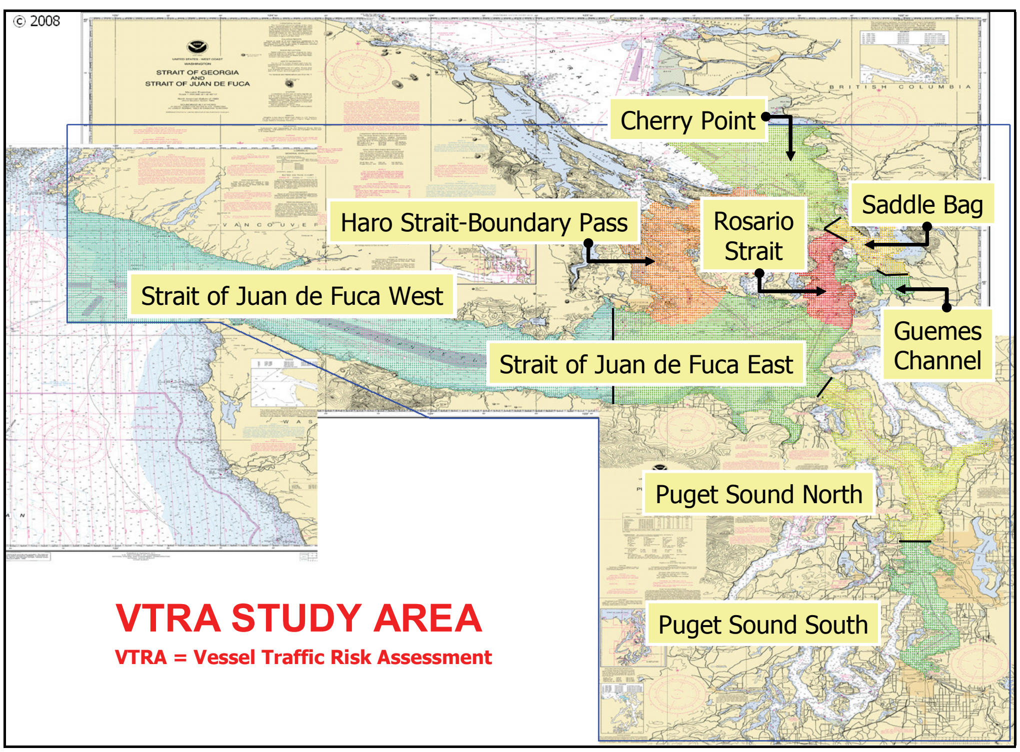


Figure D-7. The Vessel Traffic Risk Assessment (VTRA) study area and the definition of its nine different locations for expert judgment purposes

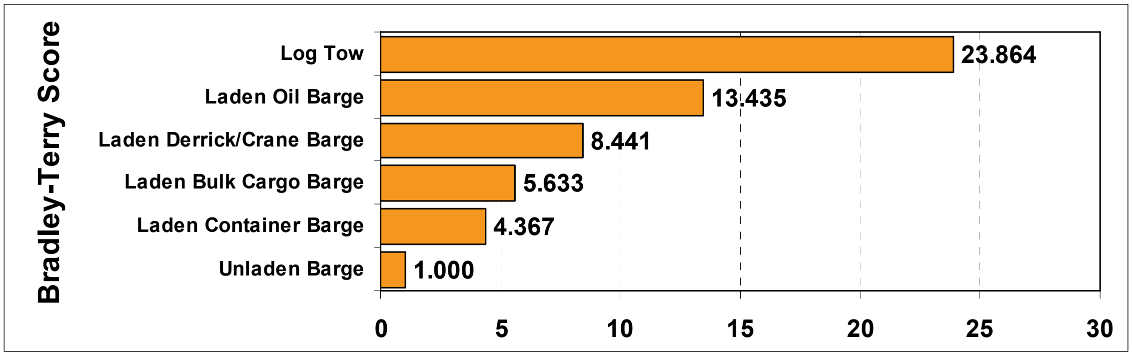


Figure D-8. Attribute scale for tug barges using tug operator responses.

Figure D-8 above provides the resulting scales from the Bradley Terry analysis for the six different barge configurations that we considered in the expert judgment elicitations. These scores followed from 8 consistent experts (with agreement amongst the experts). From Figure D-8 it follows that the "Log Tow" configuration obtains the highest level of concern (from a towing perspective) and the "unladen barge" the lowest level of concern. Laden "Bulk Cargo" and "Container" barges seem to obtain somewhat similar scores, where the "Laden Oil" barge obtain the second highest score in level of concern followed by the "Laden Derrick/Crane" barge.

The last attribute for which we constructed a scale using a Bradley Terry type analysis is the "Vessel Type" attributed listed in Tables D-2 and D-4. From Tables D-3 and D-5 it follows that we are considering 13 different vessel types in the various accident probability models. The full set of paired comparisons for that case would be $\binom{13}{2} = 78$ questions which could be considered too tasking resulting potentially in a proportionally larger number of triads and thus inconsistency in the expert judgment. In attempt to avoid such an adverse result, the development of the vessel type scale was developed using initially one questionnaire of 9 vessel types (involving 36 questions). The second questionnaire of 6 vessel types (involving 15 questions) was born from the observation amongst the Puget Sound Harbor Safety committee members that when encountering tugs that how the mariner views them depends on its tow configuration. Tables D-6 and D-7 list the classifications of vessel types provided to the experts for both questionnaires. Both questionnaires had the "tanker" and "passenger" vessel in common which allowed for the merging of the vessel scales that followed.

Table D-6. Vessel type classifications of initial vessel type scale questionnaire

	Vessel Type	Sub-Classification
1	Tanker	oil, chemical, product, LNG
2	Container	
3	Freighter	
4	Bulk carrier	
5	Tug/tow/barge/service vessel	
6	Passenger vessel	ferry, passenger ship, cruise lines, tour boat
7	Public vessel	USCG, USN, USNS, NOAA, etc.
8	Fishing Vessel	fish vessels and factories
9	Recreational Vessel	Yacht, Kayak, Jet Ski, etc.

Table D-7. Vessel type classifications to allow for a further refinement of the vessel type scale.

	Vessel Type	Sub-Classification
1	Tanker	Oil, Chemical, Product, LNG
2	Tug without Barge	
3	Tug Pushing Ahead	
4	Tug Towing Astern	
5	Tug ATB's or ITB's	
6	Passenger vessel	Ferry, Passenger Ship, Cruise Lines, Tour Boat

Figure D-8 provides the resulting scales from the Bradley Terry analysis for the initial nine different vessel types. These scores followed from nine consistent experts (with agreement amongst the experts). Figure D-9 provides the resulting scales from the Bradley Terry analysis for refinement of the vessel type scales to allow for a differentiation of tow configurations. These scores also followed from nine consistent experts (with agreement amongst the experts). Figure D-10 merges both scales from Figures D-8 and D-9. From Figure D-10 we may observes that a "recreational vessel" obtains the highest level of concern from an interaction perspective followed by a "tug towing astern". The other tug configurations "Tug Pushing ahead", "Tug ATB or ITB" and "Tug without a Barge" obtain the three smallest scores (with a much smaller score for the last one). Please note that we would not have achieved this distinction had we not further refined the vessel scale in Figure D-8. That is, the tug/tow/barge/service vessel score of 1.031 in Figure D-8 would have been the combined score for all these configurations. We can also observe from Figure D-10 that "service vessel, passenger vessel, freighter, bulk carrier, tanker and container" vessels classify in a similarity group from a vessel type perspective (when encountering them). The "Fishing Vessel" follows the "Tug Towing Astern", followed by the "Public Vessel".

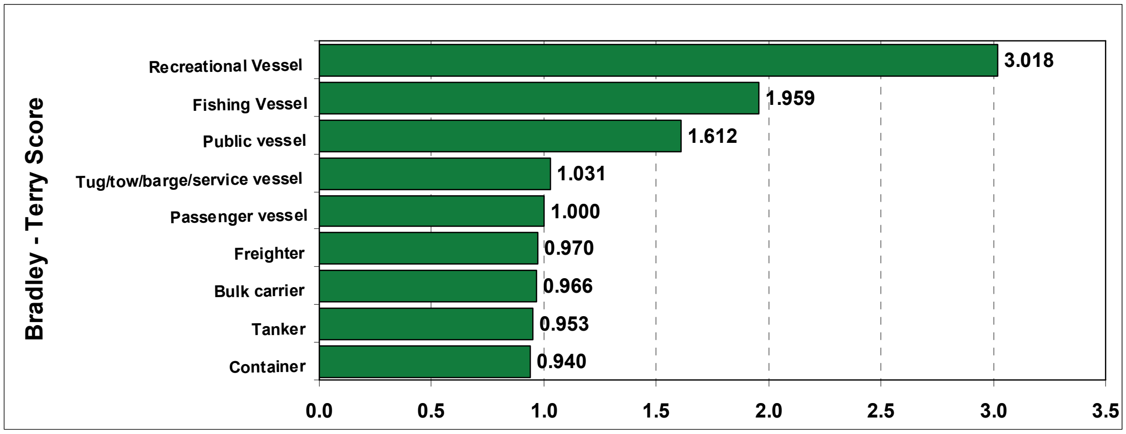


Figure D-8. Initial attribute scale for tug barges using tanker and tug operator responses.

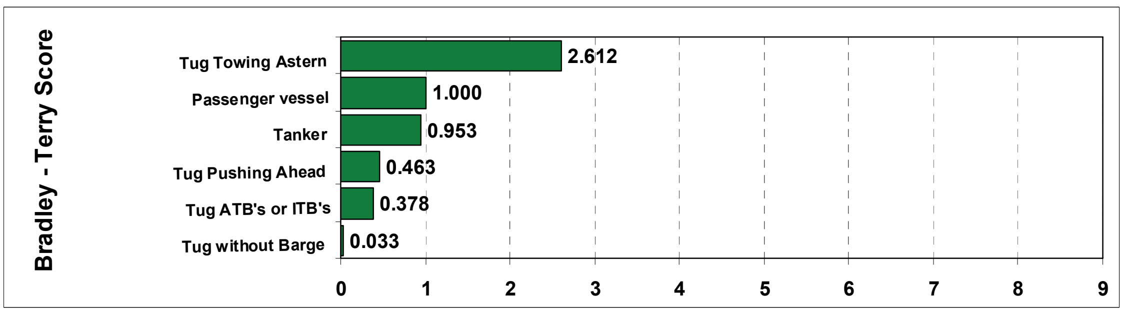


Figure D-9. Refined attribute scale for vessel types using tanker and tug operator responses.

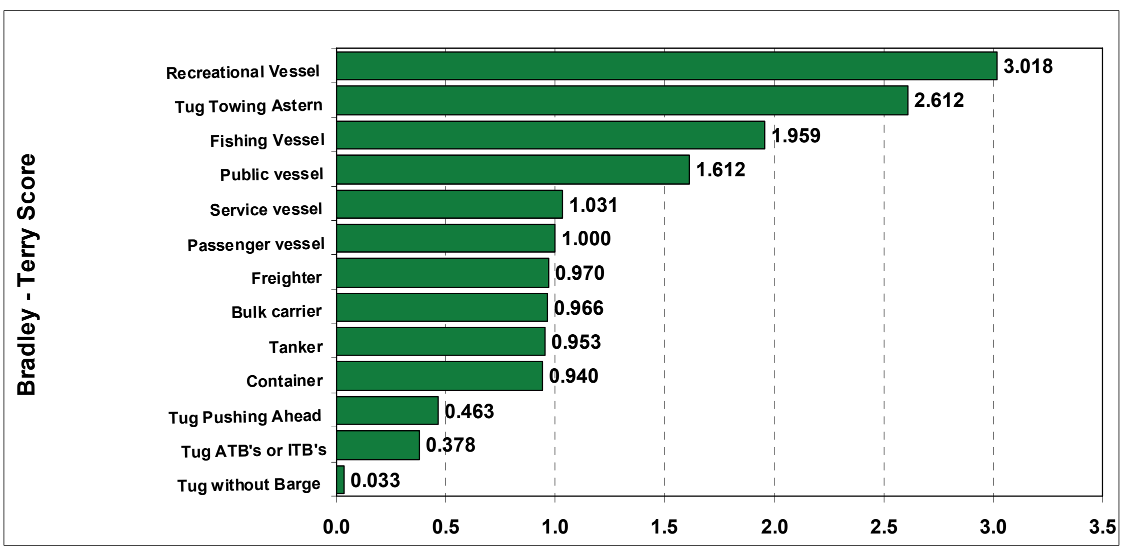


Figure D-10. Merged attribute scale for vessel types that follows from the scales presented in Figures D-9 and D-8.

Experts expressed that the US Navy vessels are of a higher concern within the "Public Vessel" classification in Table D-5, explaining possibly the relative high ranking of these vessels in the vessel type scale of Figure D-10.

While we did initially (in October 2008) arrive at individual consistency amongst 10 experts (5 pilots, 4 Ferry Masters and 1 Tug Master) in the vessel attribute scale, we unfortunately did not reach an agreement amongst these experts. When supplementing the expert judgment, however, with responses from three additional consistent tug operators (obtained in December 2008) while omitting the responses from the ferry masters, we did arrive at agreement amongst the tanker and tug operators. A possible reason for this phenomenon is that ferry masters evaluate waterway participants differently than tanker and tug operators. Following this outcome regarding vessel type scale development, it was decided for consistency to only use the tug and tank operators responses for the scale development of also the accident attributes "Location" and "Traffic Scenario" displayed in Figures D-5 and D-6. Scale developments for barges (Figure D-8) only involved tug operators from the start since only they have the appropriate experience level.

D-2.2. Attribute Parameter Assessment

Recall our collision accident probability model set-up specified by equation (D-1):

$$Pr(Collision|Incident,\ \underline{X}) = P_0 exp\left\{\underline{\beta}^T \underline{X}\right\}$$

In (D-1) an incident can either be a propulsion failure, steering failure, navigational aid failure, human error or, finally, a nearby vessel failure. The n-dimensional vector $\underline{X}$ describes particular situation on a waterway in terms of accident attributes. Attributes for the tanker and tug accident probability models are defined in Tables D-2 through D-5. The previous section discussed the development of quantitative measurement scales for the elements of the vector $\underline{X}$.

