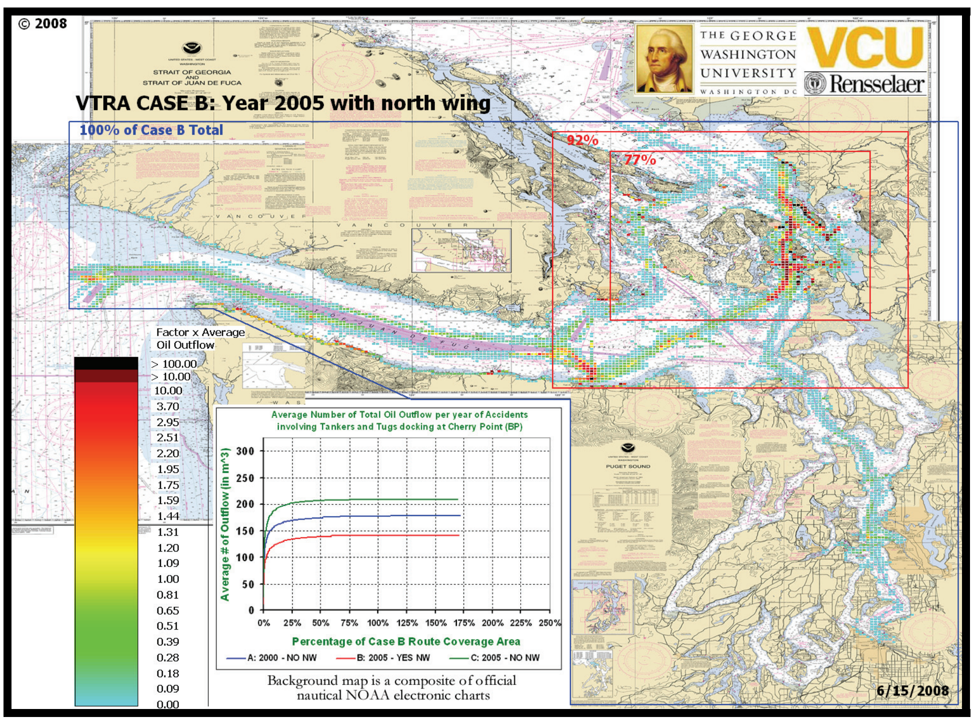
MAIN REPORT



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

PREFACE

This report is submitted by Johan Rene van Dorp, John R. Harrald, Jason R.W. Merrick and Martha Grabowski. Johan Rene van Dorp and John R. Harrald are professors at the George Washington University (GWU), Jason R. W Merrick is a professor at Virginia Commonwealth University (VCU) and Martha Grabowski is a professor at Rensselaer Polytechnic Institute (RPI). The content of the report describes a Vessel Traffic Risk Assessment (VTRA) and the team members above are referred to as the VTRA team.

The VTRA project commenced in June 2006 and spanned over a period of two years. Over the course of this project a comprehensive maritime risk management analysis tool has been developed for the VTRA study area that includes the approaches to and passages through the San Juan Islands, Puget Sound, Haro-Strait/Boundary Pass and the Strait of Juan de Fuca. However, we were tasked to only consider accident risk to vessel docking at the BP Cherry Point dock. The project was funded by BP.

From the outset of this project the support from the United States Coast Guard, Sector Seattle has been unwavering, in particular the support of Captain Stephen Metruck and Jason Tama, who was a lieutenant commander during the first year of this project in Seattle, proved instrumental. As of November 2006, the US Coast Guard introduced the VTRA team to the Puget Sound Harbor Safety committee. Since November 2006 and up to May 2008, we have been able to present our project progress during public meetings every two month held at the Army Corp of Engineers building, 4735 East Marginal Way South in Seattle, WA. During these meeting preliminary results were presented related to our base case analysis of the year 2005. Our last presentation to this community was held in May 2006 during the National Harbor Safety committee held in Seattle at that time.

The Puget Sound Harbor Safety committee, led by Bruce Reed, unselfishly extended their hospitality to allow us to present our progress over the course of this project. They provided us a public platform, missing at the outset of this project, to obtain feedback and access to

the maritime community within the VTRA study for data gathering purposes and expert judgment elicitations.

We are particularly indebted to efforts of Captain Stephen Metruck, LCDR Jason Tama, John Veentjer of the Marine Exchange and Craig Lee from BP Shipping for their efforts in soliciting experts from the maritime community. Experts were invited to and referred to the VTRA team through the United States Coast Guard and the Puget Sound Harbor Safety committee. 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. An elicitation session with ATC tanker captains was scheduled during an ATC conference in February 2007 in Portland, Oregon. In particular, the Puget Sound Pilots, led over the course of this project by Captains Richard McCurdy and Del Mackenzie, were an active participant during the elicitation sessions. None of the experts personally benefited from participating in the expert judgment elicitation. Each expert judgment elicitation session consisted of a morning and afternoon session. They donated their time for the enhancement of the safety levels in their maritime domain and they should be commended for it.

The approach for this VTRA risk assessment is builds on the methodology and the dynamic risk simulations developed for tanker operations in Prince William Sound, Alaska (1995-96), estimation of passenger risk for the Washington State Ferries (WSF) Risk Assessment (1998-1999) and the dynamic exposure simulation methodology for the San Francisco Bay Exposure Assessment (2002) also with a passenger safety focus. This methodology is described in a number of journal papers that have been reviewed by our academic peers:

- 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 off-prints of these papers are attached to this report as sub-appendices (not to be confused with the technical appendices below).

In this project, the VTRA team enhanced and improved their methodology described in the journal papers above in a variety of ways. Some of these improvements are: (1) the maritime system in the VTRA study area was modeled at unsurpassed levels of detail using both AIS and radar data to develop detailed traffic patterns of VTS reporting traffic, (2) small vessel event methodology now includes routes for sailing regattas, whale watching movements and routes from and to both tribal and commercial fishing grounds (3) the use of geographic profiles to display accident frequency and oil outflow across a geographic area using a color legend (this methodology was first used to display exposure in the San Francisco Bay Exposure Assessment), (4) enhanced collision and grounding model that take into account vessel speeds and shore-line interactions and (5) an oil outflow model that builds on those discussed in Special Report 259, *Environmental Performance of Tanker Designs in Collision and Grounding*, published by the National Research Council in 2001 with the ability to model vessel fuel losses in addition to crude oil or refined products cargo losses.

This report contains an executive summary, a main body containing several chapters and a conclusion section. Even though the report was developed to be predominantly self contained, it may refer at times to the following technical appendices A through G that describe the VTRA effort in more technical detail:

- Technical Appendix A: Database Construction and Analysis
- Technical Appendix B: System Description
- Technical Appendix C: Simulation Construction
- Technical Appendix D: Expert Judgment Elicitation
- Technical Appendix E: Oil Outflow Model
- Technical Appendix F: Future Scenarios
- Technical Appendix G: Geographic Exposure, Accident and Oil Outflow Profiles

Jason Merrick took the lead in the production and writing of the main report and Technical Appendix F, Martha Grabowski produced Technical Appendix A, Jack Harrald took the lead in developing Technical Appendix B, Jason Merrick and Johan Rene van Dorp co-authored Technical Appendix C, Johan Rene van Dorp wrote and developed Technical Appendices D, G and took the lead in the production and writing of Technical Appendix E. Finally, Johan Rene van Dorp managed the integration of these documents into a final product.

PROJECT TEAM MEMBERS

Various university members over the course of this project have contributed in a variety of ways to this report, the technical appendix and the analysis. The list below provides the affiliation of the respective university members that have contributed over the course of this two-year project and we would like to thank them for their efforts.

The George Washington University, School of Engineering and Applied Science, Engineering Management and Systems Engineering Department, 1776 G Street, NW, Washington D.C., 20052:

- Johan Rene van Dorp, Associate Professor
- John R. Harrald, Professor
- Greg Shaw, Managing Director of Institute of Crisis Disaster and Risk Management
- Adil Caner Sener, Graduate Research Assistant
- Christian Salmon, Graduate Research Assistant
- Giel van de Wiel, Exchange Visitor from the Delft University of Technology, The Netherlands

Virginia Commonwealth University, Statistical Sciences and Operations Research, P.O. Box 843083, 1001 West Main Street, Richmond, VA 23284-3083:

- Jason R. W. Merrick, Associate Professor
- Christina T. Werner, Research Associate

Rensselaer Polytechnic Institute, Department of Decision Sciences and Engineering Systems, CII 5015, 110 8th Street, Troy, New York, 12180-3590:

- Martha Grabowski, Professor
- Zhuyu You, Graduate Research Assistant
- Zhi Zhou, Graduate Research Assistant
- Huawei Song, Graduate Research Assistant
- Michael Steward, Graduate Research Assistant
- Brittany Steward, Graduate Research Assistant

EXECUTIVE SUMMARY

In June 2006, BP contracted with The George Washington University, Rensselaer Polytechnic University, and Virginia Commonwealth University to perform a vessel traffic risk assessment (VTRA) with the intent of incorporating its results into the Section 10 permit EIS for the addition of a north dock to the BP Cherry Point, Washington facility. The purpose of the VTRA study was to analyze the effects on oil spill risk of potential incremental vessel traffic projected to call at the Cherry Point dock through 2025 and to evaluate mitigation measures applicable to BP to address such impacts.

The maritime system in the VTRA study area was modeled at unsurpassed levels of detail, compared to our prior studies, and a comprehensive set of accident and incident data for all vessel types were collected and analyzed. However, the results and conclusions presented apply only to the interactions involving vessels carrying crude oil and petroleum products to and from the BP Cherry Point refinery. Tank vessels that dock at Cherry Point are articulated tug barges (ATB's), integrated tug barges (ITB's) and tankers, henceforth referred to as BP Cherry Point vessels (BPCHPT). BPCHPT vessels comprise a relatively small percentage of the total modeled vessel traffic in the study area; the time spent on the water by BPCHPT vessels accounts for 1.1% of all modeled traffic and 7 % of the modeled deep draft traffic. In contrast, the combined time on the water for all tankers, ATB's and ITB's, accounts for 3% of all traffic and 16% of deep draft traffic.

The specific scope of the study was as follows:

- The study evaluated the routes used by marine vessels to carry crude oil and petroleum products between the Cherry Point Refinery and:
    - the beginning of the Traffic Separation Scheme approximately 8 nautical miles beyond Buoy J offshore of Cape Flattery, and
    - the Puget Sound.
- The study evaluated the incremental risk of (1) an accident (collision, grounding, or other scenario) involving a tank vessel, (2) resulting in a discharge of crude oil or

petroleum products, (3) associated with reasonably foreseeable increases in vessel traffic through calendar year 2025 to and from both wings of the Cherry Point Oil Spill Risk Assessment due to increased vessel traffic calling at Cherry Point Dock Refinery Pier, (4) as compared with the baseline traffic that the pre-North Wing pier could accommodate.

- In evaluating these risks the study modeled all vessel traffic (not just vessels carrying crude oil and petroleum products) and reasonably foreseeable increases and decreases in vessel traffic along the entire pathway followed by vessels between;
    - Cherry Point and the beginning of the Traffic Separation Scheme approximately 8 nautical miles beyond Buoy J, and
    - Cherry Point and the Puget Sound,

  including but not limited to vessels calling in British Columbia, and vessels calling at the Cherry Point Refinery Pier, Conoco-Phillips, Intalco and reasonably foreseeable future marine terminal facilities in the Cherry Point area, including the proposed Gateway facility.
- The study accounted for non-VTS reporting vessels (fishing vessels and recreation traffic) using methods developed in the modeling of traffic in San Francisco Bay as far as data or expert judgment was available to model this traffic in a reasonable manner.
- The study evaluated low, medium and high traffic scenarios.
- The study considered the impact of human and organizational error on the likelihood of accidents and the effectiveness of risk reduction interventions.
- The study did not evaluate vessel traffic risks at locations other than those routes used by vessels traveling to and from Cherry Point.
- The study investigated risks associated with the Haro Strait and Huckleberry-Saddlebag approaches to and from Cherry Point.
- The study evaluated the following potential vessel traffic management protocols that potentially could reduce the risk of an accident and that can be instituted consistent with existing law: (1) use of Rosario Strait and Guemes Channel instead of the Huckleberry-Saddlebag traverse; (2) stationing a year-round prevention and response tug (of the kind currently stationed in Prince William Sound) in Neah Bay,

Washington; (3) a single tug escort requirement for the Western reaches of Juan de Fuca Strait with hand-off between prevention and response tugs stationed in Neah Bay and Port Angeles.

- The study included an impact analysis that described the outcomes of an accident as described by the location and size of oil outflows, but did not examine the fate and effects of an oil spill, a task to be performed by an independent EIS contractor based on the VTRA oil outflow analysis results.

The approach used for the VTRA extended the methodology developed by the VTRA team for the Prince William Sound Risk Assessment, the Washington State Ferries Risk Assessment, and the San Francisco Bay Water Transit Authority vessel traffic analysis. This approach recognizes that an accident is a culmination of a series of cascading events initiated by triggering mechanical failures and/or human errors. The creation of a comprehensive system simulation as the basic VTRA analysis tool ensures that the dynamic nature of system risk is captured. The system simulation provides a detailed representation of vessel traffic routes and transits for all traffic monitored by the US and Canadian Vessel Traffic Systems and an extensive capture of non VTS traffic such as fishing vessels, regattas and whale watching vessels. In addition, the system simulation represents the situational context of these transits by modeling wind conditions, visibility, and currents. The base year for the system simulation is 2005. In order for the simulation to be used as a risk management analysis tool, however, the absolute number of potential accidents and triggering incidents had to be estimated, and a method for calculating the likelihood that an incident, given a situational context, would result in an accident was required. As described in the basic report, these critical tasks were completed through extensive data analysis and the use of expert judgment where data was inadequate.

An analysis of maritime accidents and incidents in Puget Sound from 1995-2005 was completed to ensure that the maritime simulation was calibrated to historically accurate incident and accident frequencies. Accident and incident records for 1995-2005 for all vessel types and for the geographic scope of the project were solicited, and an accident-incident database was constructed. This data base consists of 2,705 events: 1462 accident events, 1159 incidents, 84 unusual events. Tank vessels accounted for 35 accidents, 111 incidents and 25 unusual events; tank barges accounted for 325 accidents, 87 incidents and 9 unusual events. The BP Cherry Point calling fleet accounted for 4 accidents and 59 incidents during the 1995-2005 time period; these events were used to calibrate the simulation for the base case year. 213 events were identified during the reporting period that were due to human error. Of the 213 human error events, 168 (79%) were unintended errors, rather than violations; of the 168 events, significantly more (52%) were due to perceptual errors, compared to skill-based errors (27%) or decision errors (21%). All of the accident and incident analyses, however, were limited by the availability of detailed information to support human and organizational error analysis.

The modeling of potential oil outflow following an accident extended work done previously by the National Research Council (NRC) and the International Maritime Organization (IMO). The oil outflow analysis estimates the probability of penetration, the number of compartments penetrated, and estimated outflow for each interaction scenario. Oil outflows were modeled for persistent oils (crude cargo and heavy fuel) and non-persistent oils (refined petroleum cargo products and diesel fuel) from vessels in the BP Cherry Point calling fleet and from other interacting vessels involved in a potential collision with the BP Cherry Point vessels.

The most likely base case accidents involving BP Cherry Point vessels were allisions followed by collisions and powered groundings. However, the average oil outflow potential was greatest from powered groundings, followed by collisions. The total potential average oil outflow from BP Cherry Point and interacting vessels in the 2005 base case was 141 cubic meters. In the base case simulation, BP Cherry Point vessels were the source of 97.5% of these oil outflows, interacting vessels accounted for 2.5%. Most (92%) of these oil outflows in the base case are concentrated in the area consisting of the approaches to and passages through the San Juan Islands and Anacortes.

Since shipping is a derived demand, projection of future vessel traffic is inherently uncertain. Actual future tanker and tank barge traffic will be dependent upon energy requirements and distribution choices. Actual future container vessel traffic and bulk cargo traffic and vessel size are dependent upon demand for imports and exports. The vessel traffic described in the base case year (2005) was projected through 2025 using 15 years historical trend data analysis by vessel type. The opening of the Gateway bulk cargo terminal (that would effect this time series projection) and statistical techniques were used to construct upper and lower bounds for future traffic. The resulting high, medium and low forecasts were used as the basis for calculating future accident frequencies and oil outflows.

Detailed accident frequencies and oil outflow volumes were calculated for all locations in the study area for 15 VTRA cases. These 15 VTRA cases describe alternative systems in the years 2000, 2005 2025, considering the presence/absence of BP north wing, the presence or absence of the Gateway terminal, and presence or absence of mitigating measure—saddlebags, extended escort, Neah bay tug. The study's conclusions are based upon the analyses of the VTRA cases.

The following summary provides significant conclusions drawn from the analysis comparison of 2000-2005 VTRA cases, the comparison of 2000-2025 VTRA cases, and the analysis of specific potential risk reduction interventions[1]:

2000-2005 comparison conclusions derived from VTRA analysis results:

- If BP had restricted operations to the south wing in 2005, it could have served 96% of the BPCHPT vessels in 2005 actually served by both wings.
- In a 2005-2005 comparison, the addition of the north wing allowed the BP Cherry Point terminal to serve slightly more calling vessels, while reducing the potential for BPCHPT vessel accidents by 21% and decreasing oil outflows by 38%[2].
- With the north wing in operation in 2005 (but not in 2000), the potential for accidents involving BPCHPT vessels decreased by 10% between 2000 and 2005, and the oil outflow potential decreased by 21% between 2000 and 2005 in spite of the changes in vessel traffic during the same period.

[1] The main report and its technical appendices provide a more detailed explanation of these results.

[2] For consistency percentages are evaluated here as percentages of 2000 levels.

- With only the south wing in operation in both 2000 and 2005, the potential for accidents involving BPCHPT vessels would have increased by 12% between 2000 and 2005 and the potential outflows would have increased by 18% between 2000 and 2005.

2000-2025 analysis conclusions derived from VTRA analysis results:

- At each of the low, medium, and high traffic scenarios for 2025, having the north wing leads to lower average accident potential and oil outflow potential for BPCHPT vessels than not having it.
- Assuming the north wing being operational in a 2025 analysis with medium traffic increases, results in a total annual average oil outflow of 174.4 cubic meters, which is quite similar to the 177.7 cubic meters of the previous 2000 analysis when the dock was not operational (but a reduction of 1.8%).
- Assuming the north wing being operational in a 2025 analysis with high traffic increases, results in a total annual average oil outflow of 229.9 cubic meters, compared to the 177.7 cubic meters of the same 2000 analysis when the dock was not present (an increase of 29.4%).
- Hence, with additional traffic increases it remains possible that even with the addition of the north wing dock, oil transportation risk rises above a level previously experienced in 2000 when the north wing dock was not operational.

Risk intervention conclusions derived from VTRA analysis results:

- At the 2005 traffic levels, and not allowing the use of the Saddlebags route from BP Cherry Point to Anacortes in our maritime risk simulation model, leads to no appreciable change in either average accident potential or average oil outflow potential. In the high traffic scenario for 2025, not allowing the use of the Saddlebags route from BP Cherry Point to Anacortes in our maritime risk simulation leads to a 2% increase in average accident potential and a 0.1% increase in average oil outflow.
- At the 2005 traffic levels, extending the escorting of BP tankers and ITBs up to Buoy J in our maritime risk simulation model, leads to a decrease in both drift groundings

and collisions in the extended escorting area. The overall effect is a 1.5% decrease in total average accident potential and a 3% decrease in total average oil outflow potential. In the high traffic scenario for 2025, these decreases are 1% and 1.5%, respectively.

- A restricted analysis of the risk reduction potential of the Neah Bay Tug, considering only BP tankers (about 1.1% of the total modeled traffic and about 7% of the total modeled deep draft traffic) within the VTRA study area (i.e. up to 8 miles of Buoy J where traffic separation commences and, more importantly, including the area consisting of the approaches to and passages through the San Juan Islands and Anacortes typically beyond the Neah Bay tug's operating range) our maritime risk simulation model evaluated that the Neah Bay tug has no appreciable effect on total VTRA study area average accident potential and reduces its total average oil outflow potential by 0.1%.
- In the restricted analysis performed, and assuming the Neah Bay tug has the capability to save any disabled[3] BPCHPT vessel that it could get to in time, regardless of the situational context, it was shown that the Neah Bay tug could reduce total average VTRA study area accident potential by 0.03% and total average VTRA study area oil outflow potential by 0.75%.