Prior to assessment of the parameter vector $\underline{\beta}$ all accident attributes $\underline{X}$ scales are pre-normalized on a $[0, 1]$ scale such that the vector $\underline{X} = \underline{0}$ describes the least "risky" situation and the vector $\underline{X} = \underline{1}$ describes the most "risky" situation. While some accident attributes have a natural ordering (such as bad visibility ($X = 1$) being worse than good visibility ($X = 0$) others required the use of expert judgment Bradley-Terry Paired comparison questionnaires to arrive at such an ordering (see Figures D-5, D-6, D-8 and D-10). From (1) we have:

$$\underline{X} = \underline{0} : \qquad Pr(Collision|Incident, \underline{X}) = P_0$$

$$\underline{X} = \underline{1} : \qquad Pr(Collision|Incident, \underline{X}) = P_0 exp\Big\{\sum_{i=1}^{n}\beta_i\Big\} > P_0 \Leftrightarrow \sum_{i=1}^{n}\beta_i > 0$$

Hence, the parameter P_0 may be interpreted as a base rate probability or the probability of a collision given the incident in the least "risky" situation. Each parameter β_i thus describes that going from best $(X_i = 0)$ to worst $(X_i = 1)$ in an accident attribute i, the base rate probability goes up by a multiplicative factor of $Exp(\beta_i) > 1$ if $\beta_i > 0$. Going from the least risky situation $\underline{X} = \underline{0}$ to the most risky situation $\underline{X} = \underline{1}$ then results in a multiplication factor of:

$$\prod_{i=1}^{n} Exp\{\beta_i\} = exp\Big\{\sum_{i=1}^{n}\beta_i\Big\} > 1 \Leftrightarrow \sum_{i=1}^{n}\beta_i > 0. \qquad \text{(D-4)}$$

The parameters β_i are estimated using expert judgment elicitation by fixing the incident type and asking a series of paired comparisons questions. In each question an experts is asked "how much more or less likely" a collision is to occur in Situation 1 ($\underline{X}_1$) compared to Situation 2 ($\underline{X}_2$) given the occurrence of an incident. His or her answer gives us, for this particular comparison of Situations 1 and 2, the value of:

$$\frac{Pr(Collision|Incident, \underline{X}_1)}{Pr(Collision|Incident, \underline{X}_2)} = \frac{P_0 exp\Big\{\underline{\beta}^T \underline{X}_1\Big\}}{P_0 exp\Big\{\underline{\beta}^T \underline{X}_2\Big\}} = exp\Big\{\underline{\beta}^T [\underline{X}_1 - \underline{X}_2]\Big\} \qquad \text{(D-5)}$$

Taking the natural logs on both sides of (D-5) results in:

$$\ln\left\{\frac{Pr(Collision|Incident, \underline{X}_1)}{Pr(Collision|Incident, \underline{X}_2)}\right\} = \underline{\beta}^T [\underline{X}_1 - \underline{X}_2]. \qquad \text{(D-6)}$$

From (D-6) it follows that the parameters β_i may now be estimated via a linear regression method on the log responses of the experts to a series of paired comparison questions. Details of our regression method are described in Szwed *et. al* (2006). The context for the example analysis in Szwed *et. al* (2006) was the Washington State Ferry Risk Assessment. In this study a total of 8 experts were used for the parameter assessment part of the collision probability model. In this VTRA study, 11 experts provided responses for the tanker collision accident probability model and 9 experts for the tug collision accident probability model.

Similar to the Bradley-Terry questionnaires the responses to the questionnaires are anonymous. Experts were told that the information they provide through the survey will be aggregated with that from other responders to link the occurrence of an incident (a failure that creates an unsafe situation) on the tanker or tug with the likelihood of a collision with another vessel. During a first questionnaire the incident in question was the propulsion failure (see Figure D-11).

Before starting the expert judgment elicitation session the graphical format (see Figure D-12) of an example question was explained to the experts. Figure D-7 was provided to explain the location attribute and Table D-6 was provided to explain the vessel type attribute in the example question of Figure D-12. It was explained that in the example question of Figure D-12 that in SITUATION 1 ON THE LEFT, an 'INBOUND' 'LADEN' tanker is en route in the 'STRAIT OF JUAN DE FUCA EAST'. It is being escorted by '1 ESCORT VESSEL' that is 'UNTETHERED'. A 'CROSSING THE BOW' situation is occurring with a 'SHALLOW DRAFT PASSENGER VESSEL' that is 'LESS THAN 1 MILE' away. The visibility is 'MORE THAN 0.5 MILE' and a wind of 'LESS THAN 10 KNOTS' is blowing with a direction 'ALONG' the tanker. The current is 'ALMOST SLACK' and the residual current is 'ALONG TANKER - OPPOSITE DIRECTION'. SITUATION 1 differs from SITUATION 2 in terms of visibility only, i.e. in SITUATION 1 the visibility is good and 'MORE THAN 0.5 MILE' vessel and in SITUATION 2 the visibility is bad and 'LESS THAN HALF MILE'.

It was explained to the experts that these situations in Figure D-12 describes the traffic scenario just before the occurrence of a COMPLETE PROPULSION LOSS on the TANKER. The expert were next asked, given the occurrence of the COMPLETE PROPULSION LOSS and the two traffic scenarios, to compare the two situations in terms of the likelihood of a collision with the interacting vessel. If they thought, given the COMPLETE PROPULSION LOSS on the TANKER, collision is equally likely in both situations they could circle 1 in the scale of Figure D-12. In Figure D-12, a six is circled towards Situation 1, which would mean that the expert would have assigned a six times higher likelihood of a collision in Situation 1 as compared to Situation 2.

Explanation of an Example Question

Your responses to this questionnaire will be used to develop an accident probability model to be used in conjunction with a Maritime Simulation for the Vessel Traffic Risk Assessment (VTRA) Study. The objective of this questionnaire is to link the occurrence of an INCIDENT (a failure that creates an unsafe situation) on the TANKER with the likelihood of a COLLISION with another vessel. The INCIDENT in this questionnaire is the occurrence of a COMPLETE PROPULSION LOSS on the TANKER while underway.

This questionnaire consists of a series of PAIRED SITUATION COMPARISONS. Each situational description consists of a set of important system characteristics. The paired situations in each question differ only in the value of one characteristic. For example, the two situations may be identical except for visibility. You are asked to compare each of the two situations described and answer the following questions:

1. Given that a COMPLETE PROPULSION has occurred on the TANKER and another vessel is close-by (this is described in more detail in the situations), in which situation would collision with this vessel be more likely?
2. How much more likely is the collision in the situation you selected in question 1, compared to the other situation.

Figure D-11. Example introduction of a paired comparison questionnaire of situations for accident attribute parameter assessment.

Q30		
Situation 1	**TANKER DESCRIPTION**	**Situation 2**
Strait of Juan de Fuca East	**Location**	-
Inbound	**Direction**	-
Laden	**Cargo**	-
1 Escort	**Escorts**	-
Untethered	**Tethering**	-
	INTERACTING VESSEL	
Shallow Draft Pass. Vessel	**Vessel Type**	-
Crossing the Bow	**Traffic Scenario**	-
Less than 1 mile	**Traffic Proximity**	-
	WATERWAY CONDITIONS	
More than 0.5 mile Visibility	**Visibility**	Less than 0.5 mile Visibility
Along Vessel	**Wind Direction**	-
Less than 10 knots	**Wind Speed**	-
Almost Slack	**Current**	-
Along Vessel - Opposite Direction	**Current Direction**	-
More? : ____	9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9	**____: More?**
Situation 1 is worse	<===================X===================>	**Situation 2 is worse**

Figure D-12. Example question of a paired comparison questionnaire of situations for tanker collision accident attribute parameter assessment given a propulsion failure.

Questionnaire 1 consisted for the estimation of the parameters of the tanker collision probability model of 44 pair wise comparison questions. The question were further subdivided in three parts. During Questions 1 through 18 the "Tanker Description" varied from Situation 1 to Situation 2 in a single attribute, whereas the description of the "Interacting Vessel" and the "Waterway Conditions" were held constant. During Questions 19 through 29 the "Interacting Vessel" varied from Situation 1 to Situation 2 in a single attribute, whereas the "Tanker Description" and the "Waterway Conditions" were held constant. Finally, during Questions 30 through 44 the "Waterway Conditions" varied from Situation 1 to Situation 2 in a single attribute, whereas the "Tanker Description" and the "Interacting Vessel" were held constant.

Questionnaire 1 consisted for the estimation of the parameters of the tug collision probability model of 47 pair wise comparison questions. The questions were further subdivided in three parts. During Questions 1 through 15 the "Tug Description" varied from Situation 1 to Situation 2 in a single attribute, whereas the description of the "Interacting Vessel" and the "Waterway Conditions" were held constant. During Questions 16 through 27 the "Interacting Vessel" varied from Situation 1 to Situation 2 in a single attribute, whereas the "Tug Description" and the "Waterway Conditions" were held constant. Finally, during Questions 28 through 47 the "Waterway Conditions" varied from Situation 1 to Situation 2 in a single attribute, whereas the "Tug Description" and the "Interacting Vessel" were held constant.

Figure D-13 shows the format of the same 44 pairwise comparison questions of a second questionnaire following Questionnaire 1. The purpose of the second questionnaire is to elicit the relative likelihood of a collision accident (of a tanker in case of Figure D-13) given the other incidents: steering failure, navigational aid failure, human error or a nearby vessel failure. Questionnaire 1 focused on the collision accident given a propulsion failure on the tanker. By separating the questionnaire in two parts the experts focus in Questionnaire 1 on the paired comparison of situations (given the propulsion failure). Before answering Questionnaire 2 experts were asked to first copy their answers from Questionnaire 1 in Questionnaire 2. Hence, this provided the benefit of having their answer of Questionnaire 1, prior to answering the paired comparison of situations for the other incident types. We believe this fosters consistency in the expert responses while experts were able to focus on the differences that the various incidents may have when answering a particular comparison of two situations.

Situation 1	TANKER DESCRIPTION	Situation 2
Strait of Juan de Fuca East	**Location**	-
Inbound	**Direction**	-
Laden	**Cargo**	-
1 Escort	**Escorts**	-
Untethered	**Tethering**	-
	INTERACTING VESSEL	
Shallow Draft Pass. Vessel	**Vessel Type**	-
Crossing the Bow	**Traffic Scenario**	-
Less than 1 mile	**Traffic Proximity**	-
	WATERWAY CONDITIONS	
More than 0.5 mile Visibility	**Visibility**	-
Along Vessel	**Wind Direction**	-
Less than 10 knots	**Wind Speed**	25 knots
Almost Slack	**Current**	-
Direction	**Current Direction**	-
	Complete Propulsion Loss	
More? : ____	**9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9**	**____ : More?**
Situation 1 is worse	<===================X===================>	**Situation 2 is worse**

	Complete Steering Loss at a Moderate Angle	
More? : ____	**9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9**	**____ : More?**
Situation 1 is worse	<===================X===================>	**Situation 2 is worse**

	Complete Navigational Aid Loss	
More? : ____	**9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9**	**____ : More?**
Situation 1 is worse	<===================X===================>	**Situation 2 is worse**

	Human Error	
More? : ____	**9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9**	**____ : More?**
Situation 1 is worse	<===================X===================>	**Situation 2 is worse**

	Nearby Vessel Incident (but you do not know the specifics)	
More? : ____	**9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9**	**____ : More?**
Situation 1 is worse	<===================X===================>	**Situation 2 is worse**

Figure D-13. Example question of a paired comparison questionnaire of situations for tanker collision accident attribute parameter assessment given all incidents.

With all the responses recorded and having the attribute scales from the previous section we assess the values of the parameters $\underline{\beta}$ following our technique detailed in Szwed et. al (2006). Besides accounting for the direct effect of the attributes in Tables D-2 through D-5, we also allowed for the potential of some interaction effects. Interaction effects modeled involved "Location", "Cargo", "Escort" and "Tethered" as a group in case of tankers, and also "Fog", "Current" and "Current Directions" as a second group. Interaction effects modeled involved "Location", "Cargo", "Hookup" as a group in case of tugs and also "Fog", "Current", and "Current Directions" as a second group.

The tanker collision accident probability questionnaire given a propulsion failure consisted of 44 questions similar to the one displayed in Figure D-12. The questions were distributed evenly over the 13 accident attributes in Table 2 (i.e. 3 to 4 questions per changing attribute). The 22×22 design matrix A of the questionnaire (see Equation (11) in Szwed et. al (2006)) is of the following form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad \text{(D-7)}$$

where A_{11} is a 13×13 diagonal matrix with diagonal elements

$$(0.186, 4.0, 4.0, 6, 1.5, 3.0, 0.4716, 2.268, 3.0, 3.0,\ 4.0, 0.281, 3.0, 3.0) \qquad \text{(D-8)}$$

and associated with the main attributes factors $X_1, \ldots, X_{13}$. (The matrix A_{11} in (D-7) is a diagonal matrix since the paired comparison scenarios $\underline{X}_1$ and $\underline{X}_2$ only differed in accident attributes (see Figure D-12)). The matrix A_{22} in (D-7) is a symmetric 9×9 matrix with elements displayed in Figure D-14 and is associated with the interaction effects $X_{14}, \ldots, X_{22}$. Finally, the matrix $A_{21} = A_{12}^T$ is a sparse 9×13 matrix with only positive elements associated with the contributing factors $X_1, X_3, X_4, X_5, X_9, X_{12}, X_{13}$ that are included in the interaction effects $X_{14}, \ldots, X_{22}$. The matrix $A_{21}$is displayed in Figure D-15.

	LOC*BAL	LOC*ESC	LOC*TETH	BAL*ESC	BAL*TETH	ESC*TETH	FOG*CUR	FOG*CD	CUR*CD
LOC*BAL	.3707	.0232	.0465	.2755	.551	.	.	.	.
LOC*ESC	.0232	.1539	.0465	.3593	.	.2755	.	.	.
LOC*TETH	.0465	.0465	.2484	.	.6109	.1377	.	.	.
BAL*ESC	.2755	.3593	.	2.5	1.	1.	.	.	.
BAL*TETH	.551	.	.6109	1.	5.	.5	.	.	.
ESC*TETH	.	.2755	.1377	1.	.5	1.25	.	.	.
FOG*CUR	.	.	.	.	.	.	2.	.	.
FOG*CUR_DIR	.	.	.	.	.	.	.	2.	.
CUR*CUR_DIR	.	.	.	.	.	.	.	.	2.

Figure D-14. Matrix A_{22} in Equation (D-7).

	LOC	DIR	BAL	ESC	TETH	TT_1	TS_1	TP_1	FOG	WD	WS	CUR	CD
LOC*BAL	.1395	.	.8863	.	.	.	.	.	.	.	.	.	.
LOC*ESC	.0697	.	.	.3593	.	.	.	.	.	.	.	.	.
LOC*TETH	.093	.	.	.	.6109	.	.	.	.	.	.	.	.
BAL*ESC	.	.	1.	1.5	.	.	.	.	.	.	.	.	.
BAL*TETH	.	.	2.	.	3.	.	.	.	.	.	.	.	.
ESC*TETH	.	.	.	1.	.5	.	.	.	.	.	.	.	.
FOG*CUR	.	.	.	.	.	.	.	.	1.	.	.	1.	.
FOG*CD	.	.	.	.	.	.	.	.	1.	.	.	.	1.
CUR*CD	.	.	.	.	.	.	.	.	.	.	.	1.	1.

Figure D-15. Matrix A_{21} in Equation (D-7).

The questionnaire was designed in a manner such that the resulting questionnaire design matrix A is positive definite (and thus invertible), but equally important, involved meaningful paired comparisons consistent with realistic scenarios in the VTRA study area. The latter required maritime knowledge about the VTRA maritime transportation system acquired by the team over the course of this project.

Table D-8 below summarizes the vector $\underline{b}$ (see Equation (11) in Szwed et. al (2006)) for each of the eleven expert responses to the 44 questions in terms of

$$\sum_{j=1}^{44} q_{ij} z_j$$

for each of the accident attributes $X_i, i = 1, \ldots, 13$ and interaction effects X_i, $i = 14, \ldots, 22$. From Table D-8 we may assess the consistency in the expert judgment with respect to the ordering of the attribute scale of the elements X_i, $i = 1, \ldots, 16$ developed in the previous section. A positive (negative) value indicates agreement with the ordering of that particular scale. For example, the row in Table D-8 associated with the contributing factor TP_1 (Traffic Proximity of interacting vessel) shows that all experts responded (not surprisingly) that vessels further away pose less (immediate) collision risk. The largest discrepancy with the ordering of an attributes scale amongst the 11 experts is observed in the TT_1 (Traffic Type of interacting vessels). Four out of the 11 experts exhibit a negative response coefficient for this particular accident attribute. The elements

$$c = \sum_{j=1}^{60} z_j^2$$

(see Equation (11) in Szwed et. al (2006)) for each individual expert are provided in Table D-9.

With the matrix A, vectors $\underline{b}$, scalars c, we can update apriori attribute parameters settings of the parameter vector $\underline{\beta}$ (see equation D-1) specified in Figure D-16 using a Bayesian analysis. The resulting aposteriori parameters settings are provided schematically in Figure D-17. The parameter ranges of ≈ 80 specified in Figure D-16 are the 80% a prior credibility intervals for the parameters $\underline{\beta}$ (i.e. the lower bound represents the 10% quantile and the upper bound the 90% quantile). Please note that apriori a zero average effect is assessed for each element in the parameter vectors $\underline{\beta}$. The posterior 80% credibility interval have a much smaller range with a maximum range of approximately 1.5. This demonstrates convergence of the expert judgment. The parameter vectors $\underline{\beta}$ for the tanker collision accident probability model (given a propulsion failure) will be set equal to the midpoints of these aposteriori 80% credibility intervals. The values of these parameter settings are summarized in the first column of Table D-10.