Quantitative results in our study are presented as average point estimates commonly used for the evaluation of alternatives in a decision analysis context. These are derived from uncertain quantities as described in each step of the analysis as described in this report and its technical appendices. As with any risk assessment model, our model too represents an abstraction of reality and its results must be interpreted with care and with awareness of scoping, data limitations and modeling assumptions. In particular, the forecasts of maritime traffic, accident frequencies, and oil outflows in 2025 must be treated with care.

One primary limitation of the VTRA study is that, due to scoping constraints, the results reflect only on a small percentage of the vessel traffic described in the maritime simulation.

[3] Our definition of a disabled BPCHPT vessel here is one that experienced either a steering or propulsion failure.

If risk interventions have an appreciable effect beyond the BPCHPT vessels analyzed in this study, they should also be tested against this larger class of vessels to determine their effects on system wide accident frequencies and oil outflows. For example, a risk intervention that reduces accident frequency and or oil outflow of BP Cherry Point vessels, but results in a larger potential increase of accident frequency and/or oil outflows from the other traffic should not be implemented. Conversely, risk mitigation measures that have little or no impact on the BP Cherry Point vessels accident frequency or oil outflow may in fact significantly reduce risk to other vessels.

As such, a full evaluation of the risk reduction potential of the Neah Bay tug was not within the scope of the VTRA, as the analysis was restricted to BPCHPT vessels in the VTRA geographic scope. A full evaluation of the risk reduction potential of the Neah Bay tug requires (1) inclusion of all non-BP vessel traffic within the VTRA study area in its effectiveness analysis and (2) inclusion of all vessel traffic beyond the boundaries of our VTRA study area (i.e. beyond the beginning of the Traffic Separation Scheme approximately 8 nautical miles beyond Buoy J offshore of Cape Flattery), but both limited to the tug's operating range. Neither was part of the scope of the VTRA study.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF FIGURES (continued)

TABLE OF TABLES

1. Introduction

The purpose of the VTRA study is to examine the changes in vessel traffic risk potentially resulting from the addition of the north-wing of the dock at BP's Cherry Point Refinery (see Figure 1). This risk is evaluated in terms of the risk of accidents involving tankers, articulated tug barges (ATBs) and integrated tug barges (ITBs) calling at BP Cherry Point and in terms of the potential for oil outflow from such accidents. The accidents included in the analysis are collisions with other vessels, groundings preceded by propulsion or steering failures (so the tanker drifts aground), groundings preceded by navigational failures or human errors (so the tanker goes aground under power), and allisions (collisions with the dock or other fixed objects). We will evaluate the changes in risk during our base case year (2005) were the dock to be used or not used. We will also evaluate the changes in risk since prior to the dock being constructed (specifically in the year 2000) and the changes in risk that could result because of future changes in traffic levels (evaluated in 2025). This will give an elaborate evaluation of the effect of the dock thus far and in to the future.

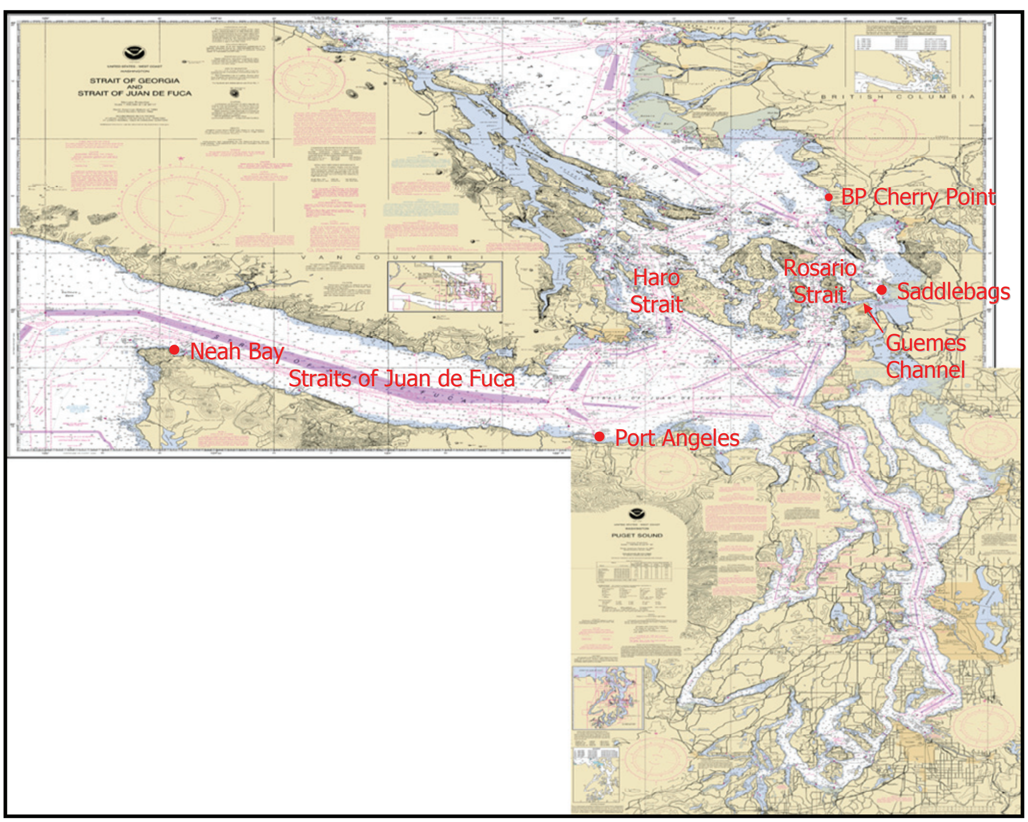


Figure 1. A chart of the area discussed in the study.

Three other changes to the risk of oil spills will be evaluated. Until 2005, tankers wanting to transit between BP Cherry Point or the Ferndale refinery and Anacortes would travel through an area known as the Saddlebags (Figure 1). We will evaluate the change in risk if the tanker were to not use Saddlebags, but instead to travel through Rosario Strait and then Guemes Channel. The other two changes involve alternatives to the current escorting system. Currently, tankers over 40,000 dead weight tons and carrying either crude oil or petroleum products must be escorted by a specialized tug from a point just after they pass by Port Angeles to the refinery or the anchorage they plan to use. They must then be escorted on any transits between anchorages and refineries until just before they reach Port Angeles on their way out through the Straits of Juan de Fuca. In addition, there is also an escort tug on 24 hour standby at Neah Bay near the entrance to the Straits of Juan de Fuca to assist any vessel within its range to assist. Two changes to this set up will be evaluated. Firstly, extending the escorts passed Port Angeles up to the end of the Straits of Juan de Fuca. Secondly, we will test the effect of removing the Neah Bay Tug to allow an evaluation of its effectiveness.

Each of the changes to be evaluated cause a change in the traffic patterns and these cause a change in the level of exposure to risk; the changes are dynamic, thus they must be evaluated using a dynamic simulation of the vessel traffic in the region. This simulation model is then integrated with models for the potential of incidents, such as propulsion failures, steering failures, navigational failures, and human errors, and with models for the potential for collisions, groundings, and allisions resulting from these incidents. A final layer of modeling takes the accident scenarios and assesses the potential for oil outflow. The form of results in this report include the aggregated potential for accidents, aggregated potential for oil outflow, and detailed maps showing the spread of each of these risk measures across the area, called geographic profiles.

In this report, we first provide the scope of the study that defines what is analyzed and what is not. We then provide a description of the system, the vessels that transit in it, the rules they follow, and the environmental factors that they encounter. The overall structure of the model used is then discussed and each of the individual pieces described, including the data that was used and the organizations that provided experts. The results are then provided and the effect of changes to the system explained. We end with a summary of our findings.

2. Scope of the Study

The scope of the study is defined by the following items:

- The study will evaluate the routes used by marine vessels to carry crude oil and petroleum products between the Cherry Point Refinery and:
  - the beginning of the Traffic Separation Scheme approximately 8 nautical miles beyond Buoy J offshore of Cape Flattery, and
  - the Puget Sound.
- The study will evaluate the incremental risk of (1) an accident (collision, grounding or other scenario) involving a tank vessel, (2) resulting in a discharge of crude oil or petroleum products, (3) associated with reasonably foreseeable increases in vessel traffic through calendar year 2025 to and from both wings of the Cherry Point Oil Spill Risk Assessment due to increased vessel traffic calling at Cherry Point Dock Refinery Pier, (4) as compared with the baseline traffic that the pre-North Wing pier could accommodate.
- In evaluating these risks the study will model all vessel traffic (not just vessels carrying crude oil and petroleum products) and reasonably foreseeable increases and decreases in vessel traffic along the entire pathway followed by vessels between;
  - Cherry Point and the beginning of the Traffic Separation Scheme approximately 8 nautical miles beyond Buoy J, and
  - Cherry Point and the Puget Sound,

  including but not limited to vessels calling in British Columbia, and vessels calling at the Cherry Point Refinery Pier, Conoco-Phillips, Intalco and reasonably foreseeable future marine terminal facilities in the Cherry Point area, including the proposed Gateway facility.
- The study will account for non-VTS reporting vessels (fishing vessels and recreation traffic) using methods developed in the modeling of traffic in San Francisco Bay as far as data or expert judgment is available to model this traffic in a reasonable manner.

- The study will evaluate low, medium and high traffic scenarios.
- The study will consider the impact of human and organizational error on the likelihood of accidents and the effectiveness of risk reduction interventions.
- The study will not evaluate vessel traffic risks at locations other than those routes used by vessels traveling to and from Cherry Point.
- The study will cover risks associated with the Haro Strait and Huckleberry-Saddlebag approaches to and from Cherry Point.
- The study will include identification and evaluation of potential vessel traffic management protocols that would reduce the risk of an accident and that can be instituted consistent with existing law. At a minimum, the vessel traffic management protocols studied will include: (1) use of Rosario Strait and Guemes Channel instead of the Huckleberry-Saddlebag traverse; (2) stationing a year-round prevention and response tug (of the kind currently stationed in Prince William Sound) in Neah Bay, Washington; (3) a single tug escort requirement for the Western reaches of Juan de Fuca Strait with hand-off between prevention and response tugs stationed in Neah Bay and Port Angeles; and (4) any additional vessel traffic management protocols or other mitigation measures selected for analysis during the scoping stage of the EIS.
- The study will include an impact analysis that will describe the outcomes of an accident as described by the location and size of oil outflows, but will stop short of examining the fate and effects of an oil spill.
- The study will use, but not be constrained by, the results of prior studies that examined various aspects of maritime risk in Washington State waters. The study will be directed by Jack Harrald and Martha Grabowski.

Figure 2 shows the geographic area included in this scope. The study will only evaluate the risk of accidents involving crude oil or petroleum product carrying vessels that call at BP Cherry Point at some point in their transit in to this geographic area. Thus it will not include collisions between two vessels that do not call at BP Cherry Point and it will not include groundings or allisions of non-BP Cherry Point vessels. However, it will include collisions between BP Cherry Point vessels and any other vessel, thus the model must include as much of the other traffic in the region as it is feasible to model, including small vessels for which transit data is more difficult to obtain.

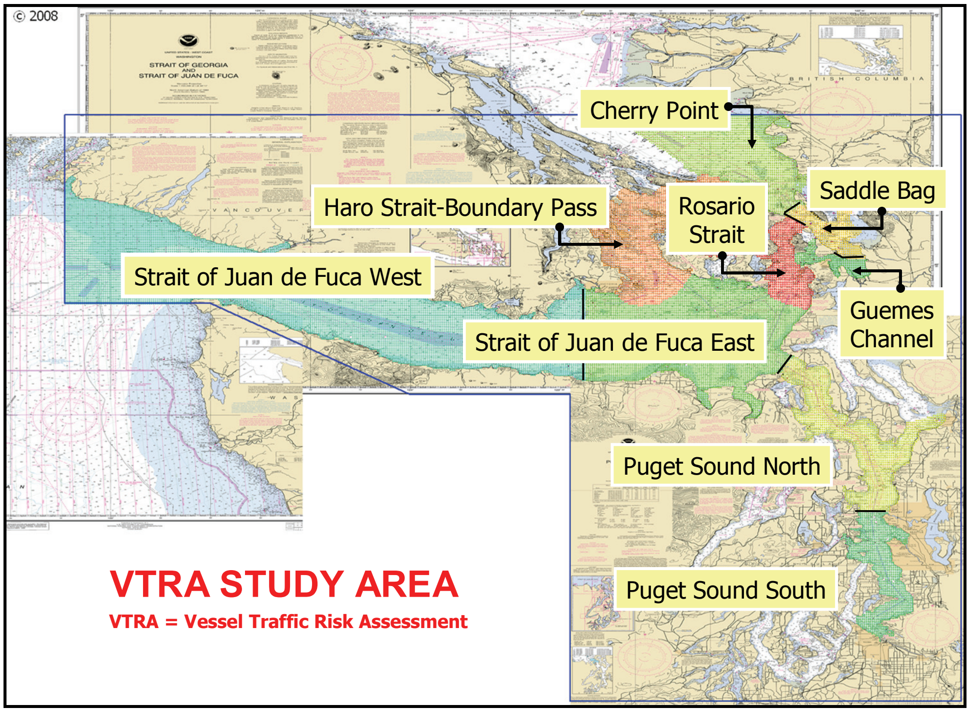


Figure 2. A chart of the geographic scope of the study.

3. Description of the System

3.1. Traffic to BP Cherry Point

Figure 3 shows a satellite map of the area around BP's Cherry Point Refinery. To the far left is Vancouver Island; the islands at the bottom are the San Juan Islands. Canada is at the top of the picture, specifically the city of Vancouver and the James River. To the lower right is Washington State. BP's Cherry Point Refinery is located at Cherry Point, which is near Bellingham.

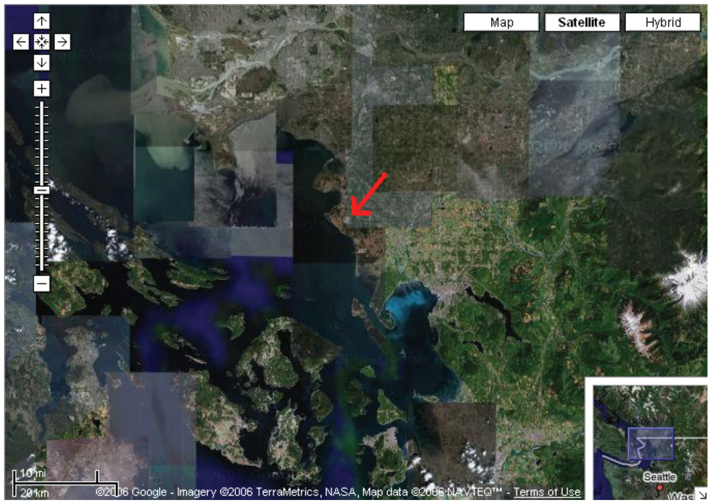


Figure 3. A satellite picture of the area around BP Cherry Point.

Figure 4 shows a satellite image of the dock at BP's Cherry Point Refinery. The dock lower in the image is the south wing which is now used mostly for tankers carrying crude oil being delivered to the refinery, but can be used for petroleum product carrying vessels taking refined products away from the refinery. The south wing was constructed at the same time as the refinery, being finished in September 1971. The dock higher in the image is the north wing that was constructed by July of 2001 and went in to service in September of 2001. The north wing is used exclusively for vessels carrying refined petroleum product away from the refinery.

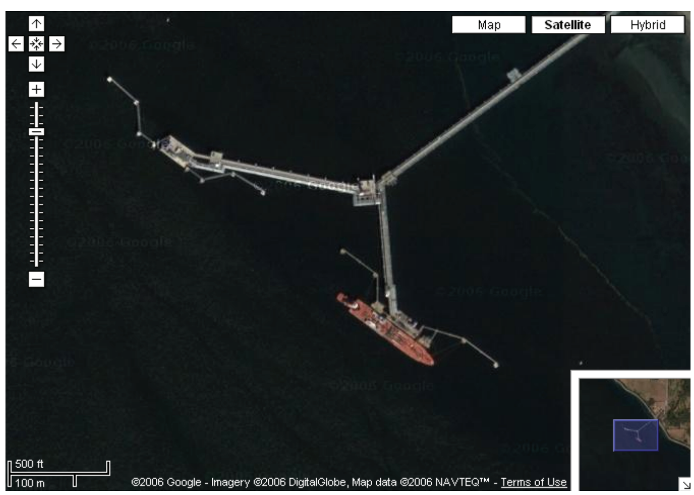


Figure 4. A satellite picture of the north and south wing of the dock at BP Cherry Point.

Figure 5 shows the number of crude oil tankers that called at BP Cherry Point each month from March 1997 to February 2008. The red line shows the point in time that BP merged with ARCO in April 2000. After that point there is a slow increase from an average of 9.2 transits per month to an average of 11.5 transits per month. It is noteworthy that this higher number of transits is actually reached before the north wing went in to operation.

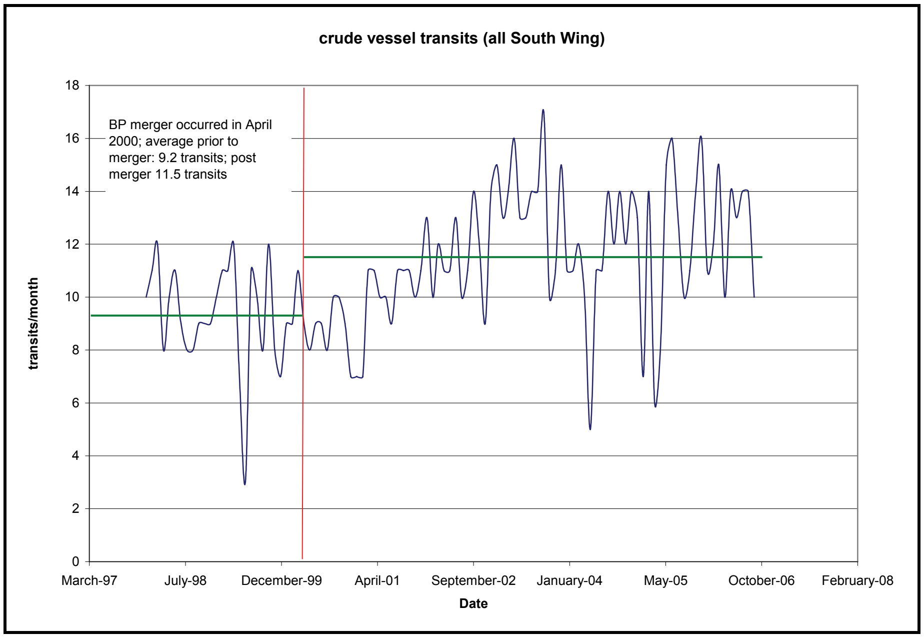


Figure 5. Crude oil tankers calling at BP Cherry Point per month.

Figure 6 shows the number of petroleum product carriers calling at BP Cherry Point from March 1997 to February 2008. Until September 2001, all product carriers had to use the south wing, but after that point they use the north wing if it is available and the south wing if the north wing is not available and no crude tanker is using the south wing. Prior to the use of the north wing, there was an average of 14 product carriers per month. With the north wing in place, an average of 12 product carriers per month used the north wing and 2 per month used the south wing. There appears to be a slight increase in 2006. This will not be reflected in our base case analysis (2005), but such trends will be considered in our analysis of future scenarios.

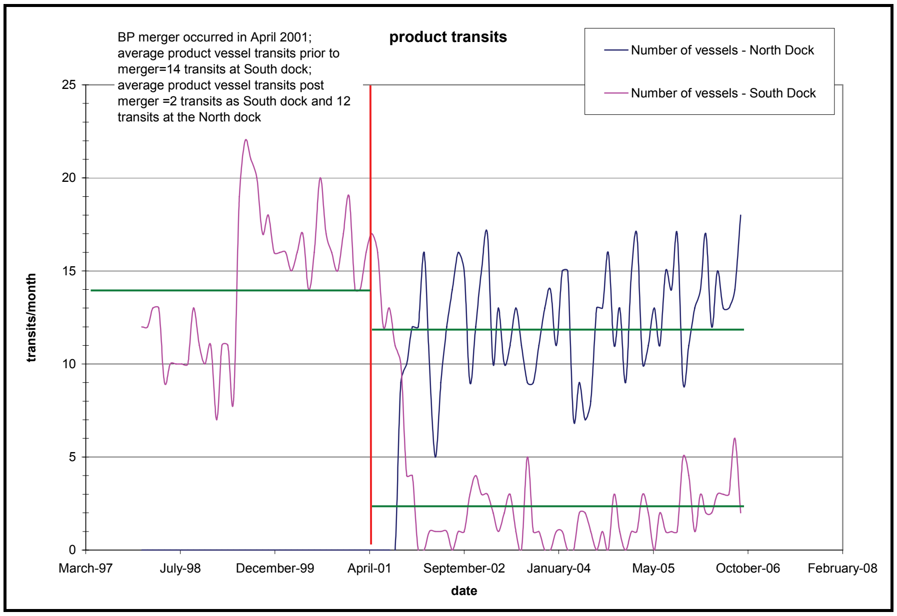


Figure 6. Petroleum product tankers calling at BP Cherry Point per month.

The routes used by tankers, ATBs, and ITBs that call at BP Cherry Point are shown in Figure 7. As can be seen, tankers transit in and out of the Straits of Juan de Fuca, south to the Puget Sound (specifically to Tacoma and Manchester), north to Vancouver, and locally to anchorages at Cherry Point, Vendovi Island, Anacortes, and Port Angeles and the refineries at Ferndale and Anacortes.

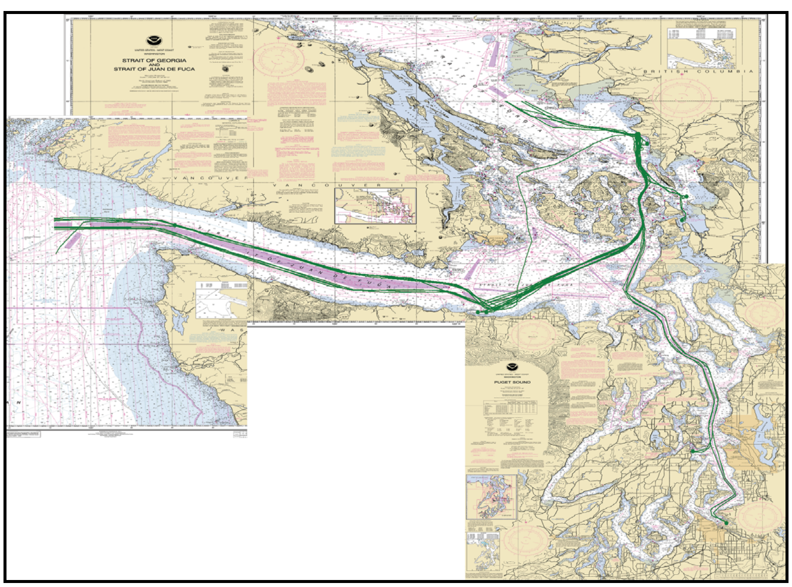


Figure 7. Representative Routes Used by Tankers Calling at BP Cherry Point.

3.2. Deep Draft Traffic

There are many other types of vessels that transit the waterways in this region. Larger vessels must report in to the Coast Guard Vessel Traffic Service (VTS). The specific requirements for a vessel to report in are:

(a) Every power-driven vessel of 40 meters (approximately 131 feet) or more in length, while navigating;

(b) Every commercial towing vessel of 8 meters (approximately 26 feet) or more in length, while navigating;

(c) Every vessel certificated to carry 50 or more passengers for hire, when engaged in trade.

The VTS records the transit and also monitors the movement of vessels on screens in their operating center. The USCG VTS in Seattle receives radar signals from 12 strategically located radar sites throughout the VTSPS area. Radar provides approximately 2,900 square miles of coverage including the Strait of Juan de Fuca, Rosario Strait, Admiralty Inlet, and Puget Sound south to Commencement Bay.

Additionally, close circuit TV provides coverage of various critical waterways. Since 1979, the U.S. Coast Guard has worked cooperatively with the Canadian Coast Guard in managing vessel traffic in adjacent waters through the Cooperative Vessel Traffic Service Puget Sound (CVTS). Two Canadian Vessel Traffic Centers work hand in hand with Puget Sound Vessel Traffic Service. The area west of the Strait of Juan De Fuca is managed by Tofino Vessel Traffic. North of the Strait of Juan De Fuca, through Haro Strait, to Vancouver Harbor, BC is managed by Victoria Vessel Traffic Service. The three Vessel Traffic Centers communicate via a computer link and dedicated telephone lines to advise each other of vessels passing between their respective zones.

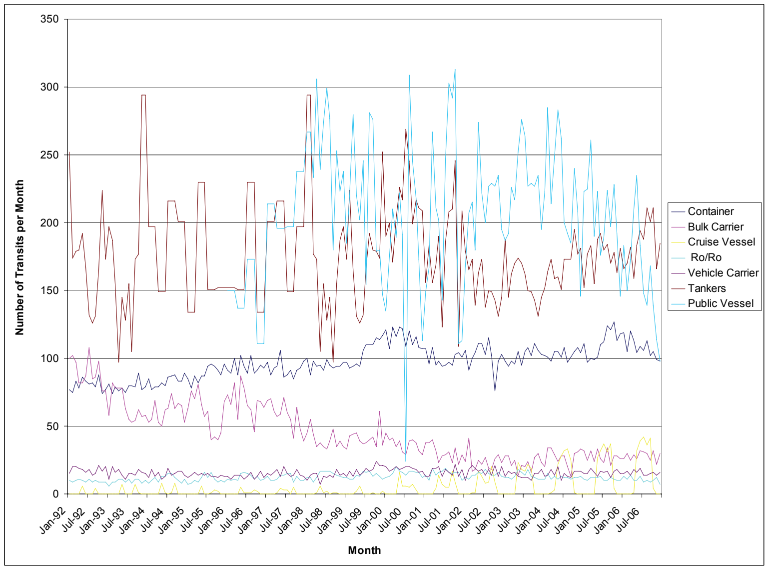


Figure 8. The number of vessel transits per month since 1992.

The number of transits per month for various types of vessels in the region are shown in Figures 8 and 9. Figure 8 shows bulk carriers, containers, cruise vessels, public vessels (navy and coast guard), roll-on/roll-off vessels, tankers, and vehicle carriers. Figure 9 shows tug tow barge transits separately as they are an order of magnitude higher than other vessels. Various trends are seen here, including a decrease in the number of bulk carriers, an increase in the number of container vessels, and a seasonal increase in the number of cruise vessels.

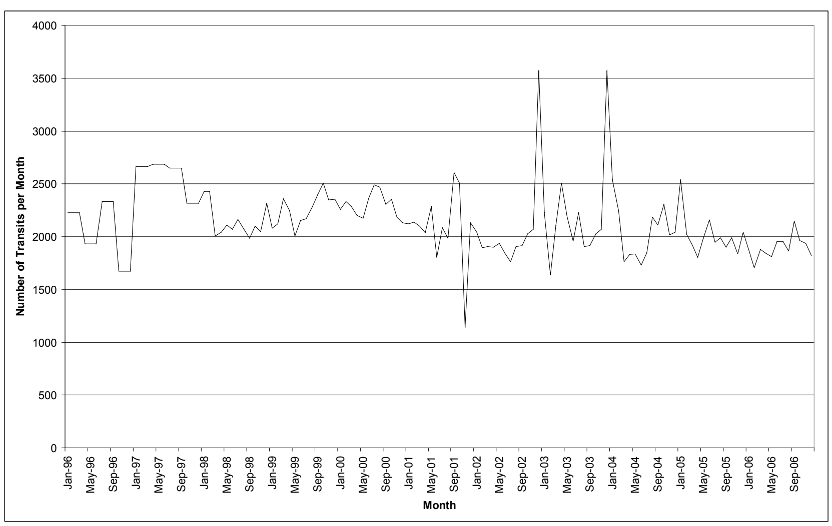


Figure 9. The number of tug transits per month since 1996.

The VTS centers also record the tracks of the vessels that report in. Figure 7, showing the routes of oil tankers, was generated by cleaning radar blips and other recording errors from these tracks for oil tanker transits and choosing representative tracks for each departure point and destination. Figure 10 through 15 show similarly generated routes for the other types of vessels that call in to the VTS.

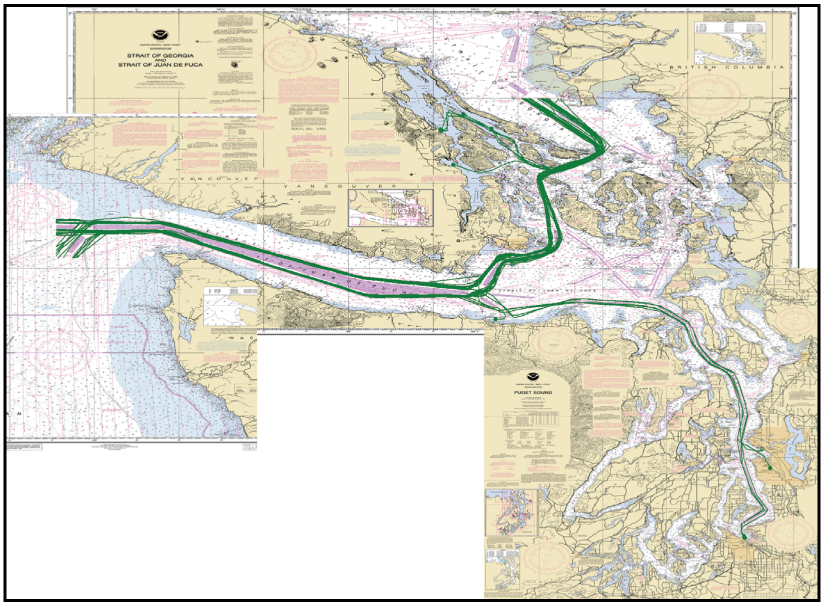


Figure 10. Representative Routes Used by Bulk Carriers.

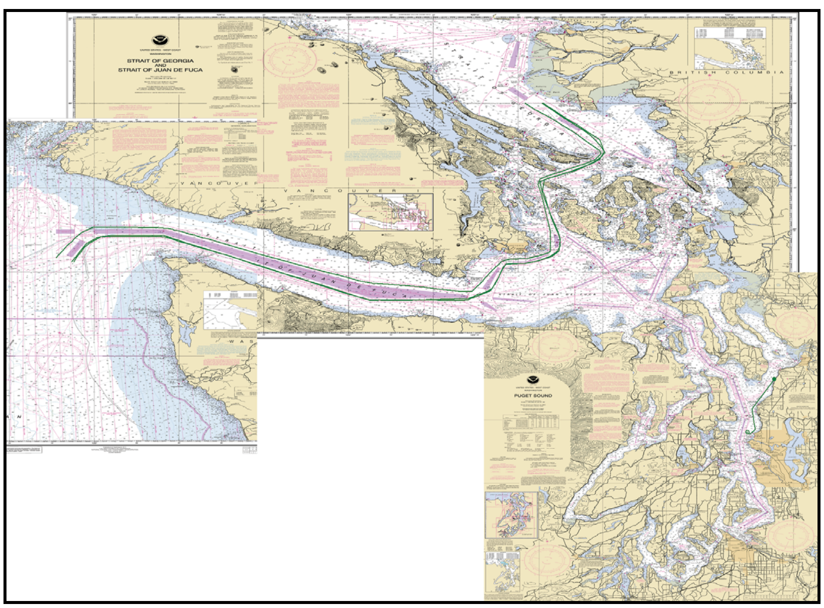


Figure 11. Representative Routes Used by Chemical Carriers.

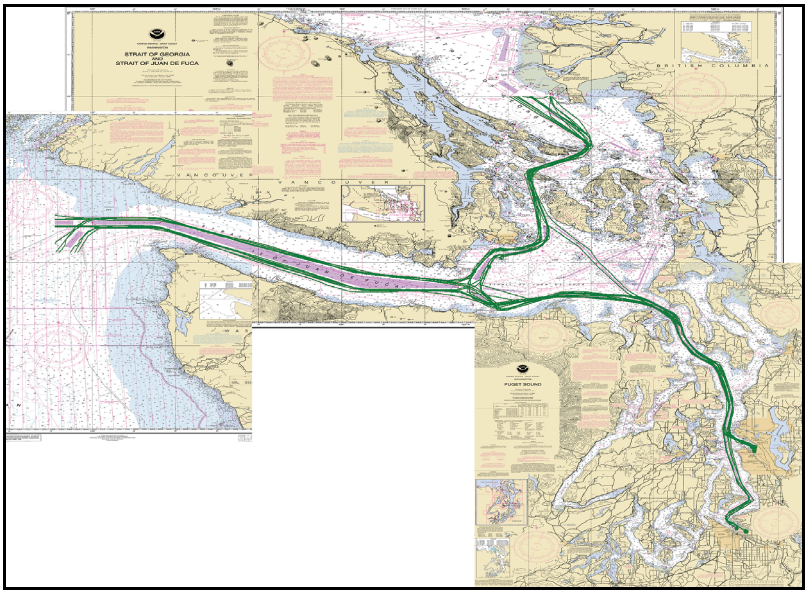


Figure 12. Representative Routes Used by Container Vessels.

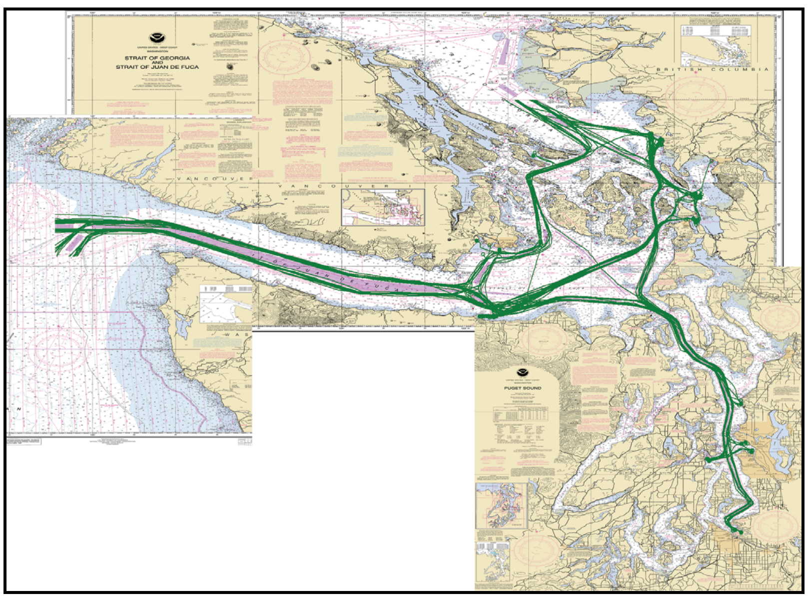


Figure 13. Representative Routes Used by all Oil Tankers.

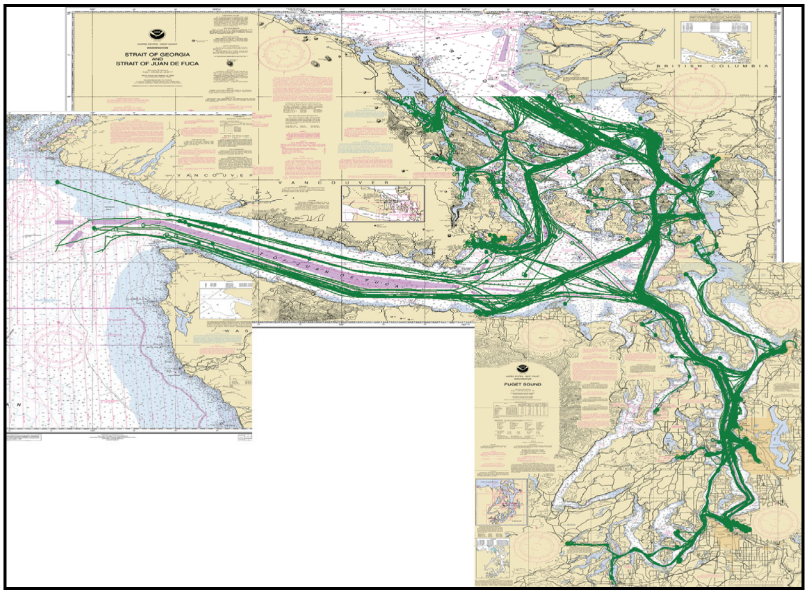


Figure 14. Representative Routes Used by Tug Tow Barges.

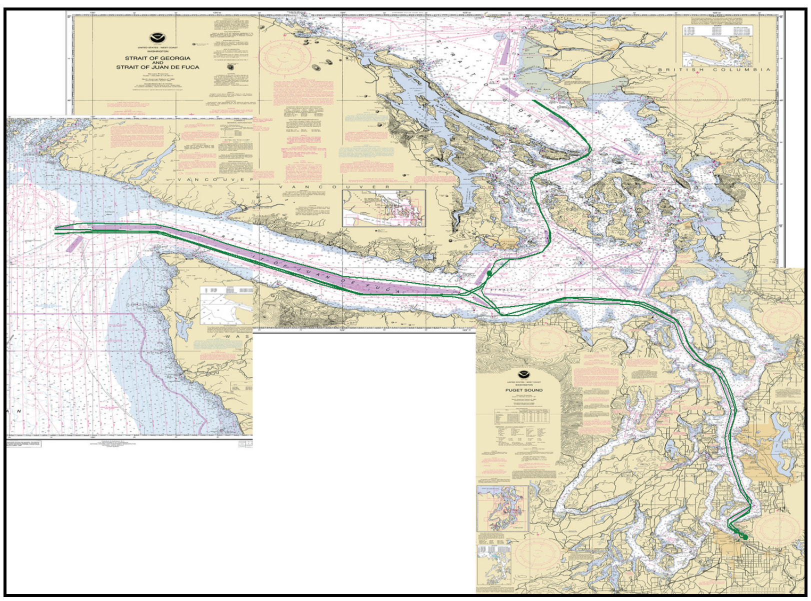


Figure 15. Representative Routes Used by Vehicle Carriers.

3.3. Ferry Traffic

The vessel type with the largest numbers of transits in this area is ferries. Ferries also call in to the VTS. The ferries in this area are operated by the Washington State Ferries (the largest ferry service in the United States), the Victoria Clipper (between Victoria and Seattle), and various smaller Canadian operators. The total number of ferry transits per month since 1996 is shown in Figure 16, varying somewhere around 15000 transits per month. This is by far the highest number of transits per month of any VTS reporting traffic, but it should be realized that most of the routes in Figure 17 for ferries are much shorter than most of the routes for other VTS reporting traffic in Figures 10 through 15. Representative ferry routes are shown in Figure 17.

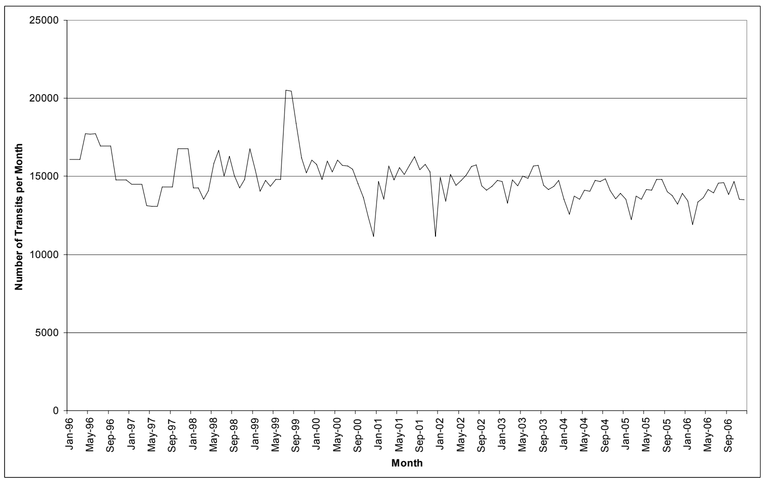


Figure 16. The number of ferry transits per month since 1992.