Table D-8. The vector $\underline{b}$ summarizing expert responses (see Equation (11) in Szwed et. al (2006)) for the tanker collision accident probability questionnaire given a propulsion failure.

	EXP 1	EXP 2	EXP 3	EXP 4	EXP 5	EXP 6	EXP 7	EXP 8	EXP 9	EXP 10	EXP 11
LOC	0.000	1.929	1.308	0.922	1.340	0.685	0.860	0.897	0.598	1.010	1.046
DIR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BAL	5.011	0.000	4.159	6.174	6.215	3.178	5.704	2.890	4.394	3.178	6.215
ESC	1.946	4.494	2.079	2.485	0.000	1.792	1.733	1.445	1.040	1.956	2.890
TETH	2.890	0.000	3.466	4.159	3.296	2.485	2.890	2.079	2.079	2.079	3.178
TT_1	1.802	-1.743	-0.309	-0.164	0.659	0.347	0.233	0.651	-0.169	0.325	0.313
TS_1	-2.536	1.091	2.773	2.808	2.306	-0.047	2.349	2.022	3.735	2.022	1.395
TP_1	5.951	4.159	4.564	5.257	7.313	2.890	2.197	2.485	1.792	2.773	5.375
FOG	5.347	6.908	4.159	4.787	3.584	3.466	3.584	2.708	3.296	2.079	5.375
WD	5.886	3.892	0.000	0.000	0.693	4.605	1.099	1.099	1.386	-1.792	1.386
WS	0.318	1.084	1.192	0.672	0.260	0.780	0.520	0.824	0.520	0.000	0.520
CUR	-1.099	0.000	5.455	3.401	1.792	2.773	2.079	1.386	0.000	1.386	4.159
CD	3.584	0.000	1.386	3.401	0.000	1.099	1.386	0.693	0.693	0.693	1.386
LOC*BAL	1.144	1.345	1.918	2.025	2.454	1.262	2.045	1.157	1.422	1.437	2.160
LOC*ESC	0.536	1.897	1.034	0.983	0.472	0.687	0.745	0.772	0.511	0.956	1.126
LOC*TETH	0.559	1.032	1.429	1.383	1.317	0.747	1.033	1.021	0.722	0.959	1.174
BAL*ESC	3.555	4.494	3.466	4.277	1.609	2.890	3.342	1.445	2.138	3.055	4.500
BAL*TETH	6.109	0.000	5.545	7.560	6.515	4.277	6.109	2.773	4.277	3.871	6.397
ESC*TETH	2.495	2.996	1.733	2.079	0.549	1.445	1.936	1.445	1.040	1.956	2.485
FOG*CUR	0.000	2.303	3.258	2.890	2.485	2.079	1.792	1.609	1.099	1.386	3.178
FOG*CD	2.890	2.303	1.386	2.708	1.099	0.000	2.079	0.000	1.099	0.693	1.792
CUR*CD	2.996	0.000	3.178	3.219	0.693	2.485	1.386	1.386	0.693	0.693	2.773

Table D-9. The scalars c summarizing expert responses (see Equation (11) in Szwed et. al (2006)) for the tanker collision accident probability questionnaire given a propulsion failure.

EXP 1	EXP 2	EXP 3	EXP 4	EXP 5	EXP 6	EXP 7	EXP 8	EXP 9	EXP 10	EXP 11
81.113	91.757	59.527	61.779	70.927	36.457	36.245	28.921	28.885	26.179	60.289

The parameter settings for the tanker collision accident probability model given the remaining incidents are solved for in a similar manner. While the paired comparison questions remained the same (and thus also the questionnaire design matrix A), a separate set of response vectors $\underline{b}$, and scalars c follow for each remaining incident type: steering failure, navigational aid failure, human error and nearby vessel failure. The parameter settings of the vectors $\underline{\beta}$ for the collision tanker accident probability model are summarized in Table D-10. Table D-11 summarizes the parameters setting for the tug collision accident probability model for each incident type.

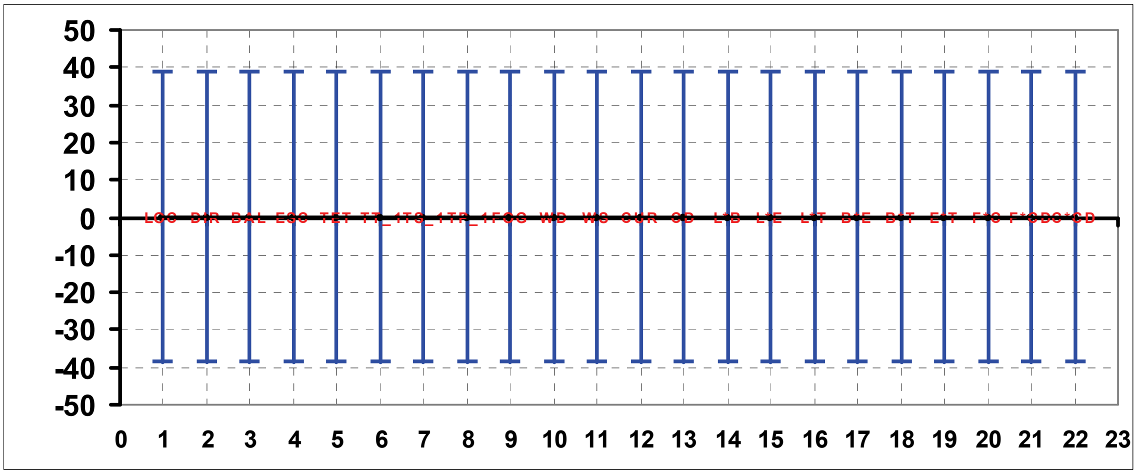


Figure D-16. Apriori specification of tanker accident attribute parameters given a propulsion failure (prior to updating with the expert judgment responses).

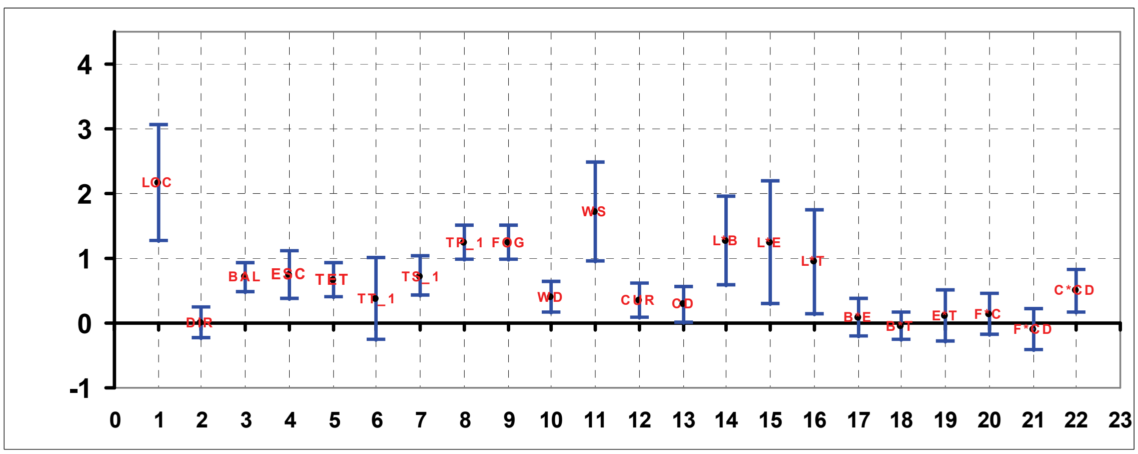


Figure D-17. Aposteriori tanker accident attribute parameters given a propulsion failure (after updating with the expert judgment responses).

The parameter P_0 in expression (D-1) does not follow from the expert judgment elicitation. Instead we solve for this parameter through a calibration step after the relative collision accident probability models for tugs and tankers have been integrated in a maritime simulation of the waterway. This maritime simulation records waterway situations as described by the attribute vectors $\underline{X}$ in Tables D-1 through D-4 in a database. Since these situations share the common base rate probability P_0 we may solve for P_0 by setting the expected number of collisions during a simulation run over a period equal to the empirical average annual number of collisions in that same period. Table D-12 provides the values for

P_0 given the different incident types for the tanker and tug collision accident probability model.

Table D-10. Attribute accident parameters for tanker accident probability models

ID	NAME	Propulsion Failure	Steering Failure	Nav. Aid Failure	Human Error	NBV Failure
1	LOC	2.164	3.038	1.642	2.969	1.785
2	DIR	0.000	-0.040	-0.015	-0.036	-0.024
3	BAL	0.700	0.876	0.542	0.937	0.691
4	ESCORTS	0.745	0.876	0.394	0.626	0.356
5	TETHERED	0.652	1.058	0.408	0.885	0.462
6	TT_1	0.369	0.185	-0.095	-0.063	0.183
7	TS_1	0.715	0.979	0.486	0.943	0.887
8	TP_1	1.231	1.607	0.847	1.387	1.138
9	FOG	1.247	1.413	1.442	1.446	1.310
10	WD	0.399	0.588	0.203	0.458	0.333
11	WS	1.708	1.631	1.037	1.440	1.429
12	CURRENT	0.345	0.831	0.565	0.814	0.599
13	CUR_DIR	0.278	0.475	0.477	0.436	0.426
14	LOC*BAL	1.271	1.746	0.879	1.631	1.020
15	LOC*ESC	1.229	1.424	0.673	1.265	0.768
16	LOC*TETH	0.940	1.554	0.803	1.242	0.743
17	BAL*ESC	0.084	0.392	0.119	0.260	0.117
18	BAL*TETH	-0.047	0.074	0.029	-0.001	0.000
19	ESC*TETH	0.104	0.104	-0.018	-0.076	-0.076
20	FOG*CUR	0.130	0.268	0.420	0.270	0.191
21	FOG*CUR_DIR	-0.109	0.024	0.261	0.171	0.130
22	CUR*CUR_DIR	0.490	0.237	0.107	0.146	0.051

Table D-11. Attribute accident parameters for tug accident probability models

ID	NAME	Propulsion Failure	Steering Failure	Nav. Aid Failure	Human Error	NBV Failure
1	LOC	0.760	0.822	0.737	1.270	0.846
2	DIR	0.028	0.194	0.032	0.184	0.067
3	Bal	1.909	1.630	1.168	1.611	1.337
4	HKP	1.336	1.482	0.865	0.981	0.876
5	TT_1	0.762	0.910	0.269	0.701	0.595
6	TS_1	0.661	0.654	0.663	0.820	0.825
7	TP_1	1.227	1.421	0.791	1.505	1.015
8	VIS	1.286	1.478	1.393	1.632	1.138
9	WD	1.145	1.024	0.558	0.862	0.701
10	WS	3.341	3.425	1.756	3.059	1.992
11	CUR	1.503	1.568	0.854	1.507	1.108
12	CUR_DIR	1.233	1.024	0.655	0.883	0.796
13	LOC*BAL	0.765	0.737	0.560	0.868	0.638
14	LOC*HKP	0.351	0.354	0.278	0.516	0.392
15	BAL*HKP	1.389	1.313	0.856	1.158	0.908
16	FOG*CUR	0.260	0.201	0.288	0.216	0.199
17	FOG*CUR_DIR	0.285	0.236	0.433	0.254	0.143
18	CUR*CUR_DIR	0.326	0.223	0.264	0.124	0.104

In our prior studies the grounding accident model had exactly the same form as (D-1), but separate grounding base rate probability P_0 were solved for by calibrating to an empirical average number of grounding accidents. Whereas the collision accident probability calibration used interaction counts with other vessel for calibration purposed, the grounding accident probability models only used the interaction counts of the tanker with the system and thus this was a purely time-based analysis. The grounding model in our prior studies was not able to directly take into account the congestion of a waterway and was only able to accommodate that indirectly through the location accident attribute. This analysis was next followed by a separate drift grounding simulation to determine likely locations of groundings.

Table D-12. Calibrations values for P_0 for the tanker and tug collision accident probability models

	Po
Propulsion Failure	1.90743E-05
Steering Failure	1.90743E-05
Nav. Aid Failure	1.90743E-05
Human Error	2.15758E-05
NBV Failure	2.15758E-05

The grounding accident probability model (D-2) in this VTRA project is improved over the grounding model in the Prince William Sound Risk Assessment in the sense that it now explicitly accounts for the congestion of a waterway. We now record, within the maritime simulation, the time t to shore depending on the distance of a vessel from the shore and whether the vessel would be drifting to shore or would be under power. A powered grounding is interpreted as a grounding preceded by a human error, navigational aid failure or a nearby vessel failure. When the vessel is under power, a 5 hour straight track line is projected in the direction of the vessel to the closest shore point and we record the shore location and the amount of time to shore t in addition to the same accident attributes $\underline{X}$ as specified in Tables D-2 through D-5. Our motivation is here that those shore points that have a tanker or tug coming directly towards it more frequently, have a higher likelihood of power grounding keeping everything else the same. If a 5 hours track line does not intersect with the shore line, we assume that no interaction with the shore is occurring resulting effectively in a zero grounding probability for that case. Within the VTRA study area it would seem reasonable that it would be highly likely that a vessel traveling in a straight line for 5 hours would obtain a course correction as result of the external vigilance from the Canadian VTS or Seattle VTS. The counting procedure above is followed except in the case when a vessel has started docking procedures for a certain dock and is within one mile of its intended dock. In the latter case, we consider the interactions above to be allision interactions since the vessel intentionally tries to get close to shore in that case.

A drift grounding is interpreted as grounding preceded by a propulsion failure or a steering failure. When the vessel is drifting we project a drifting path taking into account wind direction and speed, current direction and speed, and the vessel slowing down through the water as the result of a loss of propulsion. We evaluate the amount of time to shore t in addition to the same accident attributes $\underline{X}$ as in (D-1). We project here also a 5 hour time path. The same counting procedure as above applies to this 5 hour threshold as well. It would seem impossible given the established external vigilance within the VTRA study area that a vessel would be drifting for more than 5 hours without some form of intervention occurring in the mean time.

It is important to stress that when we evaluate the time to shore t and the location of the shore point interaction, that we make the assumption that the drift path or straight line path is not altered within this 5 hour time frame by some form of intervention. This too seems unlikely given the safeguards and vigilance already provided by the Puget Sound Pilots, the US Coast Guard, the Canadian Coast Guard and the other VTRA Study area users. However, how one would respond to an actual occurrence of an incident (as opposed to a simulated one in our simulation) involves making tactical decisions that takes the exact situation into account and not only the abstraction of reality that we have created in our maritime simulation. Indeed it would be impossible for us to model the complex human responses to such incidents occurring and evaluate the shore line interaction location accordingly. Hence, a disclaimer of our grounding analysis results is warranted in the sense that our analysis results should be used to make strategic (long term) decisions regarding waterway risk. Our geographic profile analysis results only display a tendency towards areas with higher and lower grounding accident rates keeping a broader risk management perspective in mind.