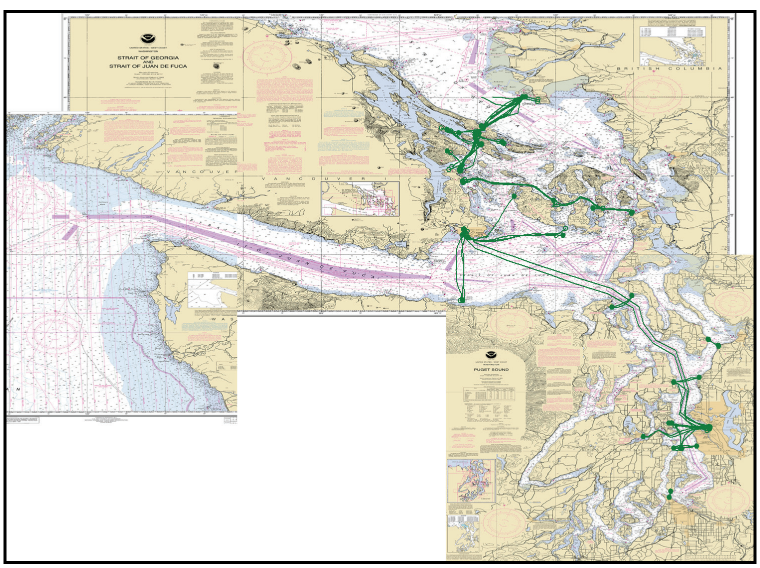


Figure 17. Representative Routes Used by Ferries.

3.4. Small Vessel Traffic

There are many other types of smaller vessels that use these waters. While data on these smaller vessels is harder to obtain, there are several groups that are recorded by various different entities.

Commercial and tribal fishing is regulated and recorded by various organizations. Canadian, US, and tribal fisheries managers provided information about the areas in which fishing occurs, the types of fishing and vessels, and the number of vessels that transit from each fishing port. Larger fishing vessels and fishing factory vessels also must report to the VTS. Figure 18 shows the sum of this information, with fishing areas shown in various colors depending on the type of fishing performed in each area and representative routes for transit to and from these areas and in and out of the region in green.

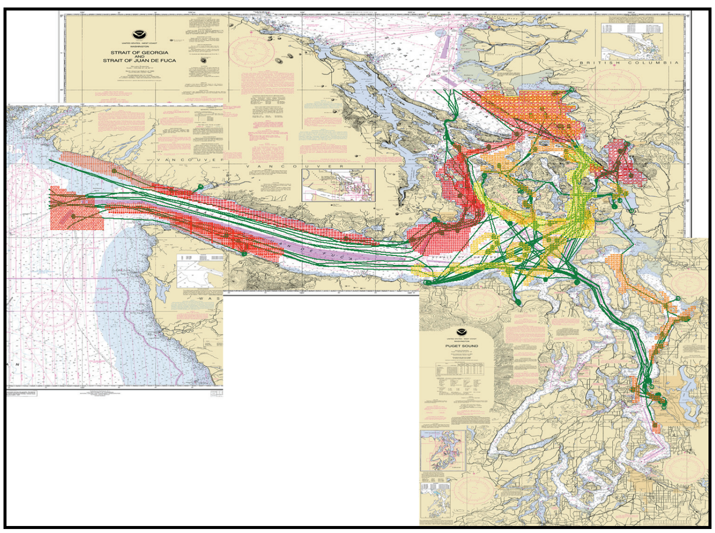


Figure 18. Fishing areas and representative routes used by fishing vessels.

Two other types of small vessel traffic for which data is recorded are regattas and whale watching. Regattas run by various yachting organizations in the area must be registered with the Coast Guard. This includes the time of the event, the route taken, and the expected number of vessels involved. Figure 19 shows the routes of regattas that took place during 2005.

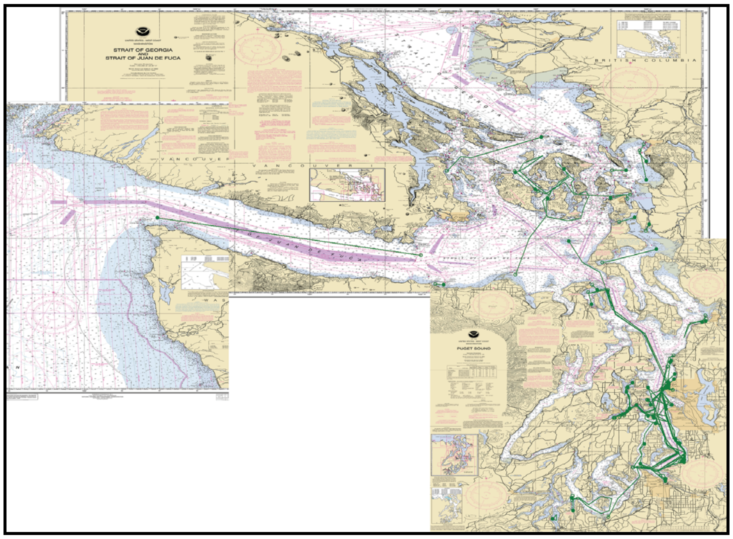


Figure 19. Representative Routes Used by USCG Registered Yacht Regattas.

Whale watching vessels follow the pods of killer whales that live in the region, allowing tourists and researchers to observe the whales. However, there are regulations that restrict these vessels from harassing the whales. Sound Watch is a non-profit organization that records the movements of whales, the number of vessels that are observing them, and any violations of the regulations. Figure 20 shows the movements of the whales and the whale watching vessels that followed them during 2005.

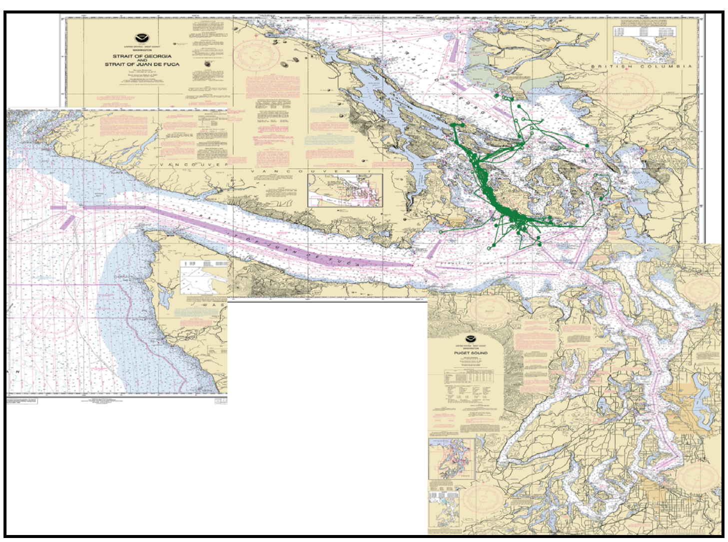


Figure 20. Routes of whale watching movements record by Sound Watch.

3.5. Traffic Rules

Reporting to the VTS is not the only requirement for vessels transiting the region. There are restrictions on where a vessel may transit, called traffic separation schemes, restrictions on speed, one-way zones, specified anchorage areas, escorting rules for oil tankers, and pilotage requirements.

Each of the charts showing representative routes also includes pink areas along certain waterways. These depict traffic separation schemes for vessels over 20 meters in length, or regions in which vessels should not travel, keep vessels transiting in opposite directions separated from each other. Areas of convergence of traffic are also depicted and caution is required in these areas. Vessels crossing the separation scheme must do so as close to a right angle as possible. No fishing or anchoring is allowed in the separation scheme area and vessels smaller than 20 meters and sailing vessels are not allowed to impede vessels in the scheme. Vessels not participating in the scheme or crossing the scheme must stay away from the areas depicted. There are also speed restrictions in various areas. In Elliot Bay, vessels are restricted to 5 knots; in Rosario Strait, deep draft vessels are restricted to 12 knots; and in the Saddlebags and Guemes Channel area, vessels are restricted to 6 knots.

The US Coast Guard has also designated a special navigation zone in Rosario Strait. This means that a vessel longer than 100 meters or more than 40,000 DWTs cannot meet, overtake, or cross within 2,000 yards of another vessel that meets these size limits within Rosario Strait. Also towing vessels cannot impede the passage of vessels more than 40,000 DWTs in this area. A similar designation is made in Haro Strait, but just applies to the smaller area at Turn Point, not the whole of Haro Strait. Guemes Channel and the area around Saddlebags and Vendovi Island are also areas where it is difficult for two vessels over 40,000 DWTs to maneuver around each other. While the area is not specifically designated as a special navigation zone, the Puget Sound VTS operates the area as if it were to avoid dangerous situations. Thus the Rosario Strait rules are essentially extended to include the waters east of Rosario Strait in practice.

Vessels requiring anchorage must get approval from the relevant VTS. There are many designated anchorage areas in the region, but four are specifically relevant to this study. Firstly, there is a large general anchorage area at Port Angeles for all deep draft vessels. There are then three anchorages with more limited capacity. Cherry Point anchorage is a short-term anchorage for tankers waiting to dock at Cherry Point or Ferndale. Anchorages around Vendovi Island can be used for longer; there are three designated anchorages for deep draft vessels and two for tugs. Finally, there are four anchorages at Anacortes, with one specifically designated for lightering operations.

The Puget Sound Pilots provide pilotage service for all U.S. ports and places East of 123 degrees 24' W longitude in the Strait of Juan de Fuca, including Puget Sound and adjacent inland waters. Pilotage is compulsory for all vessels except those under enrollment or engaged exclusively in the coasting trade on the west coast of the continental United States (including Alaska) and/or British Columbia. The pilot station is at Port Angeles, meaning that vessels picking up or dropping off a pilot will pass by Port Angeles at a slow speed, allowing a pilot boat to pull aside and the pilot to board or disembark on a pilot ladder. The pilots will navigate vessels to the dock and then back to the Port Angeles on their outbound trip.

Vessels transporting crude oil or petroleum products that are over 40,000 DWTs are required to have a tug escort beyond a point east of a line between Discovery Island and New Dungeness Light.

3.6. Overall Traffic Density

The transit counts and route maps give a general idea of traffic levels in the region, but as mentioned previously ferries have lots of transits on mostly short routes, while container vessels and bulk carriers have fewer transits, but usually on longer routes. Thus to get a true picture of the level of traffic on the water, we need to count the time that each vessel type spends on the water and where it goes. The following maps depict the time on the water with a colored legend and each colored cell is a quarter nautical mile by a quarter nautical miles. Figure 21 shows the generated density of all the traffic discussed thus far. The red areas have higher traffic levels and they include the lanes within the traffic separation schemes, the ferry lanes, and the major fishing areas. There are also some gray and black cells that have the highest levels of traffic within Elliot Bay. However, it is not only important to understand the overall levels of traffic, but it is also of interest to understand what types of traffic make up this overall pattern.

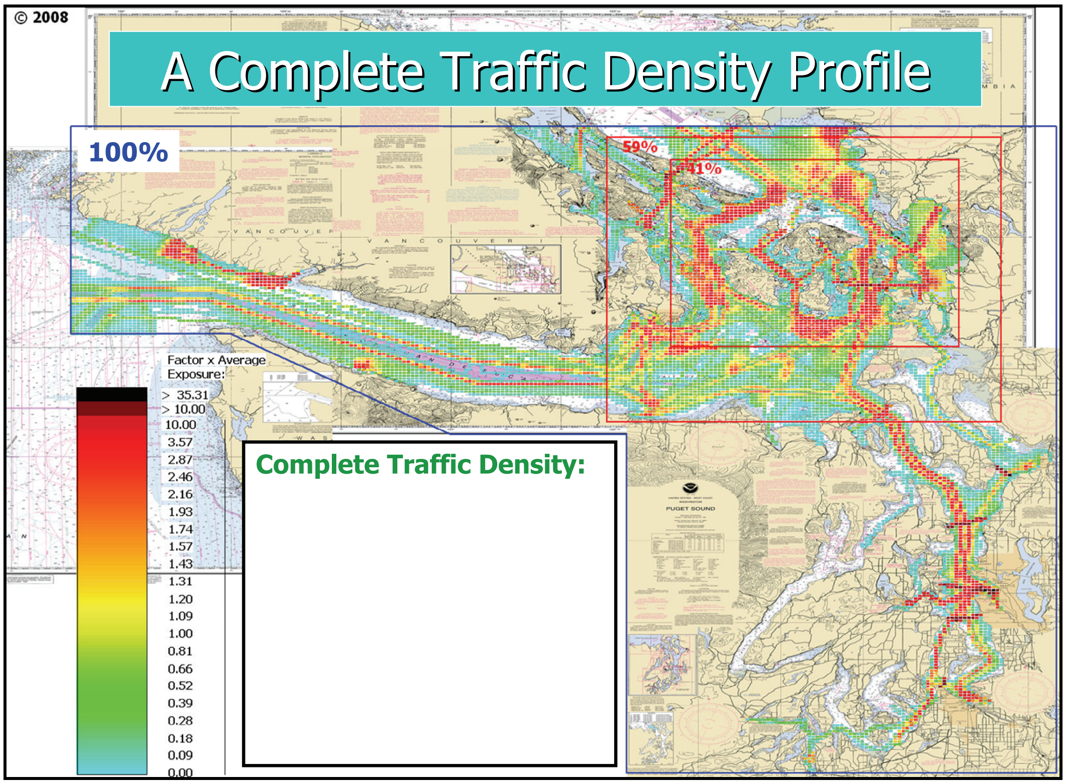


Figure 21. The density of all traffic across the region.

Let us concentrate first on the focus of this study, namely tankers, ATBs, and ITBs calling at BP Cherry Point at some point in their movements within the study area. Figure 22 shows the generated traffic density plot for BP Cherry Point traffic using the same legend as Figure 21 for all traffic. This allows us to see how much of the overall traffic picture is made up of BP Cherry Point traffic. Obviously there are specific areas that BP Cherry Point vessels transit and areas where they will not or cannot. However, of particular interest is that BP Cherry Point traffic only makes up 1.1% of the total time that vessels spend on the water in the study area.

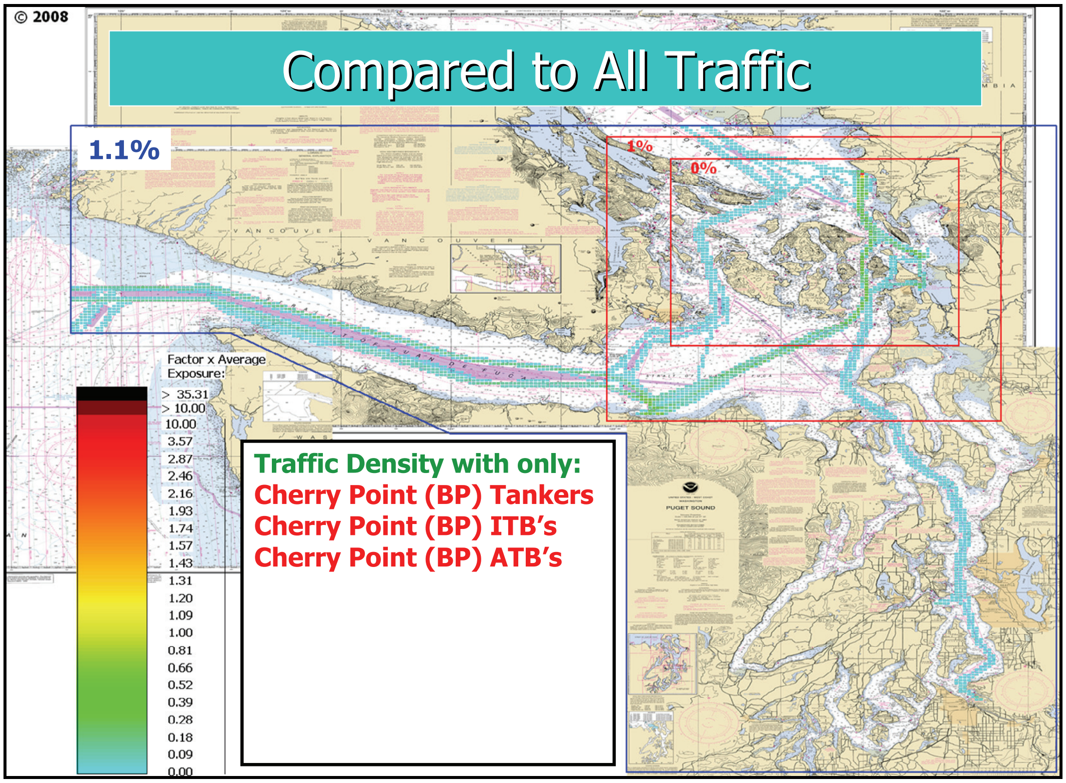


Figure 22. The density of BP Cherry Point traffic across the region.

We can generate similar statistics for other vessel types. Ferries account for 18% of the total transit time on the water even though they account for 50-75% of the transits recorded by the various VTS stations. Tug, towing vessels, and barges account for 17.1% of the total transit time on the water. Small vessel traffic, specifically commercial and tribal fishing, regattas, and whale watching vessels, make up 44.1% of the total transit time on the water. Naval, coast guard, service and supply vessels account for 4.7% of the total transit time on the water. And finally, all oil tankers, ATBs, and ITBs, not just those calling at BP Cherry Point, account for 2.6% of the total transit time on the water, thus BP Cherry Point traffic is 41% of the total for all oil tankers, ATBs, and ITBs.

3.7. Environmental Factors – Wind, Visibility, and Current

The National Climatic Data Center allows one to download hourly weather observations for the VTRA study area. Figure 23 displays seven weather stations for which we have obtained hourly wind speed and direction data based for the year 2005 on their availability and quality as well being able to map them reasonable to the locations identified in Figure 2. The length of the "wind fans" in Figure 23 represents the different wind speeds across these weather stations. As can be observed also from the wind fans in Figure 23, winds tend to be at higher levels at the entrance of the West Strait of Juan de Fuca and further inward than the other locations.

Hourly land visibility data for 2005 was available from the various airports within the study area and has been obtained from the National Climatic Data Center as well. Unfortunately, no electronic data is available for a sea fog phenomenon. Sea fog occurs on the water even with good land visibility. Conditions that determine that are dew point temperature and water temperature as well as wind speeds being below a certain level. A sea fog visibility model using these parameters as input is described in Sanderson (1982). Using the land visibility data, the Sanderson (1982) model and combining it with hourly dew point, water temperature and wind data, information from the 2006 edition of the US Coast Pilot and expert judgment, we have been able to construct hourly visibility conditions for the visibility locations in Figure 24. Figure 25 summarizes the percentage of time bad visibility occurs for these locations by month. The higher levels in the Buoy J area, West Strait and East Strait of Juan de Fuca for the months June, July, August are primarily due to a sea/channel fog phenomenon. The higher levels for the other locations towards the end of the year are primarily representative of a land fog phenomenon.

Current tables were constructed for the year 2005 from the WXTIDE 32 software and cross-checked against those available from the National Oceanic and Atmospheric Administration tides and currents website. A harmonic behavior was modeled in between the max ebb, max flood and slack times of these current tables to evaluate current speeds for these current stations at every minute. Information regarding current directions from these two data sources was integrated and cross checked with those available in the MAPTECH software. Figure 26 displays the available max ebb and max flood directions for 130 current stations within the study area. Please observe from Figure 25 that current strengths vary from one area to the other over the different stations.

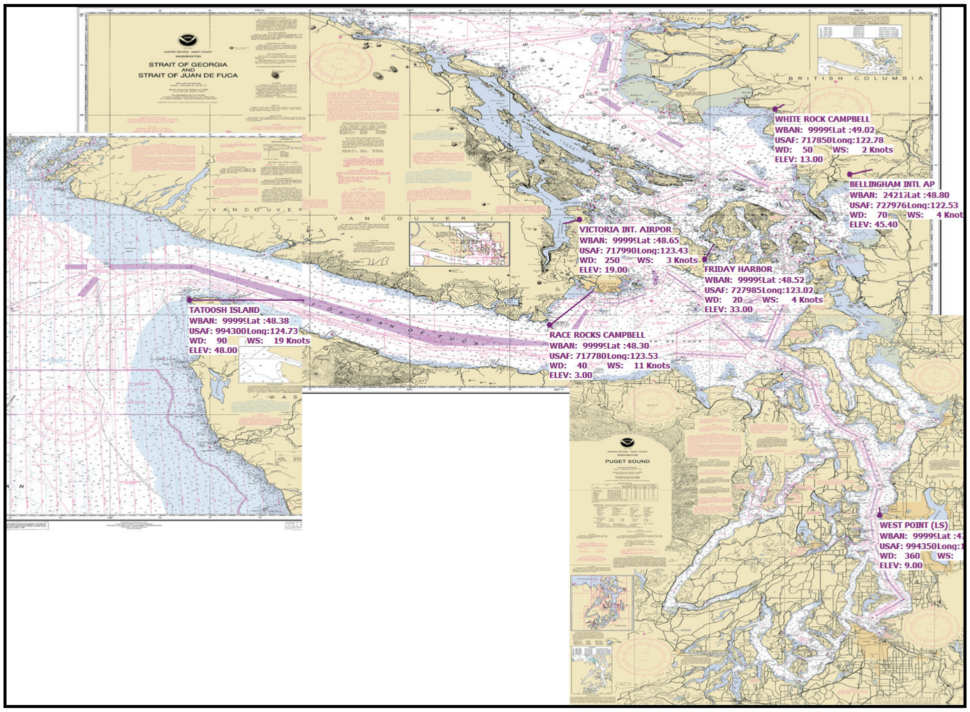


Figure 23. A map displaying the wind stations used the study.