Returning to the development of and recalling the grounding accident probability model (D-2)

$$Pr(Grounding|Incident, \underline{X}, t) = P_0 exp\Big\{ -\alpha_0[1+\underline{\gamma}^T(\underline{1}-\underline{X})]\times t\Big\}$$

$$= P_0 exp\Big\{ -\alpha_0 t - \alpha_0\underline{\gamma}^T(\underline{1}-\underline{X})t\Big\} = P_0 exp\Big\{-\alpha_0 t\Big\} exp\Big\{-\alpha_0\underline{\gamma}^T(\underline{1}-\underline{X})t\Big\}$$

we observe the probability of grounding decreasing when the time to shore t increases in the equation above. If the time to shore becomes very large (or goes to infinity) the grounding probability model goes to 0 under the conditions that

$$\alpha_0 > 0 \text{ and } \sum_{i=1}^{n}\gamma_i > 0.$$

Recalling that accident attributes $\underline{X}$ are all pre-normalized on a $[0,1]$ scale such that the vector $\underline{X} = \underline{0}$ describes the least "risky" situation and the vector $\underline{X} = \underline{1}$ describes the most "risky" situation, we have from (D-2):

$\underline{X} = \underline{0}$:

$$Pr(Grounding|Incident, \underline{X}, t) = P_0 exp\Big\{ -\alpha_0 t \Big\} exp\Big\{ -\alpha_0 \Big[\sum_{i=1}^{n} \gamma_i \Big] t \Big\} \quad \text{(D-9)}$$

$\underline{X} = \underline{1}$:

$$Pr(Grounding|Incident, \underline{X}, t) = P_0 exp\Big\{ -\alpha_0 \times t \Big\} \quad \text{(D-10)}$$

The parameter α_0 may thus be interpreted as the exponential rate of decrease in the probability of grounding as a function of time to shore in the most risky state $\underline{X} = \underline{1}$. Each parameter γ_i describes that by going from worst $(X_i = 1)$ to best $(X_i = 0)$ in an accident attribute this probability of grounding goes down by a multiplicative factor of $0 < exp\Big\{ -\alpha_0 \gamma_i t \Big\} < 1$. Going from the most risky situation $\underline{X} = \underline{1}$ to the least risky situation $\underline{X} = \underline{0}$, this probability of grounding goes down by a multiplicative factor of

$$0 < exp\Big\{ -\alpha_0 \Big[\sum_{i=1}^{n} \gamma_i \Big] t \Big\} < 1. \quad \text{(D-11)}$$

The parameters γ_i are envisioned to be estimated using expert judgment elicitation by fixing the incident type and by asking a series of paired comparisons questions. In each question an experts is asked "how much more or less likely" a grounding is to occur in Situation 1 ($\underline{X}_1$) compared to Situation 2 ($\underline{X}_2$) given the occurrence of an incident and still having t_q time to respond, where t_q is fixed for the entire questionnaire. The expert's answer would gives us, for a particular comparison of Situations 1 and 2, the value of:

$$\frac{Pr(Grounding|Incident, \underline{X}_1, t_q)}{Pr(Grounding|Incident, \underline{X}_2, t_q)} = \frac{exp\Big\{ -\alpha_0 [\underline{\gamma}^T (\underline{1} - \underline{X}_1) + 1] t_q \Big\}}{exp\Big\{ -\alpha_0 [\underline{\gamma}^T (\underline{1} - \underline{X}_2) + 1] t_q \Big\}} \Leftrightarrow$$

$$\ln \Big[\frac{Pr(Grounding|Incident, \underline{X}_1, t_q)}{Pr(Grounding|Incident, \underline{X}_2, t_q)} \Big] = -\alpha_0 t_q \Big\{ [\underline{\gamma}^T (\underline{1} - \underline{X}_1) + 1] - [\underline{\gamma}^T (\underline{1} - \underline{X}_2) + 1] \Big\} \Leftrightarrow$$

$$\ln\left[\frac{Pr(Grounding|Incident,\underline{X}_1,t_q)}{Pr(Grounding|Incident,\underline{X}_2,t_q)}\right] = \{\alpha_0\, t_q \underline{\gamma}\}^T [\,\underline{X}_1 - \underline{X}_2] \qquad \text{(D-12)}$$

Now, substituting $\underline{\beta} = \{\alpha_0\, t_q \underline{\gamma}\}$ in (D-12) yields

$$\ln\left[\frac{Pr(Grounding|Incident,\underline{X}_1,t_q)}{Pr(Grounding|Incident,\underline{X}_2,t_q)}\right] = \underline{\beta}^T [\,\underline{X}_1 - \underline{X}_2] \qquad \text{(D-13)}$$

Please note that the right hand side of expression (D-13) is exactly the same as that of the right hand side of expression (D-6) when substituting

$$\underline{\beta} = \{\alpha_0\, t_q \underline{\gamma}\}^T \qquad \text{(D-14)}$$

Hence, similar to the accident probability models in the Prince William Sound Risk Assessment (see, Merrick et. al 2002) one could significantly reduce the expert judgment elicitation burden by reusing the parameter values β_i in Tables D-9 and D-10, provided we separately recalibrate the maritime risk simulation using grounding data and a separate counting routine to record powered and drift grounding interactions of a tanker or a tug with the shoreline (while recording the time to shore t). Further substitution of (D-14) into (D-2) yields,

$$Pr(Grounding|Incident,\underline{X},t) = P_0 exp\Big\{-\alpha_0 t\Big\} exp\Big\{-\underline{\beta}^T(\underline{1}-\underline{X})\frac{t}{t_q}\Big\} \qquad \text{(D-15)}$$

and we may use t_q as a calibration constant similar to P_0 in the accident probability model.

Accessibility to experts during the Prince William Sound Risk Assessment (see, Merrick et al. 2002) was provided and guaranteed via a formal steering committee consisting of all stake holders. In the VTRA project expert judgment participation relied primarily on the willingness of experts (not directly affiliated with the project through their employer) to donate their time, without benefits to them other than that the results of the VTRA study could result in a "safer" waterway. We heavily relied on established relationships between the US Coast Guard and VTRA Waterway Participants and the Puget Sound Marine Exchange to arrange for elicitation session with tankers and tug boat operators. Despite this set-up we were able to muster the participation of 38 tanker and tug boat operators over seven separate elicitation session held over the course of one year. The participation of the experts to this study is greatly appreciated and these experts should be commended for their unselfish effort.

Unfortunately, however, the response rate to the organized elicitation sessions invitations at the US Coast Guard VTS decreased dramatically over time to the point that it became apparent that we had exhausted the available tanker and tug operator expert pool for this VTRA project. As soon as this became apparent over the course of the VTRA project, the use of expression (D-15) seems warranted. Moreover, when experts were asked informally the question (after the collision elicitation session) if their answers in the paired comparison scenario questionnaires would change if the accident scenario would have changed from a collision to a grounding, experts responded "no".

Further substitution of $\underline{X} = 1$ in (D-15) yields (see also (D-10)) :

$$Pr(Grounding|Incident, \underline{X} = 1, t) = P_0 exp\Big\{ - \alpha_0 t \Big\}. \qquad \text{(D-16)}$$

and from (D-16) we have;

$$\frac{Pr(Grounding|Incident, \underline{X} = 1, k)}{Pr(Grounding|Incident, \underline{X} = 1, (k-1))} = \frac{P_0 exp\Big\{ - \alpha_0 k \Big\}}{P_0 exp\Big\{ - \alpha_0 (k-1) \Big\}} = exp(- \alpha_0). \qquad \text{(D-17)}$$

Hence, we may interpret (D-17) such that in the worst state ($\underline{X} = 1$), the probability of a grounding reduces by a factor of $exp(- \alpha_0)$. We propose to set $\alpha = ln(2)$. Indeed, in the absence of additional information, it would seem to be reasonable to assume that in a worst case scenario there is a $50 - 50\%$ chance that one would be able to perform a save on the vessel in distress in one additional available hour of time to respond. Over five hours this yields for the worst state ($\underline{X} = 1$) :

$$Pr(Grounding|Incident, \underline{X} = 1, 5) = P_0 (\frac{1}{2})^5 = \frac{P_0}{32} \qquad \text{(D-18)}$$

For the least risky state ($\underline{X} = 0$) we obtain:

$$Pr(Grounding|Incident, \underline{X} = 0, 5) = P_0 (\frac{1}{2})^5 exp\Big\{ - \Big[\sum_{i=1}^{n} \beta_i\Big] \frac{5}{t_q} \Big\}. \qquad \text{(D-19)}$$

After calibration to 1 grounding accident per 11 years (this process will be discussed in more detail in the next section), we arrived at a value of $t_q = 0.834375$, a calibration value
$P_0 = 0.52831297$
for the incident types "Human Error" and "Nearby Vessel Failure" and a calibration value $P_0 = 0.405335373$ given the incident types "Steering Failure", "Propulsion Failure" and "Navigational Aid Failure". Hence, with the parameters settings β_i for tankers and tugs in

Tables D-10 and D-11 we evaluate from (D-18) and (D-19) the values in Tables D-13 and D-14 for $Pr(Grounding|Incident, \underline{X}, 5)$ given least risk state ($\underline{X} = 0$) and the most risky state ($\underline{X} = 1$).

Table D-13. Probabilities of grounding given an incident failure in the least risk state ($\underline{X} = 0$) and a time to shore of 5 hours. These follow from (D-18) and the tanker and tug accident accident probability parameters specified in Tables D-9 and D-10.

	Tankers	Tugs
Propulsion Failure	9.729E-41	5.991E-51
Steering Failure	5.894E-53	2.756E-51
Nav. Aid Failure	8.714E-32	6.011E-35
Human Error	3.819E-47	9.576E-50
NBV Failure	4.367E-35	4.138E-38

Table D-14. Probabilities of grounding given an incident failure in the most risk state ($\underline{X} = 1$) and a time to shore of 5 hours. These follow from (D-18) and the tanker and tug accident accident probability parameters specified in Tables D-9 and D-10.

	Tankers	Tugs
Propulsion Failure	0.0127	0.0127
Steering Failure	0.0127	0.0127
Nav. Aid Failure	0.0127	0.0127
Human Error	0.0165	0.0165
NBV Failure	0.0165	0.0165

Please observe the information in Tables D-13 and D-14 to be consistent with the modeling assumption in the maritime simulation not to count interactions of a vessel with the shore when its future drifting path or straight line projection under power does not have an intersection with the shore within a five hour time frame.

Our approach towards parameter assessment of the accident probability model for allisions (D-3) is the same as that for the grounding accident probability model. The difference being primarily in the counting procedure of allision interactions. When a vessel is within one mile of its intended dock, the projected shore interactions of a drift path and a straight line path are designated as allision interactions instead of grounding interactions. Indeed, within one mile of the intended dock, docking procedures of the tankers and tugs will have commenced, speeds are lowered, escort vessels are in place and from that point on the vessel intentionally tries to get close to the shore with a specific heading towards the shore. After calibration to 2 allision accidents per 11 years (this process will be discussed in more detail in the next section), we arrived at a value of $t_q = 0.384277$, a calibration value $P_0 = 1.039155$ for the incident types "Human Error" and "Nearby Vessel Failure" and a calibration value $P_0 = 0.894719$ given the incident types "Steering Failure", "Propulsion Failure" and "Navigational Aid Failure".

D-3. Representative results of the expert judgment

An example question in the collision accident probability questionnaire given a propulsion failure on a tanker was presented in Figure D-12. This question is repeated in Figure D-18 with underneath it the prior setting of the relative likelihood of a collision in Situation 1 compared to Situation 2. We refer to this as a prior setting, since this figure presents the relative likelihood prior to updating with the acquired expert knowledge (i.e. the expert responses). Observe from Figure D-18 that apriori we assign a 50-50% chance that Situation 1 has a higher likelihood than Situation 2 (and vice versa). In Figure D-18 the changing attribute is visibility and even though it would be quite natural to assign a higher likelihood of collision in Situation 2 (bad visibility) as compared to Situation 1 (good visibility), we still apriori assign a median likelihood of 1 to this relative likelihood. A 75% a priori credibility interval (an interval with 75% chance of falling in this interval) for the relative likelihood here equals $[1/6974,\ 6974]$. Hence, with 75% we say that the relative likelihood of Situation 1 is 6974 times higher than that of Situation 2, or vice versa. Summarizing, our apriori setting does not sway in one direction or the other regardless of the changing attribute in a particular pair wise comparison question of two situations.

Next, we update this apriori relative likelihood using the expert responses and the method described in Szwed et. al (2006). Figure D-19 provides the 11 expert responses to this particular question and even though the experts do no agree, we do notice that they all assign a higher relative likelihood to Situation 2 (bad visibility) than Situation 1 (good visibility). Included in Figure D-19 is also the empirical average (slightly larger than 4) of the average responses for this particular question. The aposteriori average in for this particular question (also indicated in Figure D-19) is slightly less than 4. The reason why this aposteriori average is different from the empirical average is that in the calculation of the aposteriori average also the responses of the experts to all the other questions are taken into account. Recalling the 75% apriori credibility interval of $[1/6974,\ 6974]$, we obtain after updating with the expert responses a 90% aposteriori credibility interval for the relative likelihood of $[2.47,\ 4.90]$ and an average aposteriori relative likelihood of 3.60.

Q30

Situation 1	TANKER DESCRIPTION	Situation 2
Strait of Juan de Fuca East	Location	-
Inbound	Direction	-
Laden	Cargo	-
1 Escort	Escorts	-
Untethered	Tethering	-
	INTERACTING VESSEL	
Shallow Draft Pass. Vessel	Vessel Type	-
Crossing the Bow	Traffic Scenario	-
Less than 1 mile	Traffic Proximity	-
	WATERWAY CONDITIONS	
More than 0.5 mile Visibility	Visibility	Less than 0.5 mile Visibility
Along Vessel	Wind Direction	-
Less than 10 knots	Wind Speed	-
Almost Slack	Current	-
Along Vessel - Opposite Direction	Current Direction	-
More? : ____	9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9	____ : More?
Situation 1 is worse	<====================X====================>	Situation 2 is worse

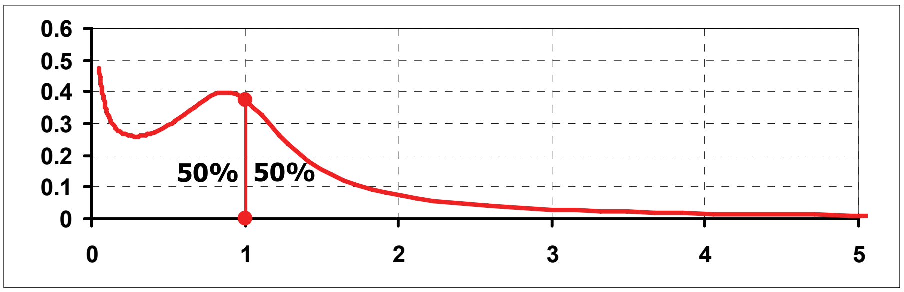


Figure D-18. Apriori setting (prior to updating with expert responses) of the relative likelihood of a collision given a propulsion failure on the tanker.

Q30

Situation 1	TANKER DESCRIPTION	Situation 2
Strait of Juan de Fuca East	Location	-
Inbound	Direction	-
Laden	Cargo	-
1 Escort	Escorts	-
Untethered	Tethering	-
	INTERACTING VESSEL	
Shallow Draft Pass. Vessel	Vessel Type	-
Crossing the Bow	Traffic Scenario	-
Less than 1 mile	Traffic Proximity	-
	WATERWAY CONDITIONS	
More than 0.5 mile Visibility	Visibility	Less than 0.5 mile Visibility
Along Vessel	Wind Direction	-
Less than 10 knots	Wind Speed	-
Almost Slack	Current	-
Along Vessel - Opposite Direction	Current Direction	-
More? : ____	9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9	10 : More?
Situation 1 is worse	<====================X====================>	Situation 2 is worse

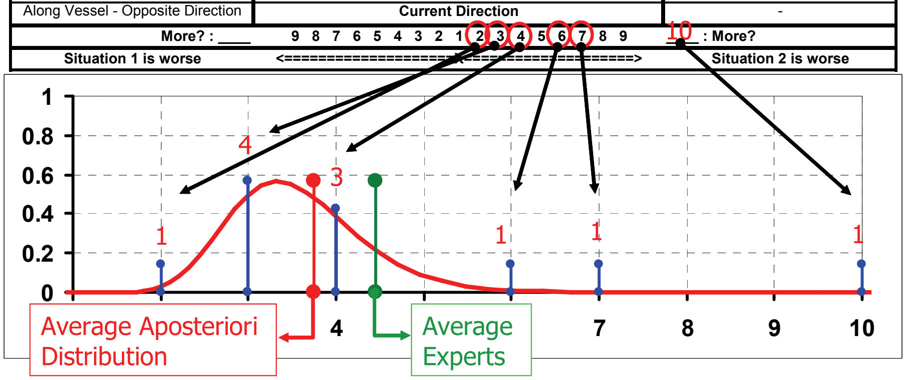


Figure D-19. Expert responses to a pair wise situation comparison to assess relative likelihood of a collision given a propulsion failure on the tanker.