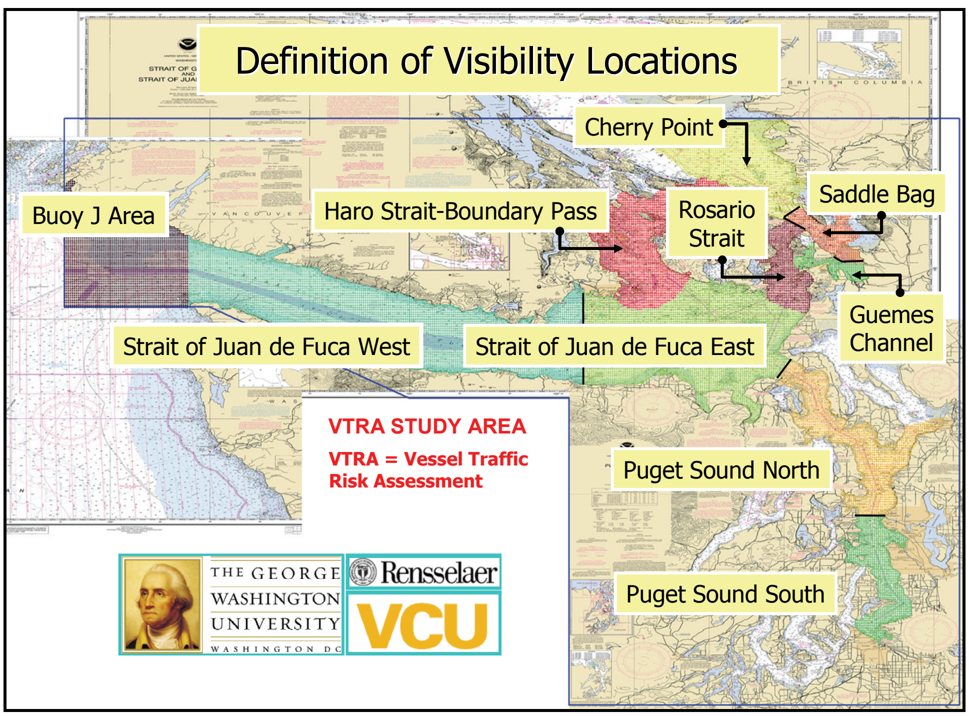


Figure 24. A map defining the visibility locations used in the study.

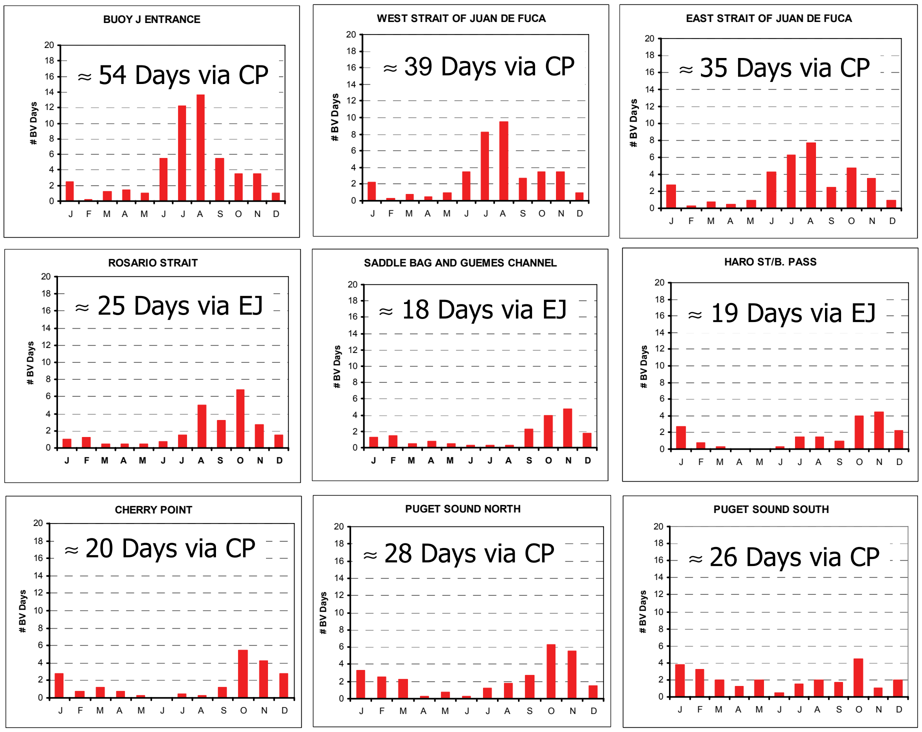


Figure 25. The total number of days with poor visibility by month and location.

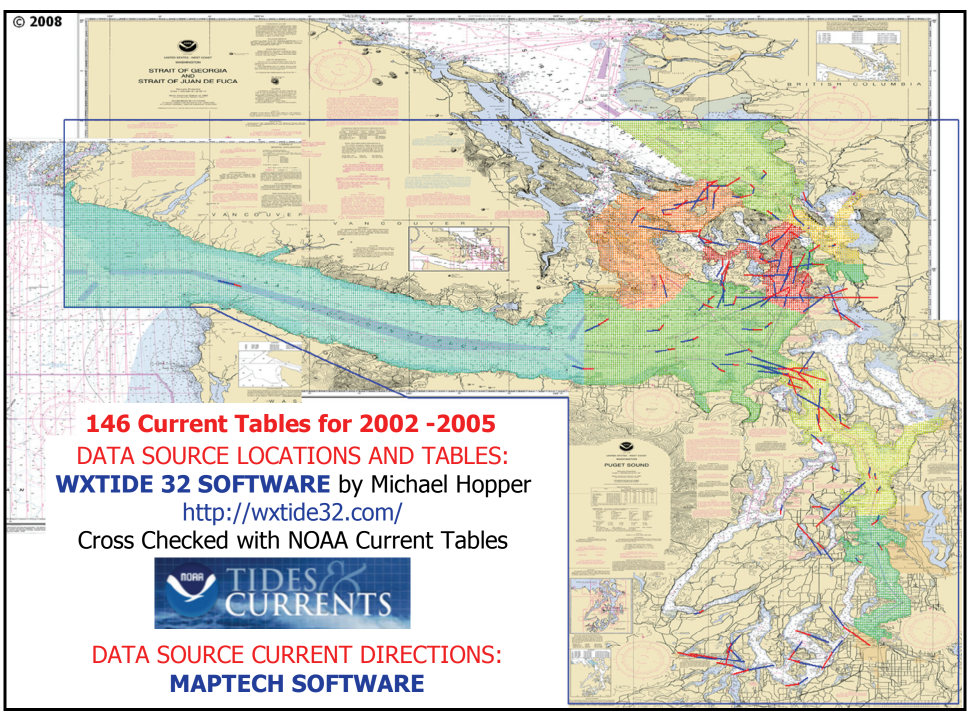


Figure 26. A map displaying the current stations used the study.

4. Model Integration and Data Sources

Our model represents the chain of events that could potentially lead to an oil spill. This model and approach has been used in the Prince William Sound Risk Assessment (Merrick et al, 2002), the Washington State Ferries Risk Assessment (van Dorp et al, 2001), and the Exposure Assessment of the San Francisco Bay ferries (Merrick et al, 2003).

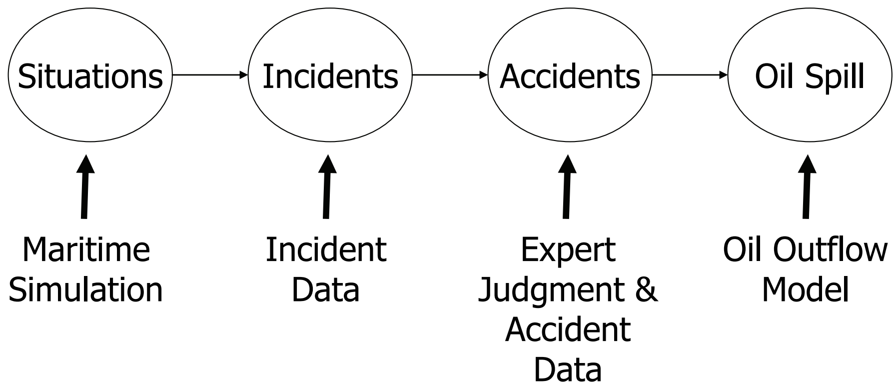


Figure 27. The chain of events that lead to an oil spill and the modeling techniques used for each step.

4.1. Interactions

Accidents can only occur when vessels are transiting through the system. Collisions can only occur when a vessel is in close proximity to another vessel. Grounding can only occur when a vessel is within close proximity (powered grounding) or drifting range (drift grounding) of shore or sufficiently shallow waters. When a BP Cherry Point tanker, ATB, or ITB is in one of these situations, we call this an interaction and the simulation is used to count these interactions. Our maritime simulation model attempts to re-create the operation of vessels and the environment within geographic scope of the study. The routes shown in Sections 3.1 to 3.3 are actually the routes used for vessel transits in the simulation model. The raw records used to obtain the transits counts in Sections 3.1 to 3.3 were used to model the transits and departure times in the simulation for the year 2005. The environmental factors modeled include wind, fog, and current. The underlying data discussed in Section 3.7 was fed in to provide dynamic environmental values in the simulation. Additional details about building maritime simulations can be found in Merrick et al (2002) and van Dorp et al (2001) and Technical Appendix C.

The interactions are counted over the course of a year of the simulation. Figure 27 shows a geographic profile of these counted interactions for the year 2005. Interactions along the shore indicate that a BP tanker, ATB, or ITB are within five hours of shore under power or within five hours of drifting ashore if they were to become disabled. Interactions on the water are with other vessels. There are also interactions with the dock. Informally, darker colors indicate more interactions and lighter colors indicate less. A black cell has one of the highest interactions of any cell in the study area; the light blue cells have the least. The light greenish-yellowish color to the left of the number 1.00 of the color legend's numerical scale represents the average number of interactions within a grid cell over the entire area.

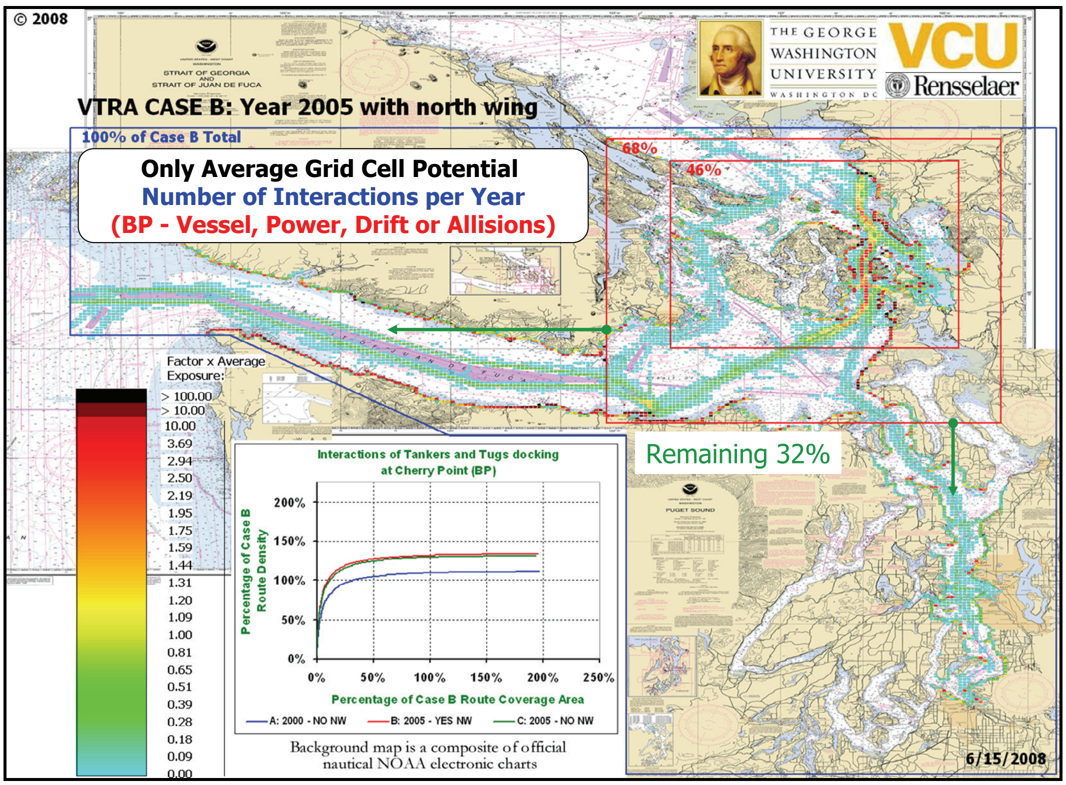


Figure 28. A geographic profile of the number of interactions counted in a simulation of 2005.

Hence, colors darker than this color (above 1.00) have more than an average number of interactions in a grid cell and colors below 1.00 have a lower than average number of interactions in a grid cell. The two red rectangles provide in the upper left corner the percentage of interactions within that rectangle. Hence, 68% of all the interactions in Figure 27 occur within the largest red rectangle and 46% of all the interactions occur within the smallest red rectangle. The remaining 32% occur outside of the largest rectangle (but within the blue border area).

4.2. Incidents

Incidents are the events that immediately precede the accident. The types modeled include total propulsion losses, total steering losses, loss of navigational aids, and human errors. The impact of each of these types of triggering events on the occurrence of accidents is estimated by examining the records of each accident that occurs inside the study's geographic scope. An exhaustive analysis of all possible sources of relevant accident, near miss, incident, and unusual event data was performed. The tanker fleet calling at BP Cherry Point has experienced xx propulsion failures, xx steering failures, and xx navigational aid failures while within the study area over the 11 year period from 1995 to 2005. The ATB and ITB fleet that call at BP Cherry Point have not been operating for as long as the tankers, just 7.5 years. Over this period they have experienced 34 propulsion losses, 13 steering losses, and 12 navigational aid failures while within the study area. These counts are used to find the probability of a propulsion failure, steering loss, or navigational aid failure during each interaction that is counted in the simulation.

Human errors are not recorded as reliably as the mechanical failures discussed above. Thus we must find another method to estimate their frequency. If we perform an error analysis of accidents that have occurred in our data collection period (1995 to 2005), we find that 75% (3 of 4) of the accidents have been preceded by human errors, while 25% (1 of 4) have been caused by mechanical failures. This is in line with such percentages found in previous studies (Grabowksi et al 2000). Thus there are 3 times as many accidents preceded by human error than accidents preceded by mechanical failure. Thus we infer that there must be 3 times as many human errors as mechanical failures and we multiply the total number of propulsion losses, steering losses, and navigational failures by 3 to obtain the number of human errors. This number is then used to find the probability of a human error during each interaction that is counted in the simulation. Appendix A discusses in more detail the collection of incident and accident data for the VTRA study area.

4.3. Accidents

The accident types included in this study are collisions between two vessels, groundings (both powered and drift), and allisions. However, as the simulation counts the situations in which accidents could occur, it also records all the variables that could affect the chance that the accident will occur; these include the proximity of other vessels, the types of the vessels, the location of the situation, and the environmental variables from Section 3.7. We know how often accidents do occur from our analysis of incident and accident data, but there is not enough data to say how each of these variables affects the chances of an accident; accidents are rare! To determine this, we must turn to the experts. We ask experts to assess the differences in risk of two similar situations that they have extensive experience of. In each question we change only one factor and through a series of questions we build our accident probability model, incorporating the data where we can. The type of incident that has occurred to lead to the possibility of an accident is also specified for each question.

Q28

Situation 1	TANKER DESCRIPTION	Situation 2
Strait of Juan de Fuca East	**Location**	-
Outbound	**Direction**	-
1 Escort	**Escorts**	-
Untethered	**Tethering**	-
	INTERACTING VESSEL	
Passenger vessel	**Vessel Type**	-
Meeting	**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 29. An example of a question used to assess the variation in accident probabilities between the different possible interaction scenarios.

Figure 28 shows an example picture; here a total propulsion loss has occurred and an oil tanker with a tethered escort is meeting a ferry. The question asks how much difference restricted visibility would make. This method has been developed over the course of over ten years of work in maritime risk assessment, has been peer reviewed by the National Research Council and our peers in the field of expert elicitation design and analysis, and has been improved thanks to funding from the National Science Foundation. The experts involved include tanker masters, tug masters, pilots, Coast Guard VTS operators, and ferry masters. Additional details about this method, how the responses are analyzed, and how the results are incorporated in to the over model can be found in Technical Appendix D and Szwed et. al (2006).

As these questions compare two scenarios, they can only be used to estimate the difference in the accident probability between two interaction scenarios. We still need to know the total number of accidents. Thus the accident probability model is calibrated to reflect the historical number of accidents that have occurred to the BP Cherry Point calling fleet within the study period over our data collection period. In all, there have been 4 accidents, 1 collision, 1 grounding, and 2 allisions.

Combining the information of geographic interaction profile in Figure 28 with the accident probability models per interaction allows one to develop a geographic profile of accident frequencies results. Figure 30 shows such a geographic profile for the year 2005. Please compare Figure 28 with Figure 30 and note that when going from interactions counts (exposure) to accident frequency the largest rectangle contains 88% of the total annual accident frequency, but contained 68% percent of the interactions. Going from interactions counts (exposure) to accident frequency the smallest rectangle contains 79% of the total annual accident frequency, but contained 46% percent of the interactions. Hence, we observe a higher concentration of accident frequency within these rectangles, compared to interaction counts. Even the though the color legends of Figures 28 and Figure 30 have different scales (since the yellow-greenish color to the left of 1.00 on the numerical scale of the color legend represents the average number of interactions over all grid cells in Figure 28 and the average accident frequency over all grid cells in Figure 30) this also follows from a lightening of colors along the coast lines in Figure 30 and a darkening of color within the smallest red rectangle in Figure 30.

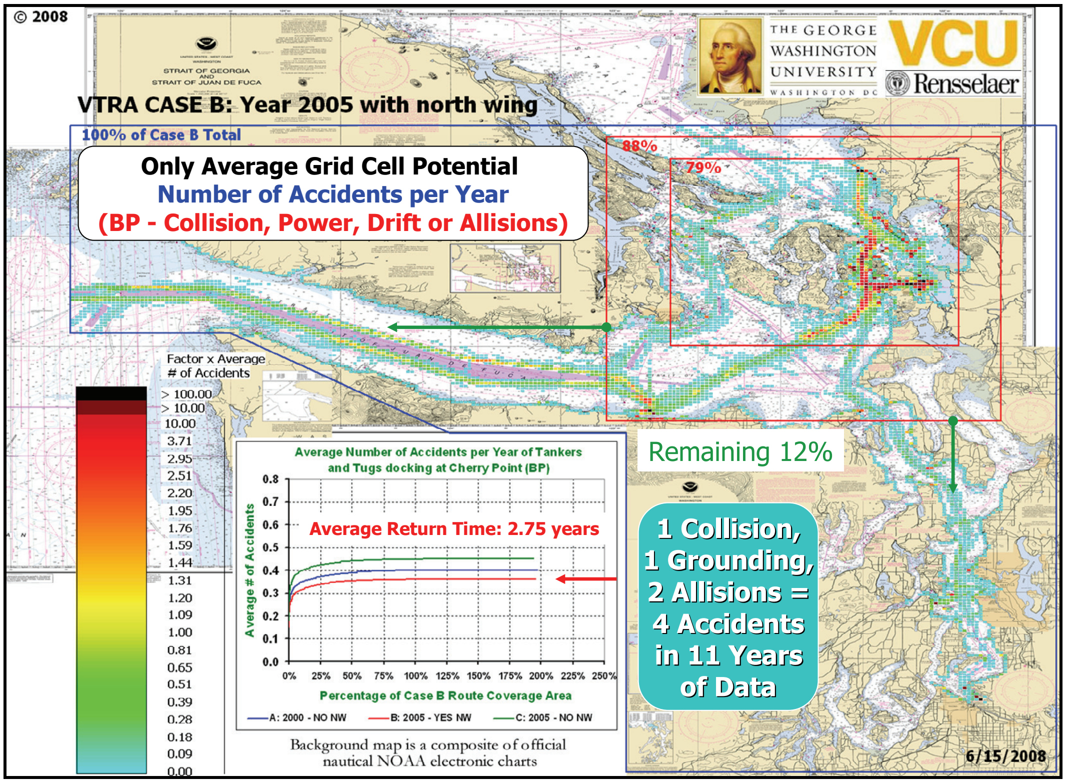


Figure 30. A geographic profile of accident frequency results for a 2005 analysis with the north wing dock in operation.

4.4. Oil Outflow

Our oil outflow methodology is derived from the one described in Special Report 259 published by the National Research Council (NRC) in 2001. For tankers, ATB's and ITB's we use the compartment configurations for single hull and double hull tankers provided in this NRC (2001) report. We make the worst case assumption that when a compartment is punctured that all its content is lost. Within the simulation, the speed and types of the vessels involved in each interaction are recorded along the angle of interaction for interactions between vessels. BP shipping provided the DWT, displacement, length, beam, and draft of each tanker, ATB, and ITB, along with the hull type. For each other type of vessels, DWT, length, beam, and draft were obtained and representative configurations of the fuel tanks developed. These were inputs to the oil outflow model.

BP could not provide specific details about how much crude or petroleum tankers carried on specific voyages. Instead, they provided the maximum capacity of each tanker and an average percentage of capacity carried based on the type of vessel and type of transit. Crude tankers are assumed to arrive in the study area full and leave empty. However, they can also make multiple calls at refinery docks during one visit to the study area. Thus they are assumed to unload an equal quantity of crude oil at each refinery. Product tankers are assumed to arrive in the study area empty and leave full. If they make multiple calls at refineries, then they are moving product between refineries in the study area. Thus they are assumed to be half full on each inter-refinery transit. All other vessels were assumed to carrying their full capacity of fuel as a worst-case assumption.

Once an accident has occurred, we must estimate the probability that the hull (or hulls in the case of double hulls) is punctured and then estimate how many compartments that are carrying crude cargo, product cargo, heavy fuel or diesel fuel have been penetrated. The speed and mass of the vessels are used to calculate the kinetic energy involved in the collision or grounding with the other vessel, the shore, or the dock, but is this kinetic energy enough to penetrate the hull of the tanker and, if so, how far in to the tanker will the penetration be? If we know this, then we can overlay this penetration on a picture of the vessel and determine which compartments are penetrated. The National Research Council 2001 study performed a large scale modeling study of oil spills for collisions and groundings using physical simulation models and described in NRC (2001). They studied both 40,000 DWT and 150,000 DWT tankers, with the smaller vessel configured like a product tanker and the larger like a crude tanker, and both single hull and double hull tankers of each size. 10,000 collision simulations and 10,000 grounding simulations of each of the four types of tankers were performed at multiple levels of each input factor. Rather than repeating these simulations, we fitted regression models to the data and then applied the fitted models to estimate the probability of penetration and the number of compartments penetrated in each interaction scenario. The regression models allow for an interpolation between the specific tankers sizes studied in the NRC (2001) report. Our oil outflow model is described in more detail in Appendix E.

Combining the information of geographic accident frequency profile in Figure 30 with the oil outflow models per accident allows one to develop a geographic profile of oil outflow results. Figure 31 shows such a geographic profile for the year 2005. Please compare Figure 30 with Figure 31 and note that when going from interactions counts (exposure) to accident frequency the largest rectangle contains 92% of the total annual average oil outflow, but contained 88% percent of the overall accident frequency. Going from accident frequency to oil outflow the smallest rectangle contains 77% of the total annual accident frequency, but contained 79% percent of the interactions.

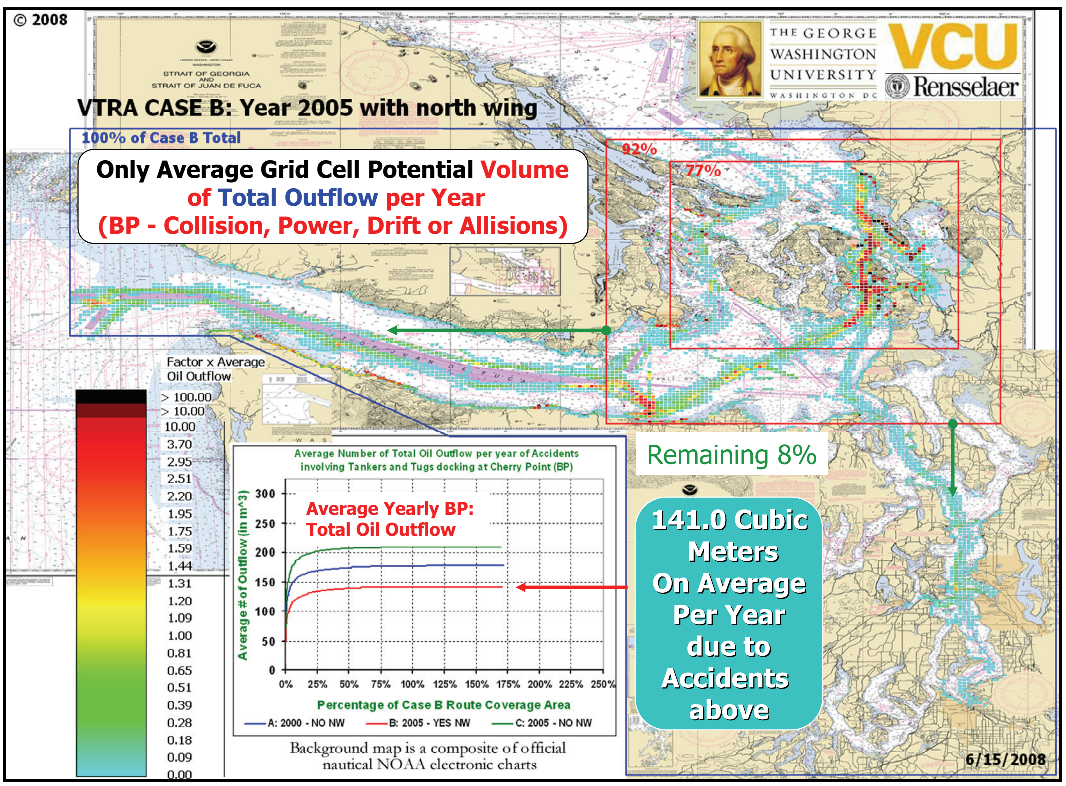


Figure 31. A geographic profile of average oil outflow results for a 2005 analysis with the north wing dock in operation.

Hence, we observe a higher concentration of oil outflow within the largest red rectangles, but a lower one in the smallest red rectangle compared to accident frequency. Even the though the color legends of Figures 28 and Figure 30 have different scales (since the yellow-greenish color to the left of 1.00 on the numerical scale of the color legend represents the average accident frequency over all grid cells in Figure 30 and the average oil outflow over all grid cells in Figure 31) also follows by and a darkening of color within the smallest red rectangle in Figure 31 and a darkening of color within the largest red rectangle but outside of the smallest one.

4.5. Organizations that Provided Experts

The organization below provided experts for the expert judgment elicitation sessions. Experts were invited to and referred to the VTRA team through the United States Coast Guard and the Puget Sound Harbor Safety committee. 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. The combined numbers of years sailing experience of the experts who participated in the elicitation process of the VTRA study area exceeds 922 years.

1. Puget Sound Pilots
2. ATC
3. BP Shipping North America
4. 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
5. The Washington State Ferries
6. Seattle sector US Coast guard VTS.

4.6. Data Sources Used

The organizations below have contributed to our data collections processes in various forms. Some provided data in electronic form, some sources were in hard copy and others assisted in a data assimilation process through personal communications. The data sources and their format are described in more detail in the technical appendices.

1. VTOS
2. The Washington Department of Fish and Wildlife
3. Washington State Fisheries
4. Canadian Fisheries
5. Native American Tribes
6. Sound Watch
7. National Climatic Data Center
8. NOAA Weather Buoys
9. NOAA Current Data
10. US Coast Pilot 7 - 2006 (38^{th}) Edition.
11. USCG Accident/Incident Data
12. USCG Small Events Permitting Data
13. Washington State Department of Ecology Accident/Incident Data
14. Puget Sound Pilots Incident Data
15. National Research Council Report (2001).
16. Fuel vessel data from various vessel brokerage web-pages e.g.: www.yachts.com and www.ship-technology.com

5. Analysis Results

In this section, we examine the results obtained using the model discussed in Section 4. We start by describing the cases analyzed using this model and then proceed to understand the differences in the level of accident potential and oil outflow potential between these cases, including changes since before the north wing, changes that might be seen in the future, and changes caused by three alternatives.

5.1. Explanation of Cases Analyzed

The analysis is based on 15 cases that represent different configurations of the simulation. The first three cases A, B, and C allow us to examine the changes from before the north wing was constructed to the year 2005 after it was operational; we also examine a hypothetical scenario were the north wing to not have been operational in the year 2005.

We then have six scenarios which examine potential changes in risk in the year 2025. Since shipping is a derived demand, projection of future vessel traffic is inherently uncertain. Actual future tanker and tank barge traffic will be dependent upon energy requirements and distribution choices. Actual future container vessel traffic and bulk cargo traffic and vessel size are dependent upon demand for imports and exports. As traffic levels are uncertain that far in to the future, we use statistical forecasts of traffic levels. However, these forecasts provide a best guess estimate, but also an assessment of the level of uncertainty, which allows us to give high and low estimates too. The vessel traffic described in the base case year (2005) was projected through 2025 using 15 years historical trend data analysis by vessel type. The opening of the Gateway bulk cargo terminal (that would effect this time series projection) and statistical techniques were used to construct upper and lower bounds for future traffic. The resulting high, medium and low forecasts were used as the basis for calculating future accident frequencies and oil outflows. Appendix F discusses the development of these future scenarios at a higher level of technical detail.

Thus cases D through I represent high, medium, and low traffic scenarios both with and without the north wing being operational. This allows us to assess the risk affect of the north wing at different levels of forecasted traffic. The final six cases evaluate changes to the operations of the BP Cherry Point tanker, ATB, and ITB traffic. Each of these changes are evaluated at the case B traffic levels and at the high forecasted level of 2025 to stress test the alternative. These changes represent risk interventions out of scope in this study if the north wing is operational, so the north wing is operational in all six cases J through O. The three risk interventions are not using Saddlebags for BP Cherry Point traffic, extending escorts to the whole area inside Buoy J, and taking out the Neah Bay tug (to assess the affect of it being included in the other cases).

The 15 cases are summarized in Table 1. We will break our discussion in to three parts, a discussion of cases A, B, and C to assess changes from 2000 to 2005 with the construction of the north wing, cases D through I to assess future changes in risk in 2025, and cases J through O to assess the affect of the three risk interventions relative to the entire VTRA study area.

Table 1. A list of all cases used in the analysis and the factors varied amongst them.

	Case	CP Traffic	Other Traffic	North Wing?	Saddlebags?	Extend Escorting?	Neah Bay?	Gate Way?
1	A	2000	2000	No	Yes	No	Yes	No
2	B	2005	2005	Yes	Yes	No	Yes	No
3	C	2005	2005	No	Yes	No	Yes	No
4	D	2025 Low	2025 Low	Yes	Yes	No	Yes	Yes
5	E	2025 Low	2025 Low	No	Yes	No	Yes	Yes
6	F	2025 Medium	2025 Medium	Yes	Yes	No	Yes	Yes
7	G	2025 Medium	2025 Medium	No	Yes	No	Yes	Yes
8	H	2025 High	2025 High	Yes	Yes	No	Yes	Yes
9	I	2025 High	2025 High	No	Yes	No	Yes	Yes
10	J	2005	2005	Yes	No	No	Yes	No
11	K	2025 High	2025 High	Yes	No	No	Yes	Yes
12	L	2005	2005	Yes	Yes	Yes	Yes	No
13	M	2025 High	2025 High	Yes	Yes	Yes	Yes	Yes
14	N	2005	2005	Yes	Yes	No	No	No
15	O	2025 High	2025 High	Yes	Yes	No	No	Yes

Let us first discuss cases A, B, and C in some detail. The base case is Case B which represents the operation of maritime traffic in the year 2005. The simulation replays VTS traffic, regattas, whale watching, and fishing traffic from the year 2005. The wind, visibility, and currents are also replayed from 2005. The north wing is operational, the BP Cherry Point tanker, ATB, and ITB traffic can use Saddlebags, and the Neah Bay tug is on standby; the tankers are escorted beyond a point east of a line between Discovery Island and New

Dungeness Light. Case A reflects the operation of the system before the north wing was constructed. The traffic levels reflect operations in the year 2000; much of the traffic has been consistent from 2000 to 2005, but product traffic at BP Cherry Point was 20% less in 2000 than 2005, while other tanker traffic was 23% higher in 2000. Bulk carriers were also 30% higher in 2000. The north wing was not constructed, so only one dock was available at BP Cherry Point. Case C is a fictional case that reflects the system in the year 2005 as in case B, but with the hypothetical scenario that the north wing was not operational.

The future scenario cases represent high, medium, and low forecasts of tanker traffic, container traffic, and bulk carrier traffic. Tanker traffic is broken in to BP crude tanker traffic, BP product traffic, and non BP traffic. For BP traffic, the low scenario represents the possibility that most crude is brought to the BP Cherry Point refinery by pipeline; the medium scenario is a moderate increase in both types of traffic; the high scenario represents the highest crude traffic level possible under currently permitted operations and a large increase in product traffic. For the low scenario, BP crude traffic is decreased to only 10% of the 2005 levels and BP product traffic is reduced by 2%. For the medium scenario, BP crude and product traffic are increased by 13%. For the high scenario, BP crude traffic is increased by 17%, but BP product traffic is increased by 90%. Other tanker traffic is forecasted to increase, but the uncertainty bounds are large due to the long forecast time horizon. Thus the low, medium, and high levels are a 52% decrease, a 55% increase, and a 162% increase. Container traffic is also forecasted to increase, but again with large uncertainty bounds. The low, medium, and high levels are a 54% decrease, a 20% increase, and a 93% increase. Bulk carrier traffic has decreased significantly over the past ten years, but for the past few years it has been consistent. However, with the renewed permit action associated with the Gateway facility, 241 additional bulk carrier transits are included from Buoy J to the Gateway facility and then back to Buoy J in the 2025 cases.

The risk interventions identified in the scope are run at case B traffic levels as well as the high future scenario. In cases J and K, BP tankers do not use the Saddlebags route to transit between BP Cherry Point and Anacortes, but instead use Rosario Strait and Guemes Channel. In cases L and M, tugs escort all tanker traffic from and to Buoy J. This is reflected in the collision and grounding probabilities. In cases N and O, the effect of the Neah Bay tug on drift grounding probabilities is removed; in other cases the time that the Neah Bay tug would take to reach a drifting tanker is calculated and the effect of a tug on the probability of grounding is applied proportionally.

5.2. Risk Changes from Adding North Wing

In this section, we will examine cases A, B, and C. The specifics for these three cases are included in Table 2 as a reminder.

Table 2. The cases used to consider changes in risk from adding the north wing.

	Case	CP Traffic	Other Traffic	North Wing?	Saddlebags?	Extend Escorting?	Neah Bay?	Gate Way?
1	A	2000	2000	No	Yes	No	Yes	No
2	B	2005	2005	Yes	Yes	No	Yes	No
3	C	2005	2005	No	Yes	No	Yes	No

Comparison of cases A and B allow us to see the change in risk from 2000 to 2005; this is caused by both changes in traffic levels and the construction of the north wing. However, using case C we can separate these two effects. Comparing cases B and C allows us to assess the effect of just construction of the north wing, but with no changes in traffic. Comparison of cases A and C allows us to assess the effect of just changes in traffic from 2000 to 2005, but without the effect of the north wing. Figure 32 shows the accident potential results for cases A, B, and C. The total accident potential is the sum of the potential of the four accident types, allisions, drift groundings, powered groundings, and collisions. Thus Figure 29 shows the accident potential for each type of accident stacked one on top of the other. This means that the total height of the bar is the total accident potential for that case. Figure 30 shows the same stacked bars for oil outflow potential in these three cases.

Comparing case A to case C, we see the expected effect of an increase in traffic levels calling at BP Cherry Point with no other changes to the system (no north wing in either case): an increase in overall levels of risk. It can be noticed that the potential number of collisions, drift groundings, powered groundings, and allisions increases. There are 25% more BP crude tankers in case C than case A and the same number of BP product vessels, so with more BP vessels there are more accidents in case C. Overall, case C has 12% higher accident potential than case A.

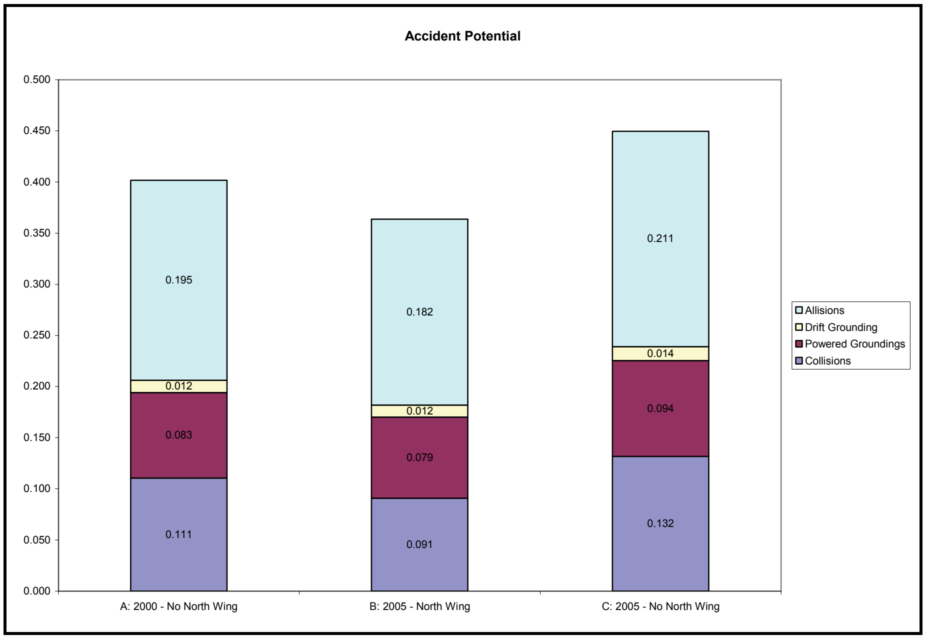


Figure 32. The accident potential by accident type for cases A, B, and C.

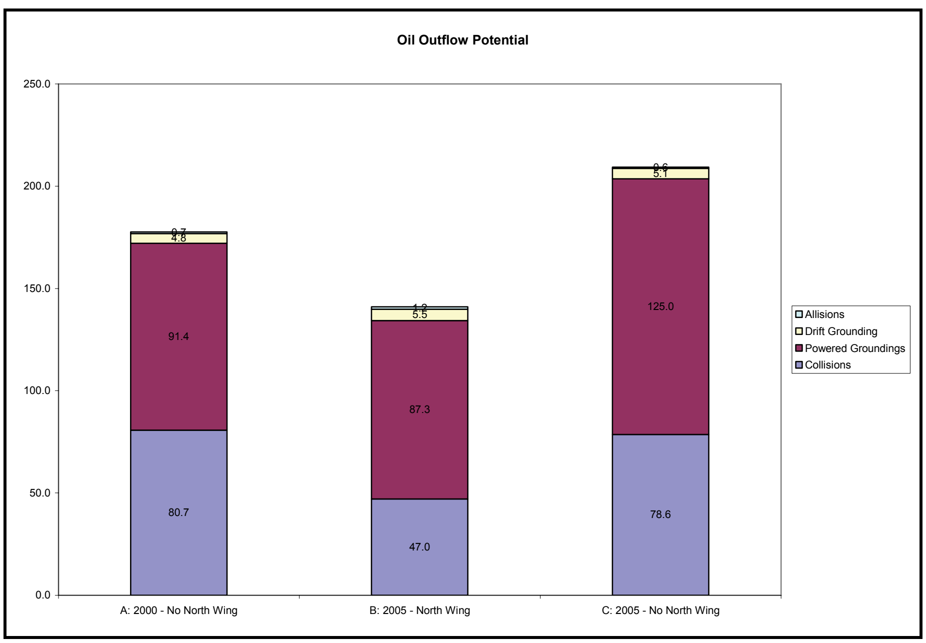


Figure 33. The oil outflow potential by accident type for cases A, B, and C.