Situation 1	TANKER DESCRIPTION	Situation 2
Rosario Strait	**Location**	Guemes Channel
Inbound	**Direction**	-
Laden	**Cargo**	-
1 Escort	**Escorts**	No Escorts
One Tethered	**Tethering**	Untethered
	INTERACTING VESSEL	
Shallow Draft Pass. Vessel	**Vessel Type**	-
Crossing the Bow	**Traffic Scenario**	-
Less than 1 mile	**Traffic Proximity**	-
	WATERWAY CONDITIONS	
More than 0.5 mile Visibility	**Visibility**	-
Along Vessel	**Wind Direction**	-
Less than 10 knots	**Wind Speed**	-
Almost Slack	**Current**	-
Along Vessel - Same Direction	**Current Direction**	-

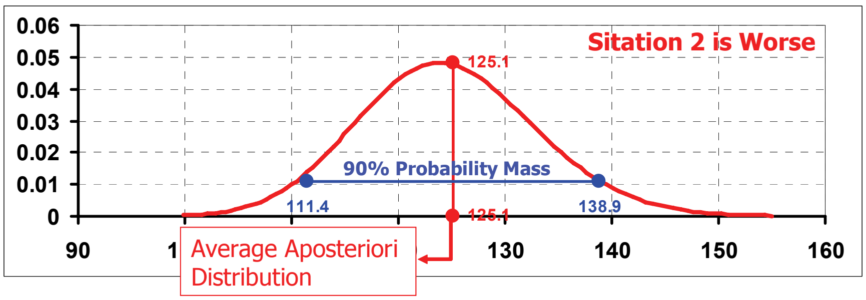


Figure D-20. Analysis of relative likelihood of a collision given a propulsion failure when three accident attributes change when going from Situation 1 to Situation 2.

Situation 1	TANKER DESCRIPTION	Situation 2
Strait of Juan de Fuca East	**Location**	Rosario Strait
Inbound	**Direction**	-
Laden	**Cargo**	Laden
2 Escorts	**Escorts**	No Escorts
One Tethered	**Tethering**	Untethered
	INTERACTING VESSEL	
Tug without Barge	**Vessel Type**	Recreational Vessel
Crossing Astern	**Traffic Scenario**	Crossing the Bow
1 to 5 miles	**Traffic Proximity**	Less than 1 mile
	WATERWAY CONDITIONS	
More than 0.5 mile Visibility	**Visibility**	Less than 0.5 mile Visibility
Along Vessel	**Wind Direction**	Abeam Vessel
Less than 10 knots	**Wind Speed**	-
Almost Slack	**Current**	Max Eb or Max Flood
Along Vessel - Same Direction	**Current Direction**	Abeam Vessel

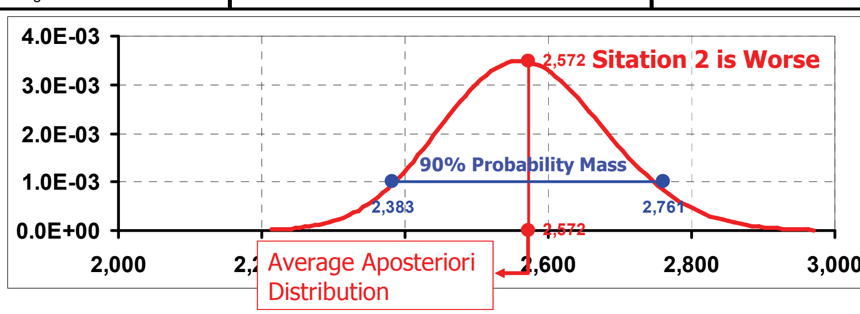


Figure D-21. Analysis of relative likelihood of a collision given a propulsion failure when eleven accident attributes change when going from Situation 1 to Situation 2.

Hence, the apriori 75% uncertainty range reduced dramatically when compared to the 90% aposteriori uncertainty range for this relative likelihood. When integrating the accident probability models with the VTRA expert judgment analysis we shall use the average aposteriori likelihoods.

After the expert judgment analysis and with the resulting parameter settings provided in Tables D-10 and D-11 our model has the ability to evaluate relative likelihoods of two situations when more than one accident attributed changes. In Figure D-20 we evaluate the relative likelihood of a collision with a shallow draft passenger vessel given a propulsion failure on the tanker when the tanker is not escorted in Guemes Channel (Situation 2) compared to the tanker being escorted and tethered in Rosario Strait. For Figure D-20 we evaluate that the collision is about 125 times more likely in Situation 2 as compared to Situation 1 (given also the settings of the remaining accident attributes in Figure D-20).

In Figure D-21 the change between Situation 2 and Situation 1 is even more dramatic. Situations 1 and 2 differ in Figure D-21 in eleven attributes. Situation 1 describes a tanker escorted by two escort vessels in the East Strait of Juan de Fuca interacting with Tug that is 1 to 5 miles away in good visibility. Situation 2 describes an unescorted laden tanker in Rosario strait with a passenger vessel crossing its bow within one mile distance in bad visibility. For Figure D-21 we evaluate that the collision is about 2572 times more likely in Situation 2 as compared to Situation 1 (given also the settings of the remaining accident attributes in Figure D-21).

With the ability of relative likelihood evaluations as in Figures D-20 and D-21, the accident probability models that evaluate accident probabilities per situation can be integrated with the VTRA simulation to evaluate annual accident frequencies. This process will be described in some detail in the next section.

D-4. Turning expert judgment into annual accident frequencies

Turning relative accident likelihoods per situation into annual accident frequencies require a calibration step and a VTRA simulation that records the values of the situation attributes needed for the accident probability models (as it simulates the maritime transportation system within the VTRA study area). In our causal chain accident probability model displayed in Figure D-1 an accident is preceded by an incident. The incidents that we have modeled are propulsion failures, steering failures, navigational aid failures, human error and a nearby vessel failure. The nearby vessel failure could either be a mechanical failure or a human error on the nearby vessel. To calibrate at the incident level we need a counting routine that is time based. The accidents that we consider in the VTRA study are collisions, drift groundings, powered groundings and allisions. Collisions involve interactions with other vessels, drift groundings involve interactions of a vessel with the shore line while adrift, powered groundings involves interactions of a vessel with the shoreline while under power and allision interactions involve interactions of a vessels with its intended dock. The separate counting mechanisms within the VTRA simulation tool will be described in the next section. In the sections thereafter we shall discuss incident and accident calibration for our simulation for the year 2005 (i.e. VTRA Case B). For this year we are effectively replaying the movement of vessels rather than having to make use of additional probabilistic traffic arrival generators. Hence, VTRA CASE B is a natural calibration scenario.

D-4.1. Simulation Counting

Consider a hypothetical interaction of a vessel with a tanker as depicted in Figure D-22. Observe that both vessels cross in Figure D-22. Informally, the level of risk could follow a profile over time as depicted in Figure D-23. That is, when the vessel are far way from another the risk is low and the closer they get it increases. At some point the risk of the vessel interaction will attain its maximum value after which it will continue to decrease and eventually return to zero when the vessels have well passed the crossing point. We attempt to capture the behavior of such a time profile in the VTRA maritime transportation simulation by discretizing the time in intervals of finite length. Whereas during the Prince William Sound Risk Assessment (see, Merrick et al. (2002)) computation efficiency only allowed us to take a snapshot of the simulation once every five minutes, we are able to take a snapshot of the VTRA simulation once every minute.

During such a snapshot the variety of interactions are evaluated and written to their "counting databases". The VTRA study area is indicated by the blue border area in Figure D-24. A counting grid is overlaid on top of this VTRA study area with grid cells that are 0.5 nautical miles by 0.5 nautical miles wide. There is a separate counting database for "route interactions" that count the amount of minutes that a vessel of interest appears in a certain grid cell. Figure D-24 displays the counting profile of route interactions when the vessel of interests are tankers, ATB's and ITB's that dock at the BP Cherry Point terminal (hereafter referred to as CHPT vessels). This counting or exposure geographic profile will be used to calibrate the probability of a propulsion failure, steering failure, navigational aid vessel or human error on a CHPT vessel during a route interaction.

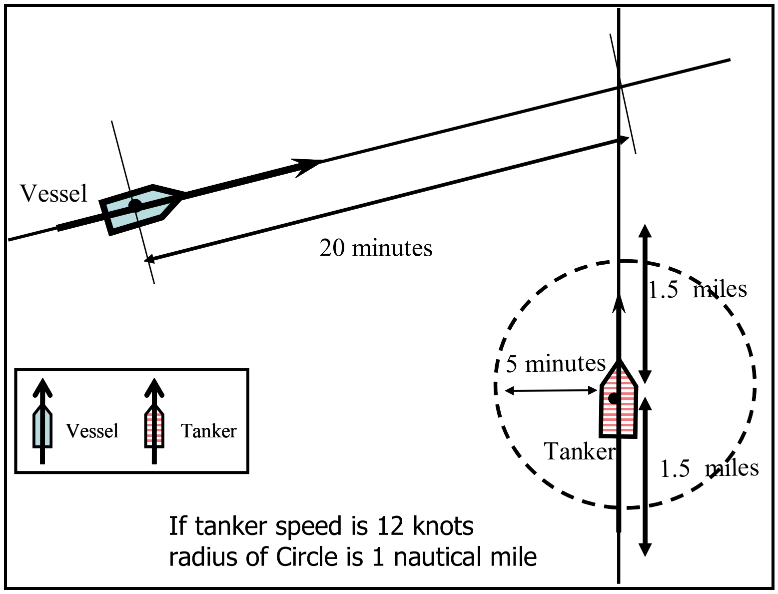


Figure D-22. Schematic of counting procedure for vessel interactions

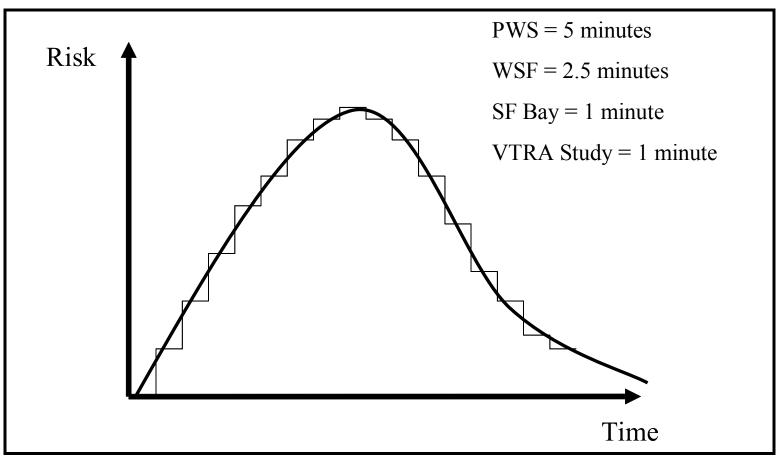


Figure D-23. A risk profile as a function of time when two vessels cross.

Figure D-25 displays the counting profile of route interactions when we count route interactions of all vessels. The counting profile in Figure D-25 is used to calibrate the probability of a nearby vessel failure during a route interaction. Please note that the color legends in Figures D-24 and D-25 are different, indicating a higher number of counts in a grid cells on both scales by a darker color.

Other "counting databases" capture vessel interactions, drift interactions, power interactions and allision interactions. A vessel interaction between a tanker and a vessel is always counted by the VTRA simulation when the interacting vessel is within a distance that

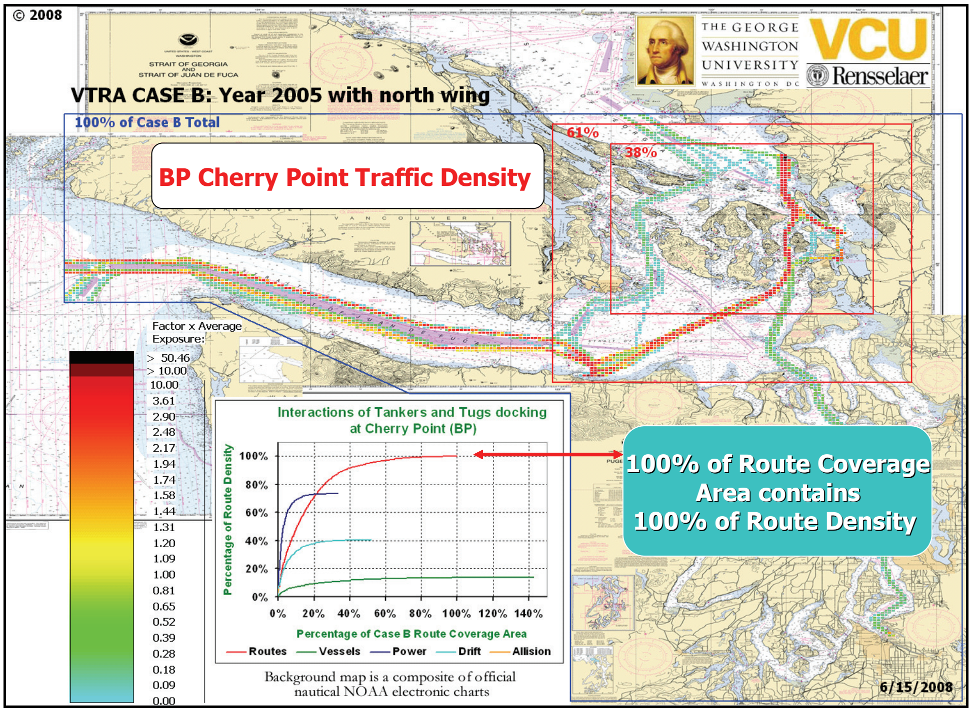


Figure D-24. Exposure Counts of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

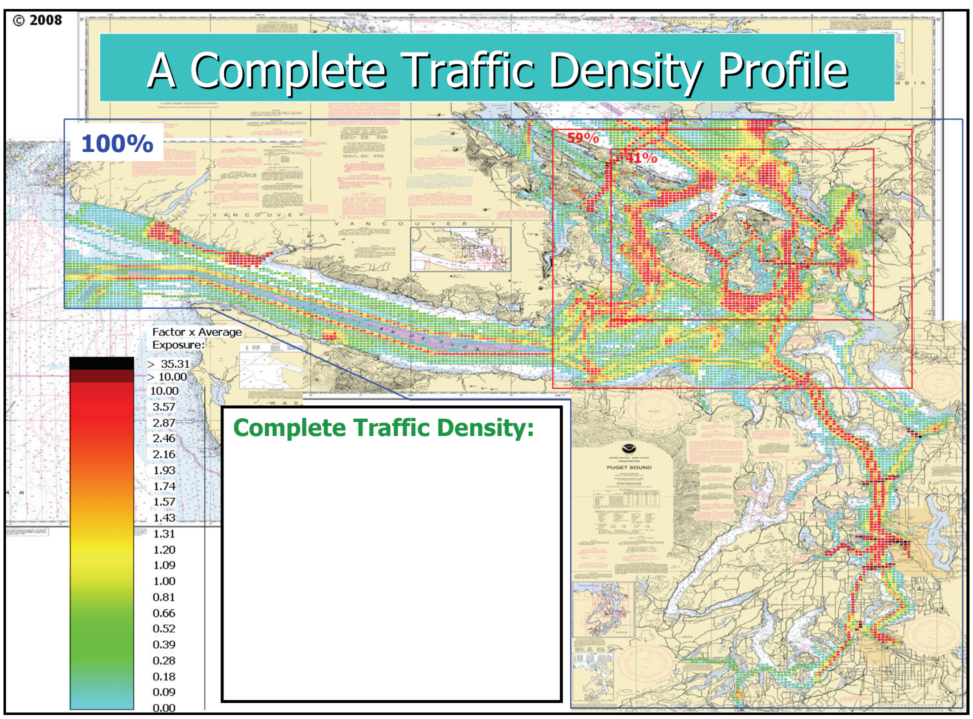


Figure D-25. Exposure Counts of all vessels in the calibration case: VTRA CASE B.

the tanker can travel within 5 minutes. Hence, when the speed of the tankers is 12 knots this distance would be one nautical mile (see Figure D-22). In previous studies this distance was fixed regardless of the speed of the vessel of interest. Our enhanced counting procedure enlarges or shrinks the vessel counting circle as function of the speed of the vessel of interest.