We see some different effects in oil outflow as not all accidents cause the same level of potential oil outflow. Allisions have a much lower amount of oil outflow than the other three accident types. Oil outflow from potential powered groundings increases from case A to case C due to the increase in the number of BP crude tankers and the associated increase in the use of anchorages[4]. However, we can also see that the potential oil outflow from collisions is about the same though the potential number of collisions increases from A to C. Oil outflow from collisions depends on what the other vessel involved is. In case C, the BP vessels interact more with fishing vessels and less with ferries than case A. Increasing interactions with fishing vessels leads to more collision potential, but not a lot more oil outflow potential. Increased interactions with ferries leads to both collision and oil outflow potential. Thus case C has more collision potential because of the fishing vessels, but the oil outflow potential is a wash compared to case A because of the higher number of ferries interactions in Case A. Putting the effects on all types of accidents together though, there is a total increase of 18% from case A to case C in potential oil outflow.

Comparing case B to case C allows us to see the effect of the north wing with no other changes to the system[5]. We can see that the levels for each accident and the total level are lower in case B than in case C. The oil outflow potential is also higher without the north wing. This is because incoming tankers, ATBs, and ITBs do not have to transit at reduced speed or go to anchorage as often because the north wing is available. In fact the number of trips to anchorage for vessels inbound to BP Cherry Point is reduced by 40% if the north wing is available. This reduces the time that each tanker, ATB, and ITB spends in the study area on a given visit. However, taking a look at the other side of this same coin, in case C we use the same schedule for tankers as case B[6] but without the north wing, there is a 4% reduction in the number of BP tanker, ATB, and ITB visits in the one year of simulation as they can't pass through the study area as quickly. Despite handling slightly more BP vessels

[4] The trips to anchorage take vessels through waterways where the shore is closer.

[5] The traffic levels are the same

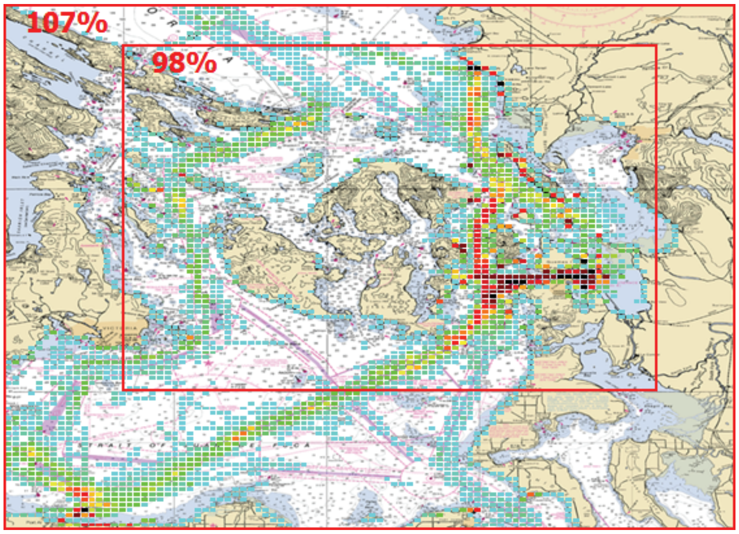
[6] Meaning each vessel spends the same amount of time out of the study area between calls and the same amount of time at each dock they visit

with the north wing than without it, having the north wing in 2005 reduces the accident potential by 21% and the oil outflow potential by 38%[7].

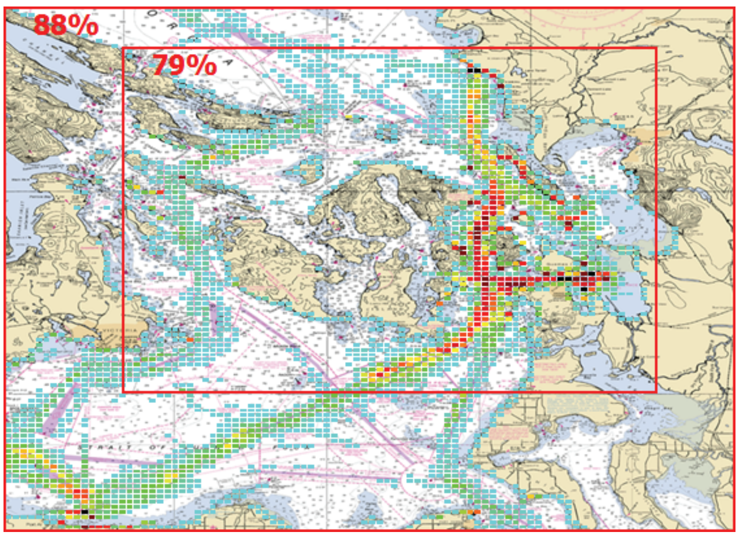
What then is the change from 2000 to 2005? The changes in traffic over this period increased accident and oil outflow potential in our case A and case C comparison. But at fixed traffic levels, having the north wing is better than not having it in our case B and case C comparison. What is the change in risk caused by both the changes in traffic levels and the addition of the north wing? Comparing case A to case B, we can see that the potential for all types of accidents decreases. The level of oil outflow potential is also lower. Overall, there is a 10% reduction in accident potential and a 21% reduction in oil outflow potential. Thus we can say that the addition of the north wing has reduced the risk to BP vessels in the study area despite the increase in crude tankers calling at BP Cherry Point from 2000 to 2005. Another way of saying this is that the addition of the north wing has mitigated the effect of traffic changes from 2000 to 2005.

It is useful to see the changes between these three cases on a geographical profile (Figures 34). A geographic profile is generated by counting the potential number of accidents or summing the potential volume of oil outflow in a grid of cells and then overlaying these amounts on a map of the study area. The cells are colored to indicate higher of lower amounts. The color scheme goes from blue for the lowest amounts, through green to yellow for average amounts, through orange, red and brown to black for the highest amounts. Figure 34 shows three such maps for the potential number of accidents, one each for cases A, B, and C. The maps show an area that includes BP Cherry Point to the top right and Port Angeles to the bottom left. In case B, this area includes 88% of the total potential number of accidents for the whole study area. In case A, this area has 7% more accidents than the whole study area does in case B. In case C, this area has 12% more accidents than the whole study area does in case B. These percentages are shown on the maps along with percentages for a smaller area shown as a red box.

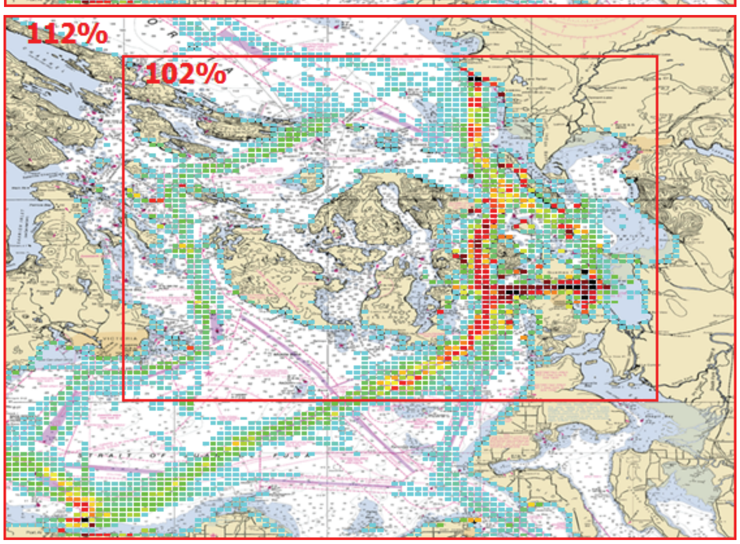
[7] For consistency percentages are evaluated here as percentages of 2000 levels as well.



Case A: Accident Potential

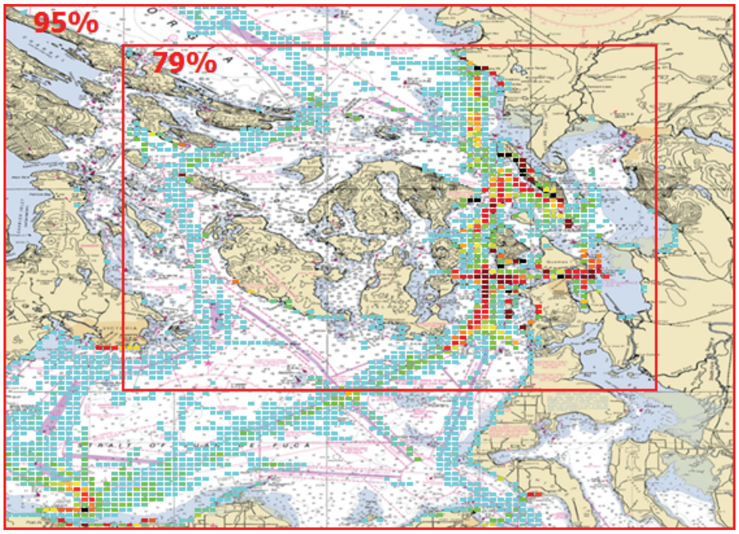


Case B: Accident Potential

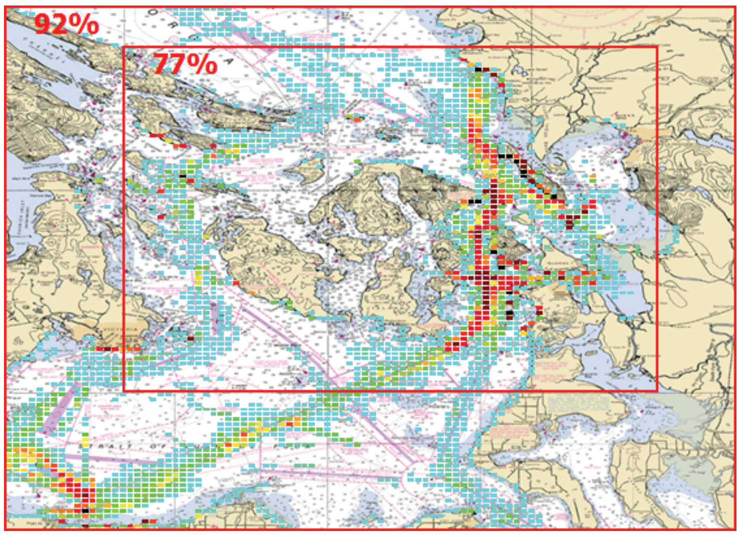


Case C: Accident Potential

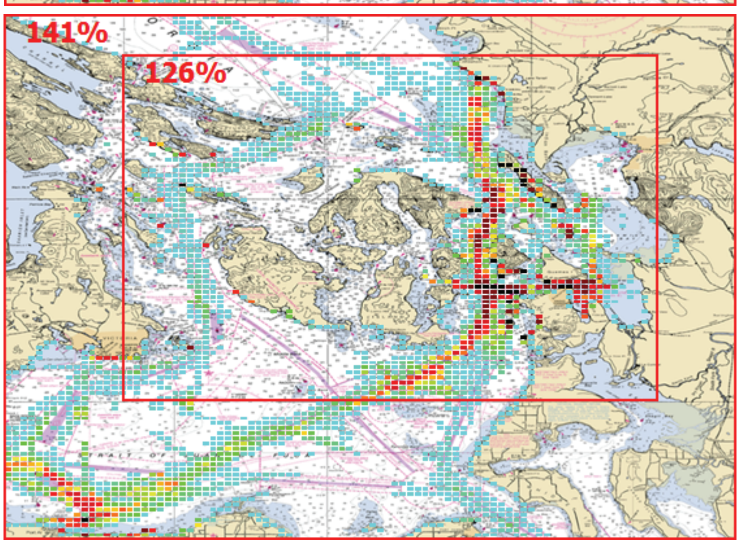
Figure 34. Geographic profiles of accident potential for cases A, B, and C.



Case A: Oil Outflow



Case B: Oil Outflow



Case C: Oil Outflow

Figure 35. Geographic profiles of oil outflow potential for cases A, B, and C.

Comparing the three maps in Figure 34 shows that cases A and C have more dark cells in Guemes channel, around Saddlebags and its approach from Rosario Strait. This is because tankers, ATBs and ITBs have to use the anchorages more often when there is no north wing. Cases B and C also have some brown cells higher up Rosario Strait where case A has red cells. There are also more red cells in case B and C than in case A on the approach to Rosario Strait at its south end. These effects are both because there are more BP tankers transits in cases B and C than in case A. Figure 35 shows similar maps to Figure 34 but showing a color scheme that depicts the potential volume of oil outflow. The interpretation of the color scheme is the same, but it is now showing higher and lower levels of potential oil outflow rather than potential numbers of accidents. Examination of these geographic profiles shows the same patterns of behavior in terms of oil outflow as the accident profiles.

5.3. Future Changes in Risk

In this section, we will examine cases D through I, the 2025 scenarios, and compare them to cases B and C for the year 2005. The specifics for these eight cases are included in Table 3 as a reminder.

Table 3. The cases used to consider future changes in risk.

	Case	CP Traffic	Other Traffic	North Wing?	Saddlebags?	Extend Escorting?	Neah Bay?	Gate Way?
2	B	2005	2005	Yes	Yes	No	Yes	No
3	C	2005	2005	No	Yes	No	Yes	No
4	D	2025 Low	2025 Low	Yes	Yes	No	Yes	Yes
5	E	2025 Low	2025 Low	No	Yes	No	Yes	Yes
6	F	2025 Medium	2025 Medium	Yes	Yes	No	Yes	Yes
7	G	2025 Medium	2025 Medium	No	Yes	No	Yes	Yes
8	H	2025 High	2025 High	Yes	Yes	No	Yes	Yes
9	I	2025 High	2025 High	No	Yes	No	Yes	Yes

Comparing case B to cases D, F, and H allow us to see the changes in risk from 2005 to 2025 if the north wing is operational, along with the range of uncertainty about the risk levels in the future with the north wing. Comparing case C to cases E, G, and I allow us to see the changes in risk from 2005 to 2025 if the north wing is *not* operational, along with the range of uncertainty about the risk levels in the future without the north wing. The uncertainty in future risk levels is derived from the level of uncertainty in the traffic levels that may be seen in 2025, including BP tanker, ATB, and ITB traffic and other types of traffic.

We may also make another comparison; we have already compared cases B and C that have the same 2005 levels of traffic, one with the north wing operational and one without it. This showed that the level of accident potential and potential oil outflow is lower if north wing is operational. However, we may also assess the effect of the north wing being operational in each possible future traffic level by comparing cases D and E, cases F and G, or cases H and I. In each of these comparisons the traffic levels are kept the same and only the operation of the north wing differs. The results for each of the cases in terms of accident potential are shown in Figure 36. Cases B and C are shown as individual points as the traffic levels are not uncertain for 2005. However, cases D, F, and H are shown on a line as they all correspond to 2025 with the north wing, but for different possible traffic scenarios that are possible in that year. This shows the range of uncertainty in the level of accident potential in 2025 if the north wing is operational. Cases E, G, and I are shown in a similar fashion to show the range of uncertainty in 2025 if the north wing is not operational. The results for oil outflow potential are shown in the same fashion in Figure 37.

Comparing case B to cases D, F, and H, we see that the potential number of accidents is higher in cases F and H (the medium and high traffic cases) than in case B. However, the potential number of accidents is lower in case D (the low traffic case) than in case B. So there is no guarantee that the risk will increase from 2005 to 2025; it depends on what happens to the traffic levels and the number of vessels that call at the BP Cherrypoint dock.. Comparing case C to cases E, G, and I shows the same is true were the north wing to not be operational in 2005 and 2025.

It is tempting to now compare the range of cases D, F, and H to the range of cases E, G, and I in Figure 36. In a statistical sense, if these ranges overlap then we might conclude that we do not have enough evidence to say that the north wing will reduce risk levels in the future as it does in 2005[8]. However, this is not the correct approach.

[8] Case B has lower accident potential and oil outflow potential than case C

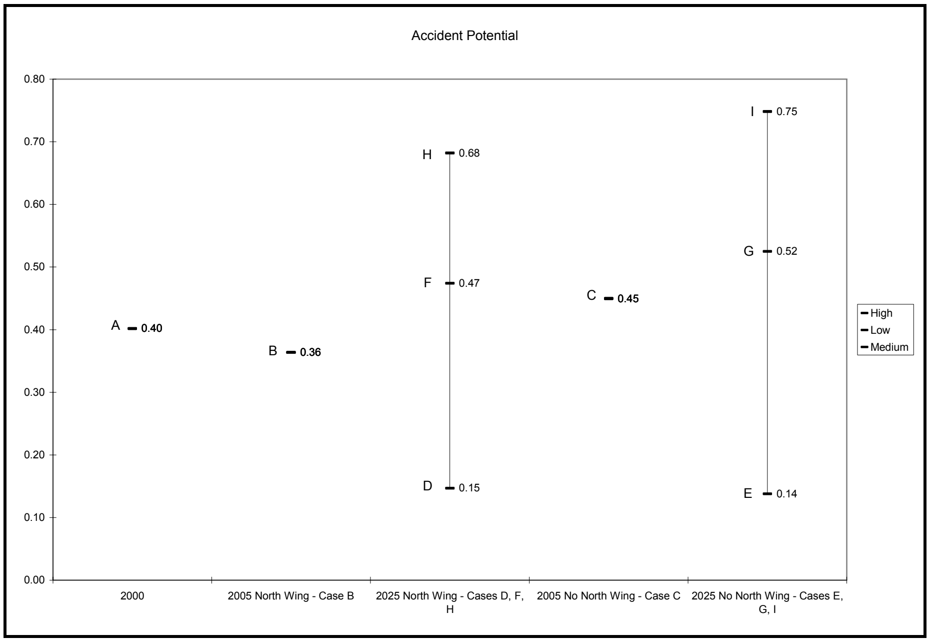


Figure 36. Accident potential in 2005 and 2025 with and without the north wing.