The left snapshot of the VTRA maritime simulation in Figure D-26 demonstrates an interacting vessel in the vessel of interest (134) "counting zone". The counting color scheme changes dynamically as the simulation counts while continuing to assign darker colors to those grid cells with a higher number of vessel interactions. The right snapshot of Figure D-26 demonstrates that it is also possible for an interaction to occur when a vessel is not within the immediate "counting zone" of the interacting vessel. This happens when the future crossing point of the interacting vessel is within 1.5 nautical miles from the front or the back of the vessel of interest and the crossing would occur within the next 20 minutes. This 1.5 nautical mile distance is set to capture the behavior of Figure D-23 when two vessels cross over time and depends on the grid cell size.

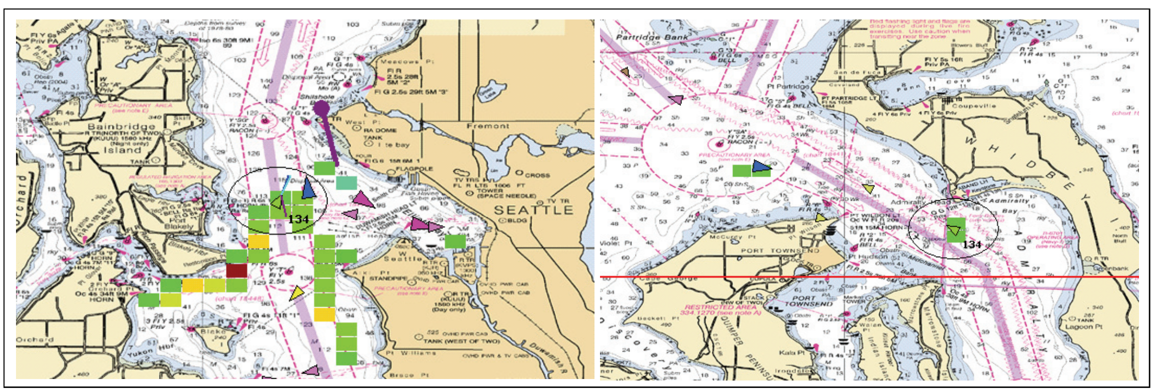


Figure D-26. Examples of vessel interaction counting in the VTRA maritime simulation.

To count drift or power interactions with the shore line when a vessel of interest is underway we first need to "define" the shore line. Figure D-27 provides this definition where each shoreline grid cell indicated in red is also 0.5 nautical miles by 0.5 nautical miles. To count drift interactions we predict the drifting path of tanker five hours out. This drifting path takes into account future wind speeds, currents and slows the tanker down over time as it drifts. The calculated future drift path follows the drift model of the NOAA (1997) publication. A drift interaction is recorded by the VTRA maritime simulation for the first grid cell that falls on this drifting path and is part of the shore definition in Figure D-27. If a five hour project drift path does not intersect with the shoreline definition, no drift interaction is counted. Both snapshots of Figure D-28 below show a drifting path of a tanker as well as the grid cells of the shore definition. In both figure a drift interaction is recorded when this drifting path intersects the shore line definition for the first time. Similar to the vessel interaction counting, the color coding of the shore line grid cells that do have drift interactions are darker when it relatively encounters a higher number of drift interactions.

To count interactions with the shore line definition in Figure D-27 for powered groundings we project a straight line following the current direction of the vessel of interest. The assumption here is that those shoreline grid cells that have more frequently a vessel of interest coming directly towards them will also have a higher powered grounding risk. These straight line projections are drawn for a distance that the vessel of interest can travel in a five hour time frame (assuming its current speed over that time frame). The first grid cell of the shoreline definition that intersects this straight line projection will obtain a power interaction count. The two snapshots of the VTRA maritime simulation in Figure D-29 demonstrates the power interaction counting algorithm for two snapshots taking shortly after one another.

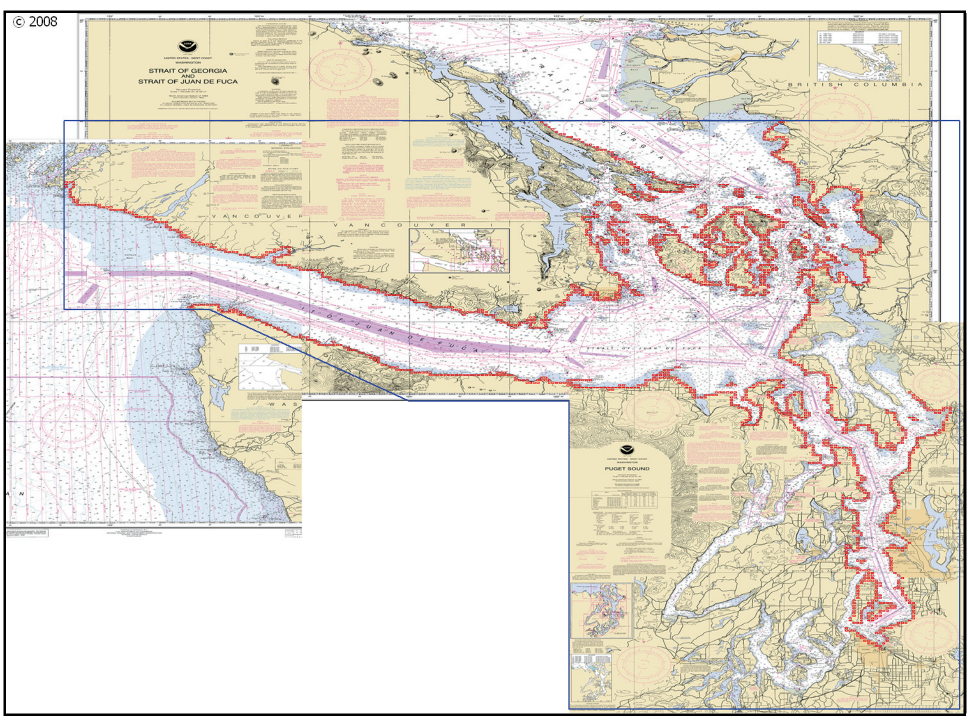


Figure D-27. Shore line definition in the VTRA maritime simulation.

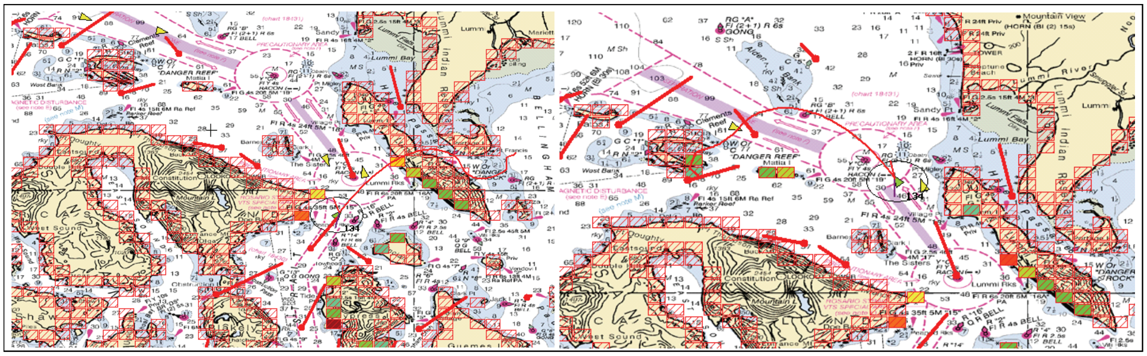


Figure D-28. Examples of drift shore line interaction counting in the VTRA maritime simulation.

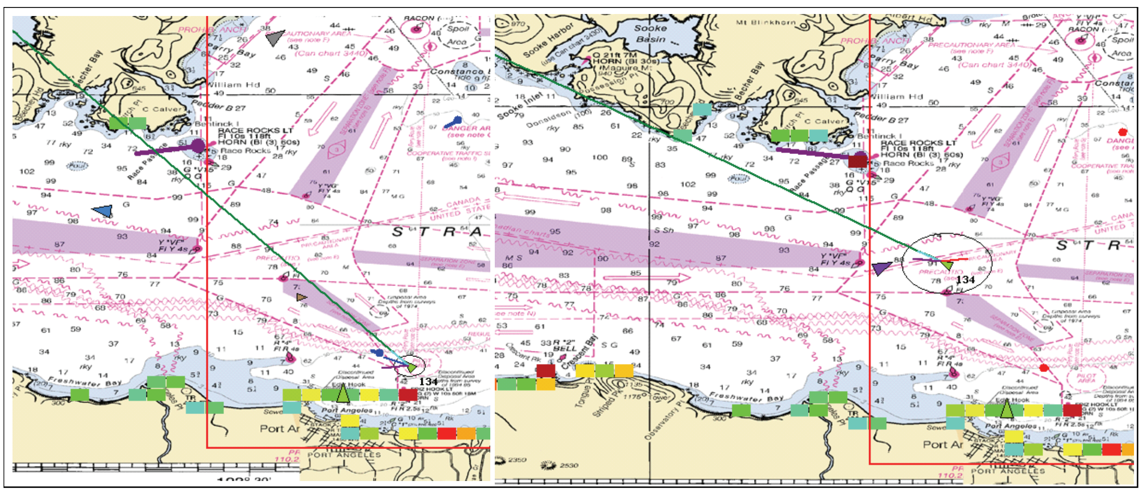


Figure D-29. Examples of power shore line interaction counting in the VTRA maritime simulation.

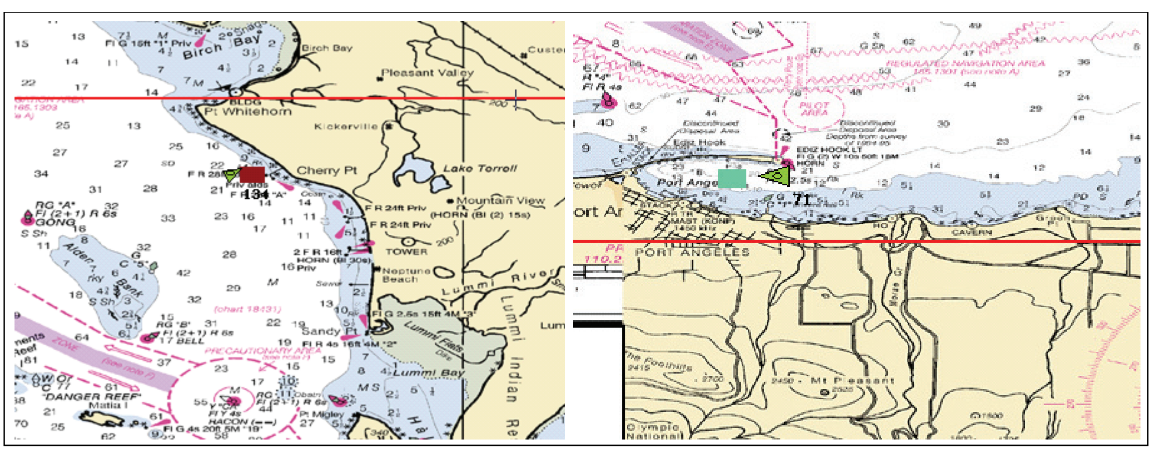


Figure D-30. Examples of allision interaction counting in the VTRA maritime simulation.

Finally, we have also implemented a counting algorithm for allision accidents. When a vessels is within one mile of its intended dock we use the straight line projection approach of the powered interactions to count allision interactions with the shore line definition in Figure D-27. From that point on neither drift or powered grounding interactions are counted anymore. Indeed, within one mile of the intended dock, docking procedures of the tankers and tugs will have commenced, speeds are lowered, escort vessels are in place and from that point on the vessel intentionally tries to get close to the shore with a specific heading towards its intended dock. Figure D-30 above shows two snapshots of the allision interaction algorithm implemented in the VTRA maritime simulation for the BP Cherry Point dock and a dock at Port Angeles.

D-4.2. Incident Calibration

From the analysis of accident and incident data it followed (See Appendix A) that over an 11 year period (1995-2005) the VTRA study area experienced 31 steering, 11 propulsion and 10 navigational aid failures on CHPT tankers totaling 52 mechanical failures. Over a 7.5 year period (the first ITB sailed about mid 1998) the VTRA study experienced 3 propulsion, 2 steering and 2 navigational aid failures on CHPT ATB's and ITB's totaling 7 mechanical failures. The data collection process in Appendix A demonstrated that human error incidents are rarely reported.

On the other hand over the data collection period (1995-2005) 4 accidents occurred, three of which were preceded by a human error. Hence, since human error incidents were rarely reported we also applied a three to one ratio (experienced at the accident level) at the incident level. Hence, with this assumption we obtain 156 human error for CHPT tankers over an 11 year period and 21 human errors for CHPT ATB's and ITB's over a 7.5 year period. These counts can next be converted to average yearly incident rates. For example, we arrive at an average annual total number of mechanical failures (i.e. propulsion, steering and navigational aid) of 5.661 and an average annual total of human errors for CHPT vessels of 22.642.

Dividing average yearly incident rates by the total number of interactions for CHPT tankers and separately by the total number of interactions for CHPT ATB's and ITB's over one year, yields the incident rates per interaction. The total number of interactions for CHPT tankers for the calibration year VTRA CASE B was 271526 and for ATB's and ITB's 172087. These counts combined resulted in the route interaction count distribution of Figure D-23. Table

D-15 present the incident rate analysis by incident type per interaction for CHPT Tankers and CHPT ATB's and ITB's to arrive at these annual totals.

Table D-15. Incident rates per route interaction for CHPT Tankers, ATB and ITB's.

	CHPT Tankers	CHPT ATB's and ITB's
Propulsion (per Year)	2.818	0.400
Steering (per Year)	1.000	0.267
Nav. Aid (per Year)	0.909	0.267
Human Error (per Year)	14.182	2.800
Annual Interactions	271526	172087
Propulsion (per Interaction)	1.038E-05	2.324E-06
Steering (per Interaction)	3.683E-06	1.550E-06
Nav. Aid (per Interaction)	3.348E-06	1.550E-06
Human Error (per Interaction)	5.223E-05	1.627E-05

Similarly, the accident incident analysis in Appendix A over the period from (1995-2005) showed a record of 1100 mechanical failure incidents and a worst case ratio of 78 to 1369 accidents that were preceded respectively by a mechanical failure or a human error. Hence, applying this worst case ratio of 17.6, we arrive at an annualized number of incidents for all vessels in the VTRA simulation of about 1855. Given a total number of interactions of all the vessels modeled in the VTRA simulation of 34519581 we arrive at an overall nearby vessel incident incident rate of 5.374E-05 per interaction. This results in a total number of nearby vessel failures for the vessels modeled in the VTRA simulation during the time that a CHPT vessel is underway of 23.84.

D-4.3. Accident Calibration

After the calibration of the VTRA simulation at the incident level, we can start the calibration process at the accident level. To calibrate VTRA simulation for a particular accident type to a given annual average number of accidents, we first need to evaluate for each recorded interaction the probability that an incident occurs and next evaluate per interaction the probability of an accident given an incident using the probability models (D-1), (D-2), (D-3). Evaluating the product of these two probabilities and summing them over all simulated interactions over one year of simulation time, yields the average annual number of accidents of that type generated by the VTRA maritime simulation. To be able to evaluate the accident probabilities given an incident using the models (D-1), (D-2), (D-3) the simulation records the accident attributes of these models. Figure D-31 displays a screen shot of this recording process for the transit of the vessel of interest 134 identified in Figure D-31. 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 during these vessel interactions. These interactions are recorded separately for route interactions, vessel interactions, drift interactions, power interactions and allision interactions as per the counting algorithms discussed in the previous sections.

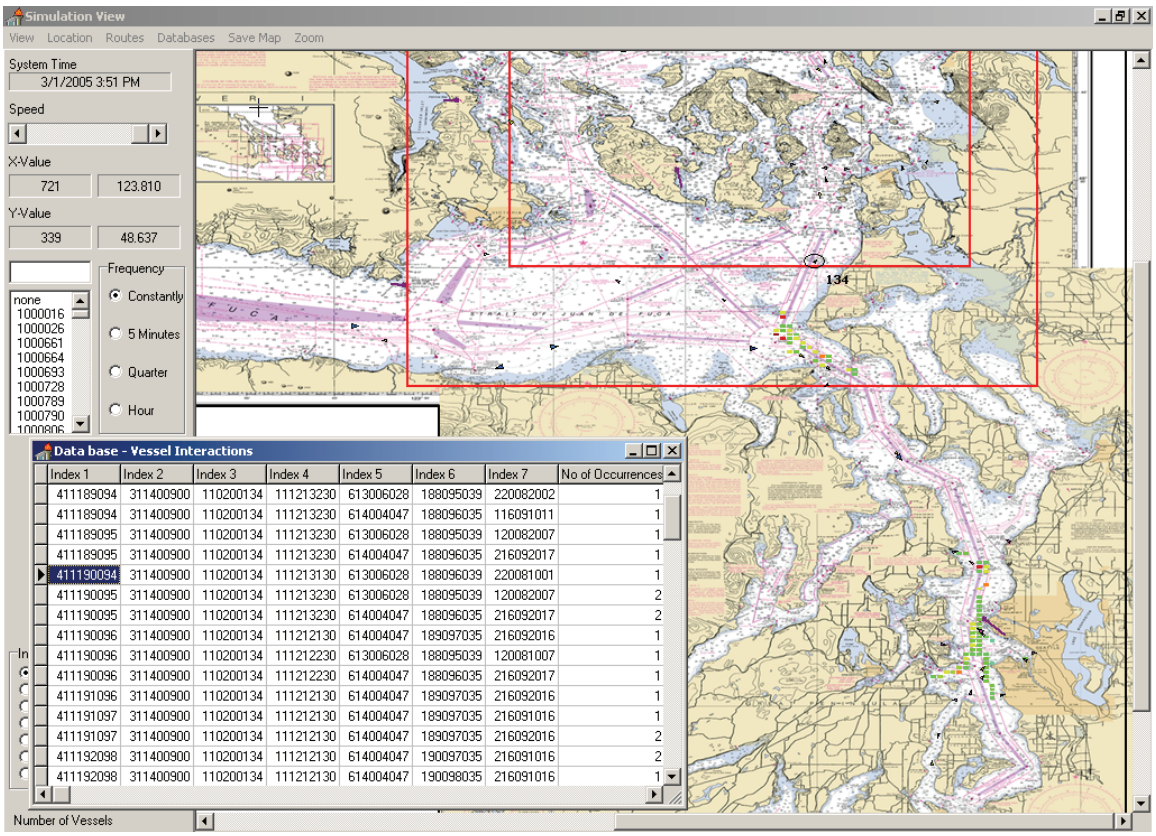


Figure D-31. Encoding of interactions by the VTRA maritime simulation.

Our accident collection process for the time period from (1995-2005) recorded 4 accidents for CHPT vessels (1 collision, 1 grounding and 2 allisions) and 3 of these accidents were preceded by a human error and 1 by a mechanical failure. During the calibrations process of our various accident type we shall maintain this ratio of 3 to 1 of average annual frequency caused by a human error or mechanical failure, for all accident types. To calibrate collisions we first ensure a ratio of 1 to 3 of frequency of collisions caused by mechanical failure compared to human errors. Next, we calibrate the collision model given a nearby vessel failure by ensuring that the average frequency of collisions caused by an incident on the CHPT vessels is the same as the average frequency of collisions caused by an incident on the near by vessel (NBV), when restricting the nearby vessel to the CHPT vessels. By a symmetry argument these average annual frequencies have to be the same. Finally, we calibrate the VTRA simulation such that on average the number of collisions per year equals $1/11$. The respective calibration values for P_0 for the collision accident probability models given an incident are provided in Table D-12.

Figures D-32 and D-33 summarize the result of the calibration step for collisions. First note that the graph in Figure D-33 shows an average return time of collisions of 11 years (equivalent to an average annual frequency of collision of about 0.09 or 1 collision in 11 years). We may also observe from this graph that approximately 42% ($\approx 60\%/140\%$) of all the grid cells that have vessel interactions, account for almost all of the total average frequency of collisions per year. The 140% in the previous calculation implies that the area of the grid cell coverage of vessel interactions is about 1.4 times the area of the grid cells through which CHPT vessel travel as displayed in Figure D-24. This follows since we do not only record the location of the CHPT vessel in these collision geographic profiles but also the location of the interacting vessel. Also observe from Figures D-32 and D-33 that the smallest red-square in Figure D-33 captures 63% of the collision frequency, whereas in Figure D-32 this red-square only captures 57% of the total vessel interactions. This difference is a direct results of overlaying the calibrated collision accident probability model (D-1) on top of the vessel interaction exposure profile of Figure D-32. When studying the color changes when going from Figure D-32 to D-33 we observe a darkening effect at the entrance from Rosario Strait to Guemes Channel, in Guemes Channel and Rosario Strait. Moreover, we observe a lightening effect at Port Angelas, East Strait of Juan de Fuca, from Rosario Strait onwards to the Cherry Point dock and possibly also a minor lightening effect in the Puget Sound area.

To calibrate to 1 grounding accident over an 11 year period of collected accident data, we first need to join the power and drift interaction database. Drift groundings in our model are those groundings that are preceded by a propulsion failure or a steering failure. The operating assumption for steering failures here is that when a steering failure occurs on a tanker, that one shuts down the propulsion and thus the vessel effectively starts drifting. Powered groundings in our models are those groundings that are preceded by a human error, navigational aid failure or a nearby vessel failure. Hence, the later incident accounts for those grounding scenarios where a vessel has to avert the nearby vessel.

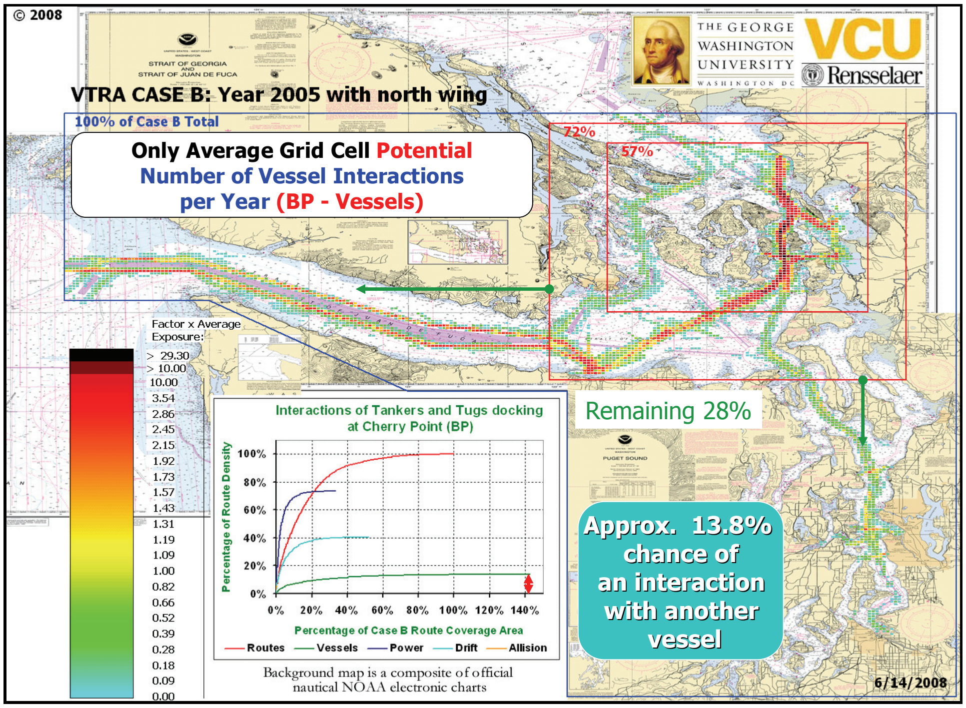


Figure D-32. Vessel interaction counts of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

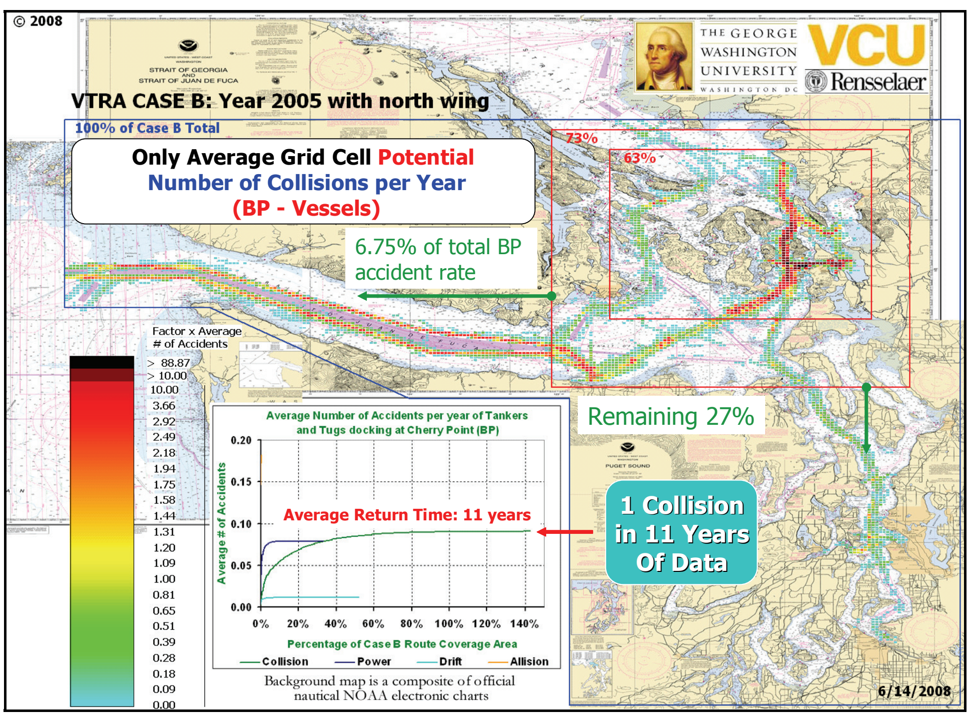


Figure D-33. Annual collision frequencies of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

The calibration process of the grounding model (D-15) is considerably more complicated than the collision model as a result of the additional time-to-shore variable. Whereas in the collision model calibration involves solving a linear equation in a closed form, calibration the grounding model involves solving a non-linear equation using a bisection routine. Each iteration of this routine involves a complete run through all grounding interactions and hence this step is quite computationally intensive. We first solve for the calibration constant t_q in (D-15) by setting the annual frequency of groundings equal to $1/11$ (we observed 1 grounding in 11 years of data) using this bisection method. This results in a value $t_q = 0.834375$. However, after this step the ratio of groundings preceded by human error as compared to mechanical failures turns out to be 2.26 in stead of the desired ratio of 3. To correct this we solve for the remaining calibration constants P_0 by incident type in a similar manner as the collision model calibration. This step results in a calibration value $P_0 = 0.52831297$ for the incident types "Human Error" and "Nearby Vessel Failure" and a calibration value $P_0 = 0.405335373$ given the incident types "Steering Failure", "Propulsion Failure" and "Navigational Aid Failure".

Figures D-34 and D-35 summarize the result of the calibration step for drift groundings. Figures D-36 and D-37 summarize the result of the calibration step for powered groundings. 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 $\approx$ 85 years)
Powered Grounding: ≈ 0.079 (average return time of $\approx$ 13 years)

This coincides with a ratio of 6.8 of powered groundings to drift groundings. Hence, our model evaluates a much higher frequency of powered groundings as compared to drift groundings. This is explained primarily by the ratio of combined incident rates of human error, navigation aid failure and nearby vessel failure to combined incident rates of propulsion failure and steering failure, which is about 4.7 to 1. This takes into account that when a CHPT vessel is underway it has approximately a 13% chance of interacting with another vessel. The remain difference between the ratio of 4.7 to 1 compared to 6.8 to 1 is primarily explained by the time to shore variable in the grounding model. Indeed, given a steering failure or a propulsion failure the time to shore on average is higher than when the vessel remains under power given a human error, navigational aid failure or a near by vessel failure. Thus, the time to shore variable for the drift grounding is higher on average than for

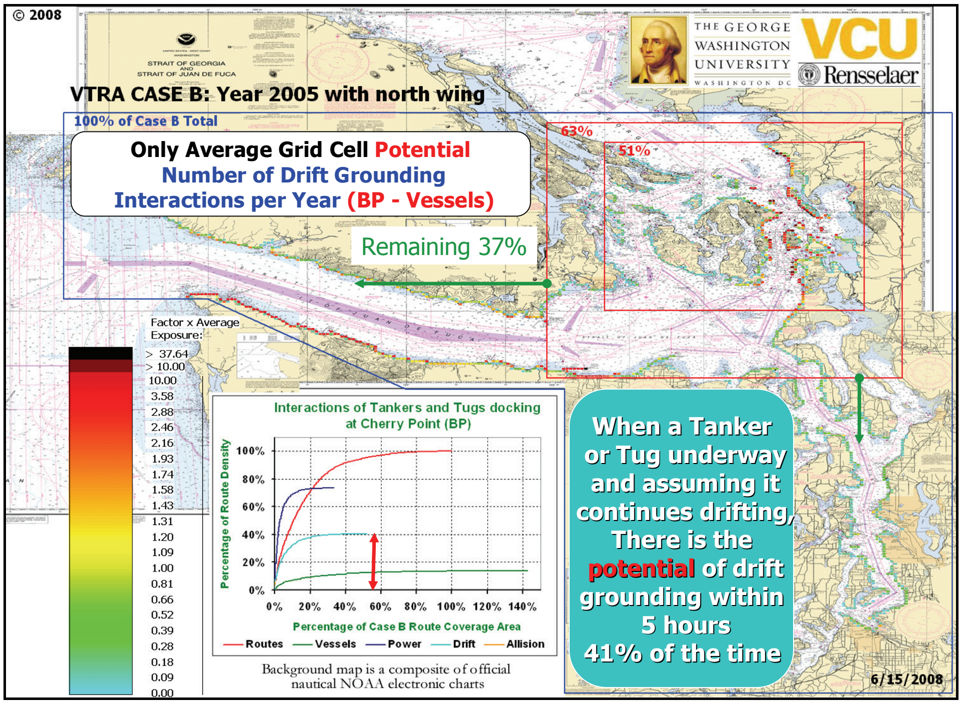


Figure D-34. Drift interaction counts of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

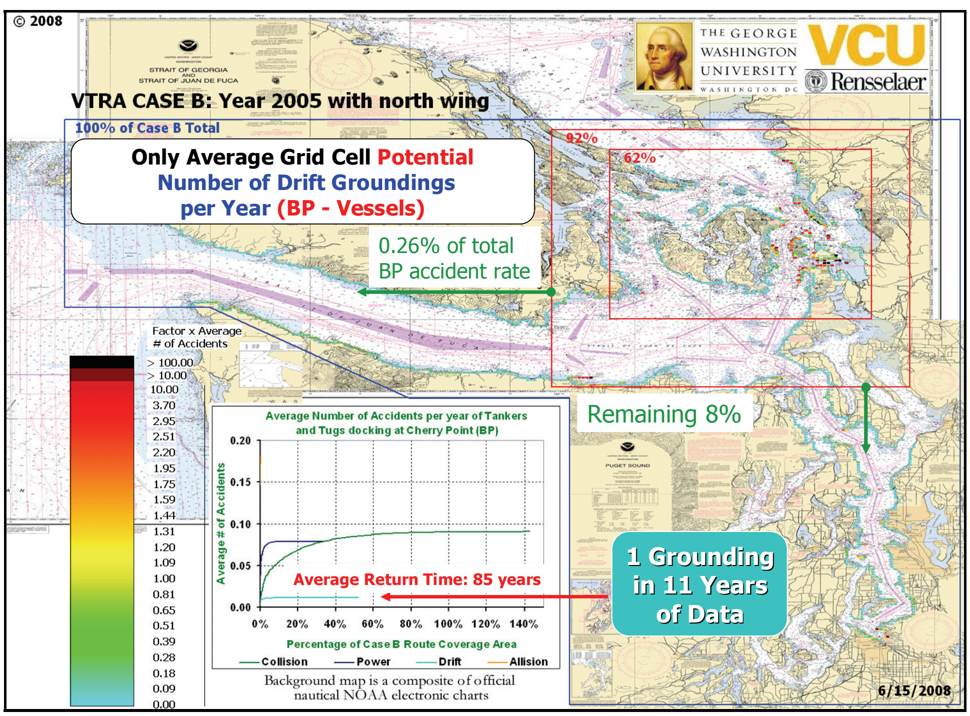


Figure D-35. Annual drift grounding frequency of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

powered groundings resulting on average in a lower drift grounding accident rate per drift interaction than a powered grounding accident rate per power interaction. Observe from Figure D-34 that within our model about 41% of the time there is the potential that a CHPT vessel will run aground within a five hour time frame while adrift, whereas this percentage is 73% when the CHPT vessel is under power (see Figure D-36).

A further effect of the time to shore variable can be observed by comparing Figures D-34 and D-35. Note that while we observe 37% of the drift interactions outside the largest red square, we only observe an 8% of the overall drift grounding accident frequency outside this red square. This follows from larger time-to-shore drifting times overall in the areas outside this red square (especially in the West Strait of Juan de Fuca) compared the area within this red square (especially, the Guemes Channel and Rosario Strait areas and entrances). Indeed, of the total CHPT vessel annual accident frequency of $4/11 \approx 4/11$ (combining collisions, groundings and allisions), we evaluate that only 0.26% is represented on the average by drift groundings outside the largest red square!

The powered grounding analysis displays a similar behavior (see Figures D-36 and D-37). Note that while we observe 33% of the power interactions outside the largest red square, we only observe a 7% of the overall powered grounding accident frequency outside this red square. This too follows from larger time-to-shore times overall in the areas outside this red square (especially in the West Strait of Juan de Fuca) compared the area within this red square (especially, the Guemes Channel and Rosario Strait areas and entrances). Here, of the total CHPT Vessel annual accident frequency of $4/11 \approx 4/11$ (combining collisions, groundings and allisions), we evaluate that only 1.53% is represented on the average by powered groundings outside the largest red square.

Summarizing, the 92% percentage of annual frequency of drift groundings within the largest red square in Figure D-35 and the 93% of annual frequency of powered groundings in this red square in Figure D-37, demonstrates that comparatively within the VTRA study area the grounding risk is confined to this red square (although the remaining 8% and 7% outside should not be considered negligible).

The calibration process for allision is the same as that of groundings. The primary difference between these two accident probability models is the interaction counting as

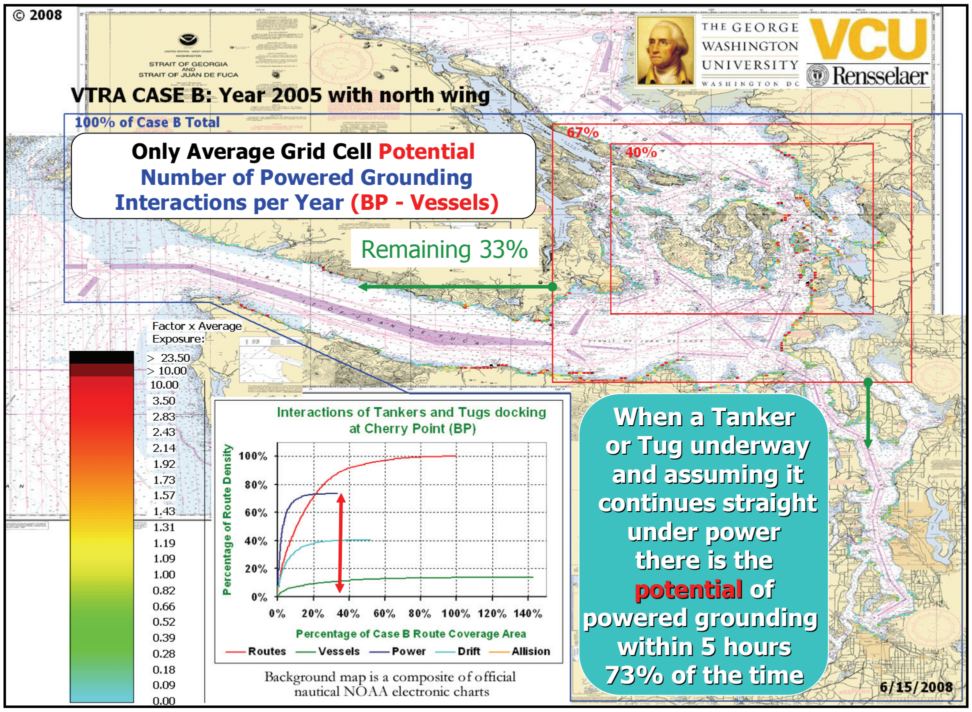


Figure D-36. Power interaction counts of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

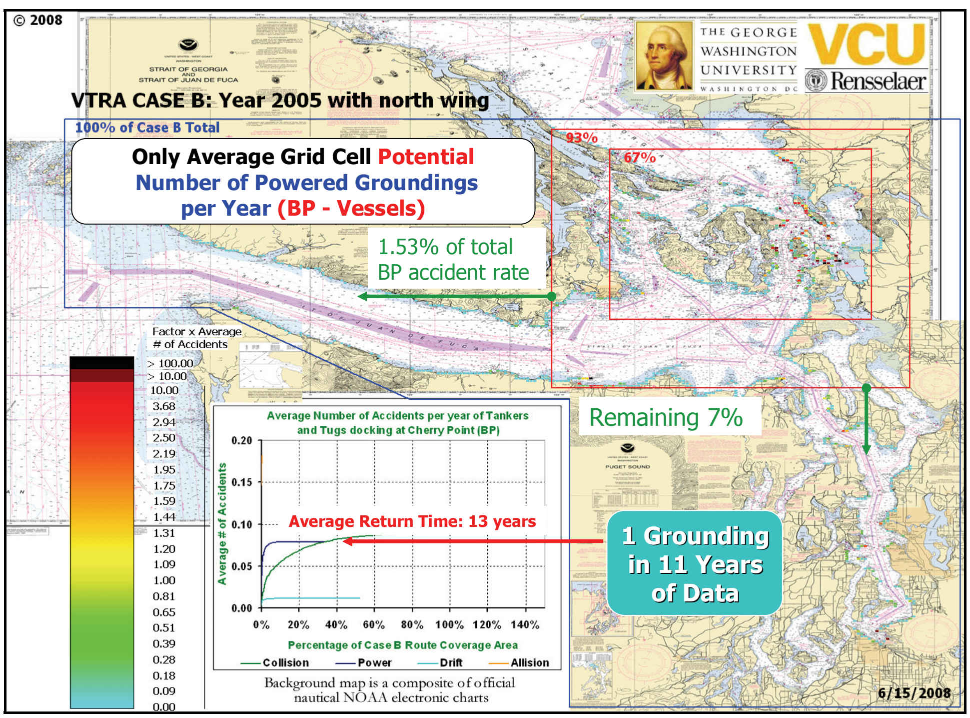


Figure D-37. Powered grounding frequency of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

explained in the previous section. When a vessel is within one mile of its intended dock, the projected shore interactions of a straight line path are designated as allision interactions instead of power interactions. Indeed, within one mile of the intended dock, docking procedures of the tankers and tugs will have commenced, speeds are lowered, escort vessels are in place and from that point on the vessel intentionally tries to get close to the shore with a specific heading towards the shore. After calibrating to 2 allision accidents per 11 years, we arrived at a value of $t_q = 0.384277$, a calibration value $P_0 = 1.039155$ for the incident types "Human Error" and "Nearby Vessel Failure" and a calibration value $P_0 = 0.894719$ given the incident types "Steering Failure", "Propulsion Failure" and "Navigational Aid Failure". The interaction count geographic profile for allisions is presented in Figure D-38 and the allision accident frequency profile is presented in Figure D-39.

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. Next, we can compare the aggregate analysis results of VTRA Case C to those of VTRA Case B and draw overall conclusions regarding the aggregate effect of potentially removing the North wing in our model.

The geographic profiles allow us to further zoom-in on these aggregate effects by compare those of VTRA Case B (see Figures D-32 to D-39) to those of VTRA Case 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 from one case to another comes from. Visual comparison of these geographic profiles allows one to draw 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

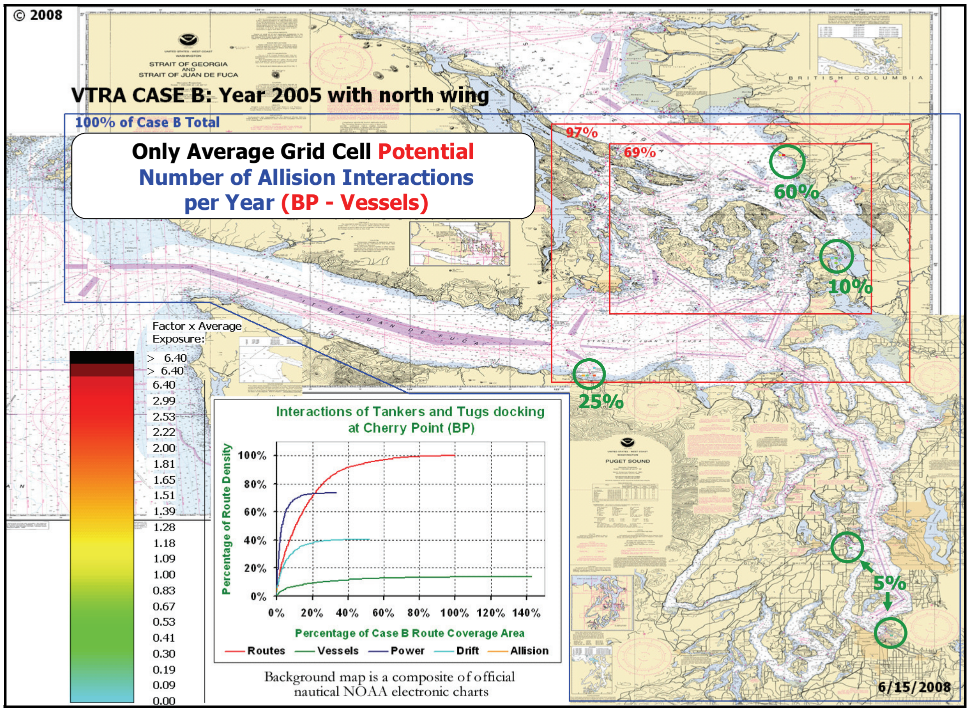


Figure D-38. Allision interaction counts of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

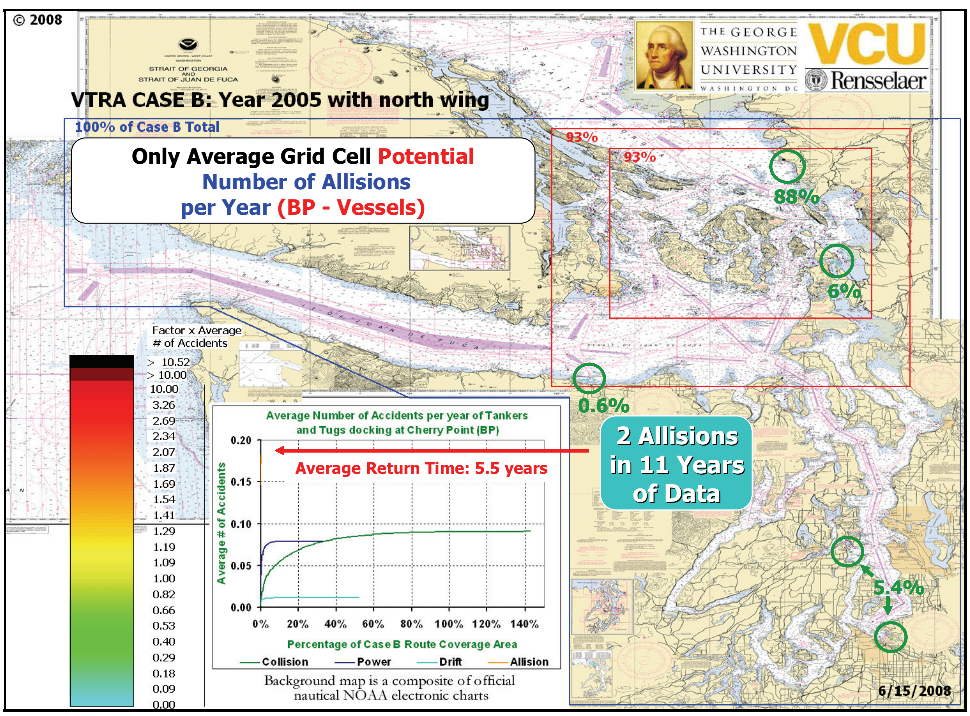


Figure D-39. Allision frequency of Cherry Point Tankers, ATB's and ITB's in the calibration case: VTRA CASE B.

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

R. Bradley (1953). "Some statistical methods in taste testing and quality evaluation", *Biometrics*, Vol. 9 (1): pp. 22-38.

R.M. Cooke (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science*, Oxford University Press, Oxford, U.K.

J.R.W. Merrick, J.R. van Dorp, J.P. Blackford, G.L. Shaw, T.A. Mazzuchi and J.R. Harrald (2003). "A Traffic Density Analysis of Proposed Ferry Service Expansion in San Francisco Bay Using a Maritime Simulation Model", *Reliability Engineering and System Safety*, Vol. 81 (2): pp. 119-132.

J.R.W. Merrick, J. R. van Dorp, T. Mazzuchi, J. Harrald, J. Spahn and M. Grabowski (2002). "The Prince William Sound Risk Assessment". *Interfaces*, Vol. 32 (6): pp.25-40.

National Oceanic and Athmospheric Administration (1997). *Ship Drift Analysis for the Northwest Peninsula and the Stait of Juan de FUca.* HAZMAT Report 97-3.

P. Szwed, J. R. van Dorp, J.R.W.Merrick, T.A. Mazzuchi and A. Singh (2006). "A Bayesian Paired Comparison Approach for Relative Accident Probability Assessment with Covariate Information", *European Journal of Operations Research*, Vol. 169 (1): pp. 157-177.

L.L. Thurstone, (1927a). "A law of comparative judgment", *Psychology Review*, Vol. 34: pp. 273-286.

L.L. Thurstone, (1927b). "Psychophysical Analysis", *American Journal of Psychology*, Vol. 38: pp. 368-389

J.R. van Dorp, J.R.W. Merrick , J.R. Harrald, T.A. Mazzuchi, and M. Grabowski (2001). "A Risk Management procedure for the Washington State Ferries", *Journal of Risk Analysis*, Vol. 21 (1): pp. 127-142