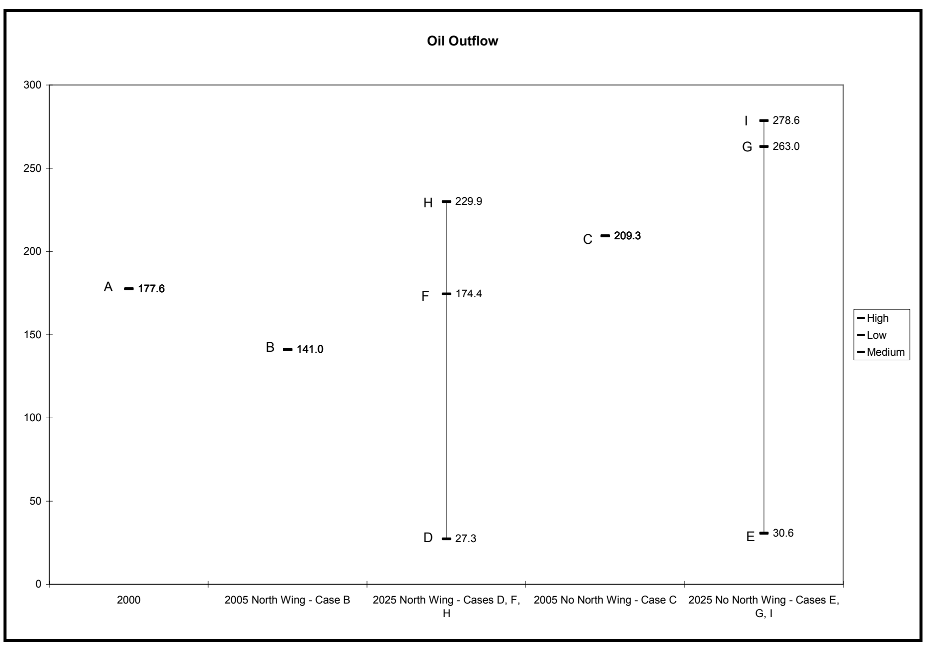


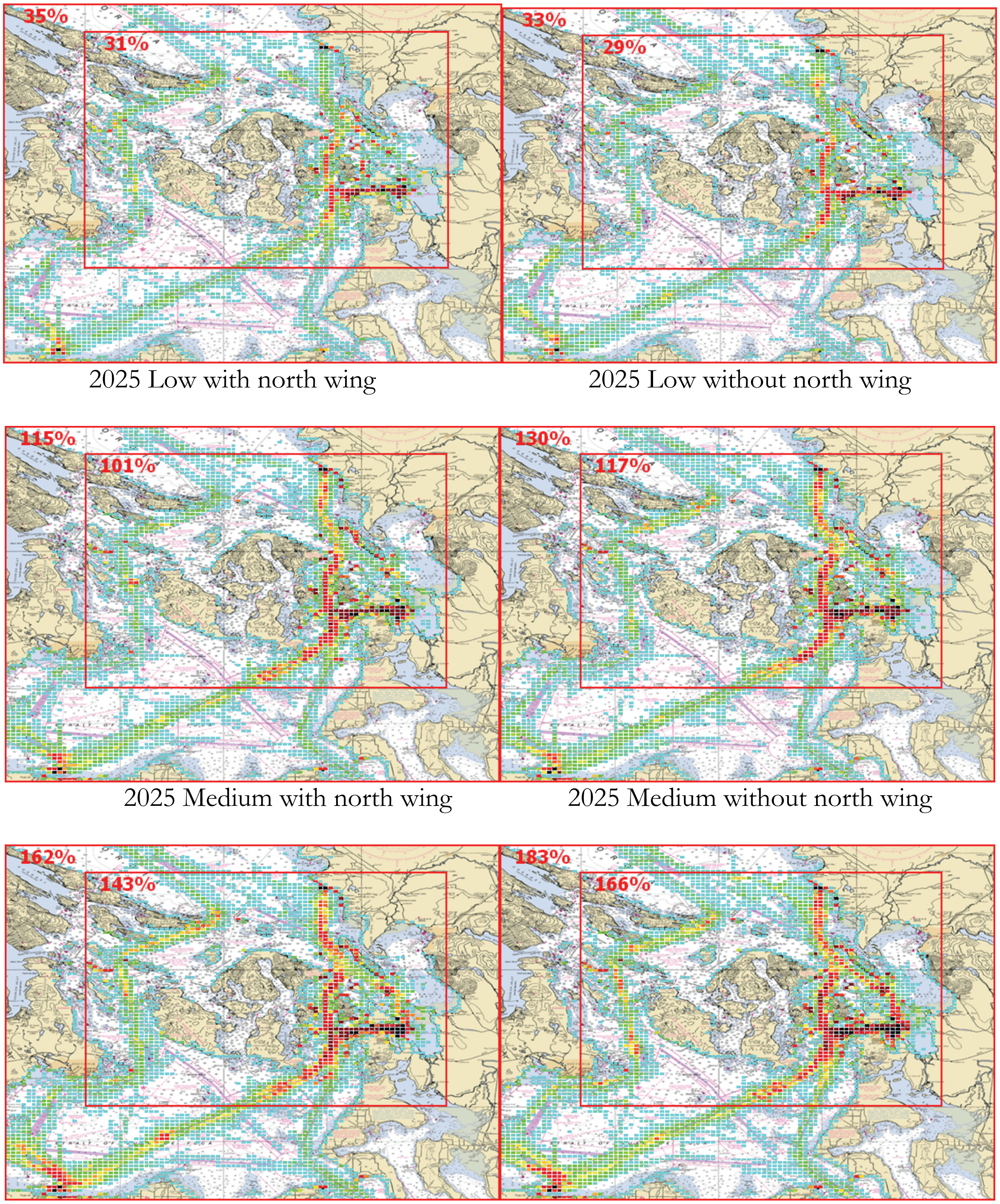
Figure 37. Oil Outflow potential in 2005 and 2025 with and without the north wing.

Instead we must consider each potential traffic level in 2025, low, medium, and high, and compare the case with the north wing operational to the case without it. Thus we must compare cases D and E for the low traffic scenario in 2025, cases F and G with the medium traffic scenario in 2025, and cases H and I with the high traffic scenario in 2025. Case F has a lower accident potential than case G, meaning that the north wing will reduce accident risk if the traffic levels are at the medium scenario in 2025. Case H has a lower accident potential than case I, meaning that the north wing will reduce accident risk if the traffic levels are at the high scenario in 2025. However, case D actually has a higher accident potential than case E, although the difference is much less than the other comparisons. Thus only if the traffic levels are at the low scenario in 2025 will the north wing lead to a slightly higher accident potential. This is because the number of BP tanker visits are slightly higher in case D than in case E, but as the total amount of traffic is low, there are no congestion problems in case E that lead to higher overall risk levels like it does in the medium and high traffic scenarios.

Figure 37 shows the same comparisons, but for oil outflow potential. We seem the same results, although there is one interesting result. The accident potential for case I was quite a bit higher than that for Case G, but there is little difference in oil outflow potential. Cases G and I each make the north wing not operational, but case G is for the medium traffic scenario in 2025 and case I is for the high traffic scenario. Recall that the increase for BP crude tankers from case B to Case G and I are 13% and 17% respectively and the increase for BP product tankers from case B to Case G and I are 13% and 90% respectively. Thus the large increase in BP product tankers leads to a large increase in accidents, but apparently not an associated large increase in oil outflow. Product tank vessels have a much lower carrying capacity than crude tankers. The small increase in BP crude tankers from case G to case I lead to a small increase in oil outflow potential.

Recall that the 17% increase in crude tankers calling at BP Cherry Point is the highest modeled future increase under when the north wing is not operational. Thus under these conditions there appears to be somewhat of an upper limit to the increase in oil outflow potential, but an increase in oil outflow of 97.5% as compared to case B levels.

Accident Potential



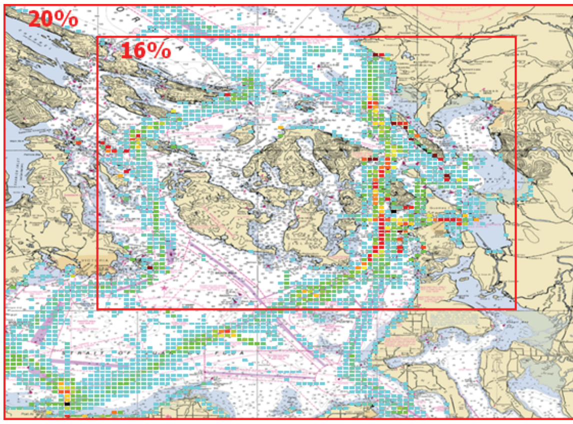
2025 Low with north wing | 2025 Low without north wing

2025 Medium with north wing | 2025 Medium without north wing

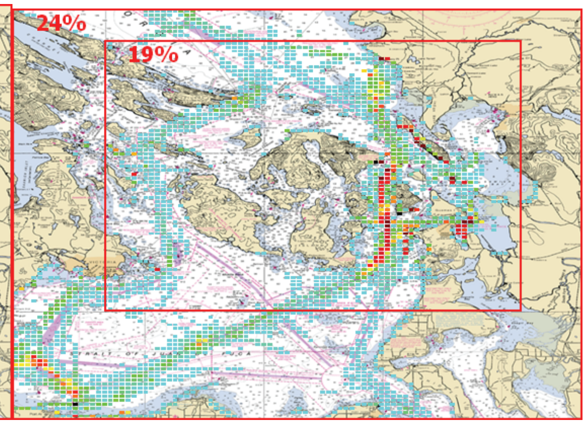
2025 High with north wing | 2025 High without north wing

Figure 38. Geographic profiles of accident potential for the low, medium, and high traffic scenarios for 2025 both with (left) and without (right) the north wing.

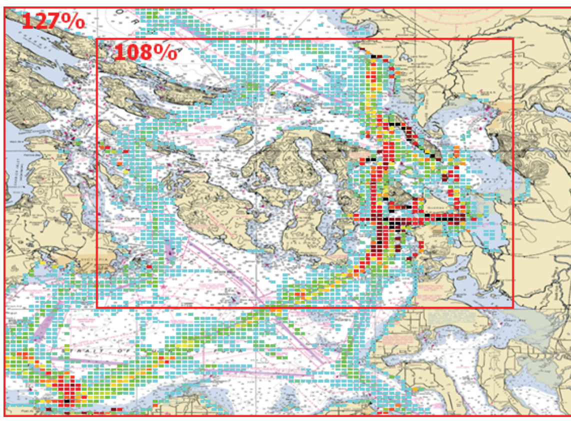
Oil Outflow Potential



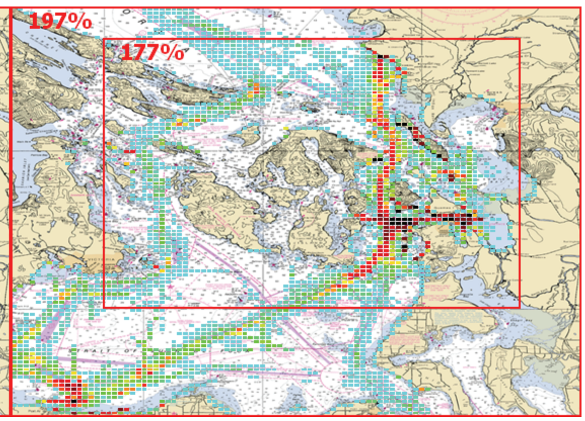
2025 Low with north wing



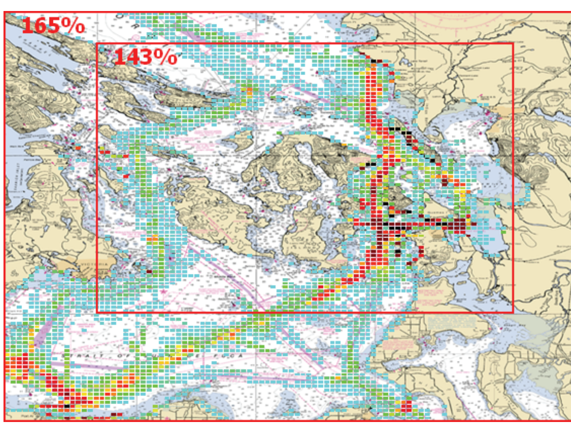
2025 Low without north wing



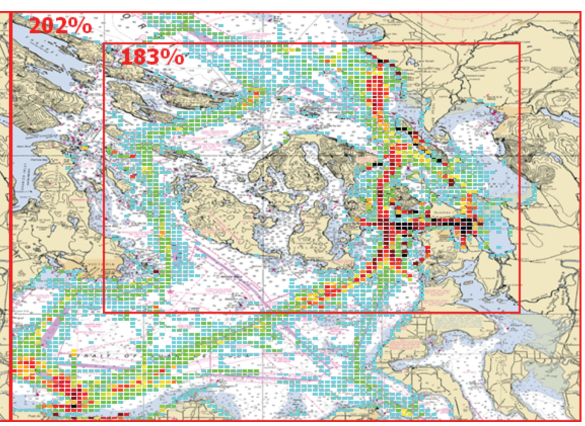
2025 Medium with north wing



2025 Medium without north wing



2025 High with north wing



2025 High without north wing

Figure 39. Geographic profiles of oil outflow potential for the low, medium, and high traffic scenarios for 2025 both with (left) and without (right) the north wing.

The geographic profiles of accident potential and oil outflow potential for cases D through I are included in Figures 38 and 39 respectively using the color scale for the case B results. While we observe lower percentages of accident potential and oil outflow potential in the 2025 low scenarios as compared to case B in the largest red rectangular area (88% and 92% for accident potential and oil outflow potential, respectively, see Figures 34 and 35), one observes dramatic increases within these areas in the medium and high traffic scenarios.

In the medium 2025 scenario we observe 15% more accident potential within the largest red rectangular area alone as within the entire VTRA study area in case B when the north wing is operational. We have about the same (1% more) accident potential within the smallest red rectangle as compared to the total accident potential within case B. Hence, take the total accident potential in case B over the entire VTRA study area (see Figure 2), multiply it by a factor of 1.15 and this gives you the accident potential in the largest rectangular area alone under this 2025 medium scenario. In the case that the north wing is not operational the accident potential in the largest red rectangle increases to 30% more than the total accident potential in case B. In terms of oil outflow potential we observe from Figure 39 27% more oil outflow within this area alone as compared to the total oil outflow in case B when the north wing dock is operational and 97% more than the total oil outflow in case B when it is not.

In the high 2025 scenario we observe 62% more accident potential within the largest red rectangular area alone as the total accident potential in case B when the north wing is operational and 43% more in the smallest red rectangular area. Hence, take the total oil outflow potential in case B over the entire VTRA study area (see Figure 2), multiply it by a factor of 1.62 and this gives you the accident potential in the largest rectangular area alone under this 2025 high scenario. In the case that the north wing is not operational this even increases to 83% more than the total accident potential in case B. In terms of oil outflow potential we observe from Figure 39 65% more oil outflow within this area alone as compared to the total oil outflow in case B when the north wing dock is operational and 102% more when it is not.

5.4. Evaluation of scope risk interventions

In this section, we will examine cases J through O for the risk interventions identified in the scope. The specifics for these three cases are included in Table 4 as a reminder. Cases B, J, L, and N are set at the 2005 traffic levels. Case H, K, M, and O are set at the high scenario for 2025 to stress test the interventions. Each of these cases includes the north wing as these interventions were tested to see if they would mitigate risk if the north wing is operational.

Table 4. The cases used to consider changes in risk from three risk interventions.

	Case	CP Traffic	Other Traffic	North Wing?	Saddlebags?	Extend Escorting?	Neah Bay?	Gate Way?
2	B	2005	2005	Yes	Yes	No	Yes	No
10	J	2005	2005	Yes	No	No	Yes	No
12	L	2005	2005	Yes	Yes	Yes	Yes	No
14	N	2005	2005	Yes	Yes	No	No	No
8	H	2025 High	2025 High	Yes	Yes	No	Yes	Yes
11	K	2025 High	2025 High	Yes	No	No	Yes	Yes
13	M	2025 High	2025 High	Yes	Yes	Yes	Yes	Yes
15	O	2025 High	2025 High	Yes	Yes	No	No	Yes

Figures 40 shows the accident potential results for cases B, J, L, and N all with 2005 traffic levels. Each accident is shown separately, but the columns are stacked so the total can also be seen. Figure 41 shows the corresponding oil outflow results. Figure 42 shows the accident potential results for cases H, K, M, and O all with traffic levels from the high scenario for 2025. Figure 43 shows the corresponding oil outflow results.

Let us first consider the use or not of Saddlebags. In cases B and J, the traffic is set to 2005 levels. Case B has BP tankers using the Saddlebags route to transit between BP Cherry Point and Anacortes, while case J has BP tankers using the Rosario/Guemes route. The total potential number of accidents shown in Figure 40 is the same in these two cases, although there are slightly more allisions in case J and slightly less in case B. When we examine the potential oil outflow in Figure 41, it is again the same in total, but case J has slightly more oil outflow from powered groundings and slightly less from collisions. The accident potential and oil outflow geographic profiles look seemingly identical for these two cases and are thus not included.

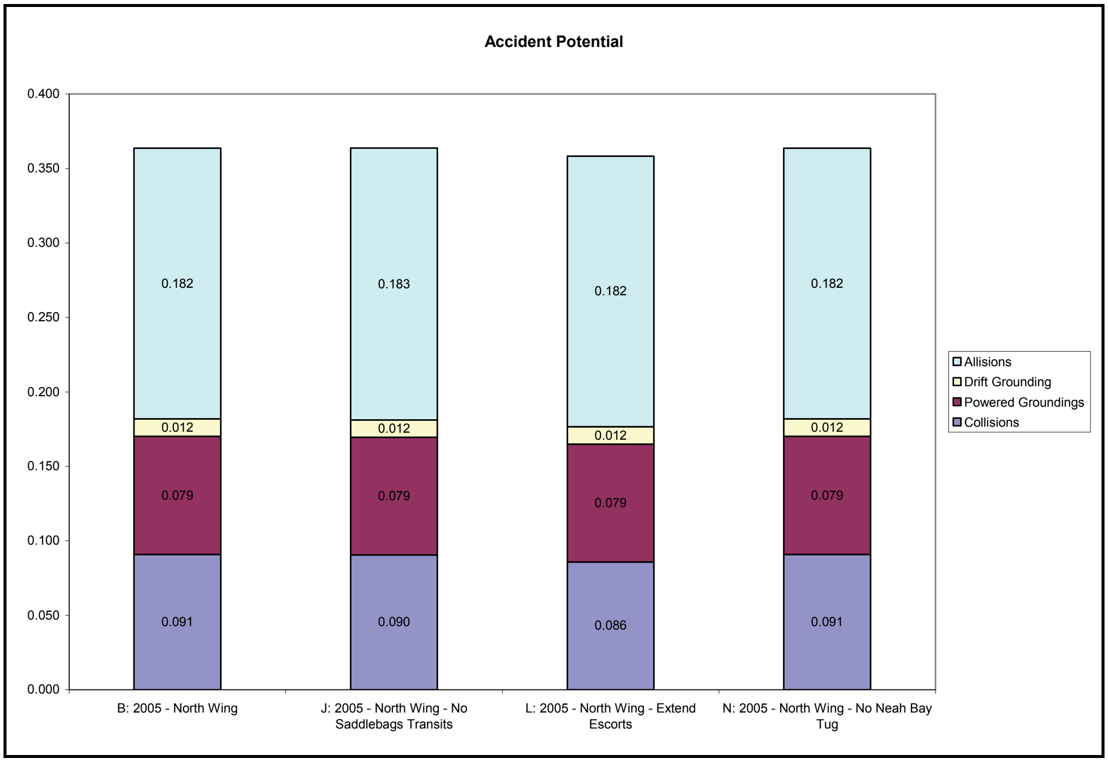


Figure 40. The accident potential by accident type for each intervention in 2005.

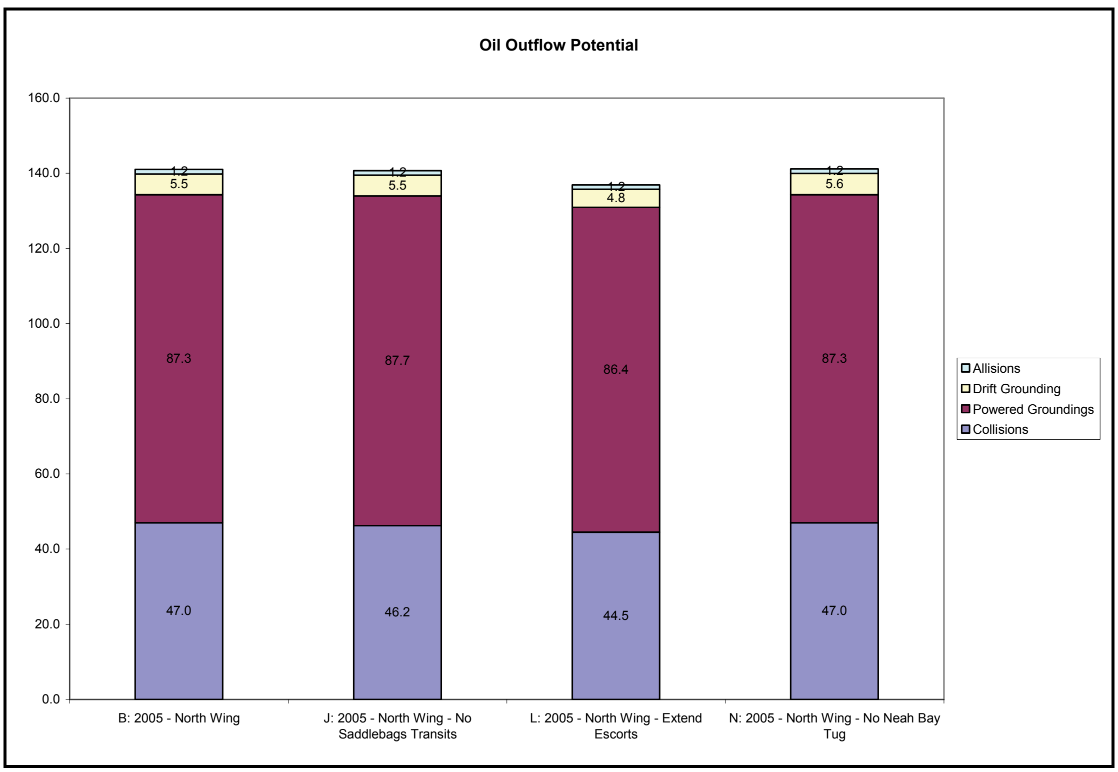


Figure 41. The oil outflow potential by accident type for each intervention in 2005.

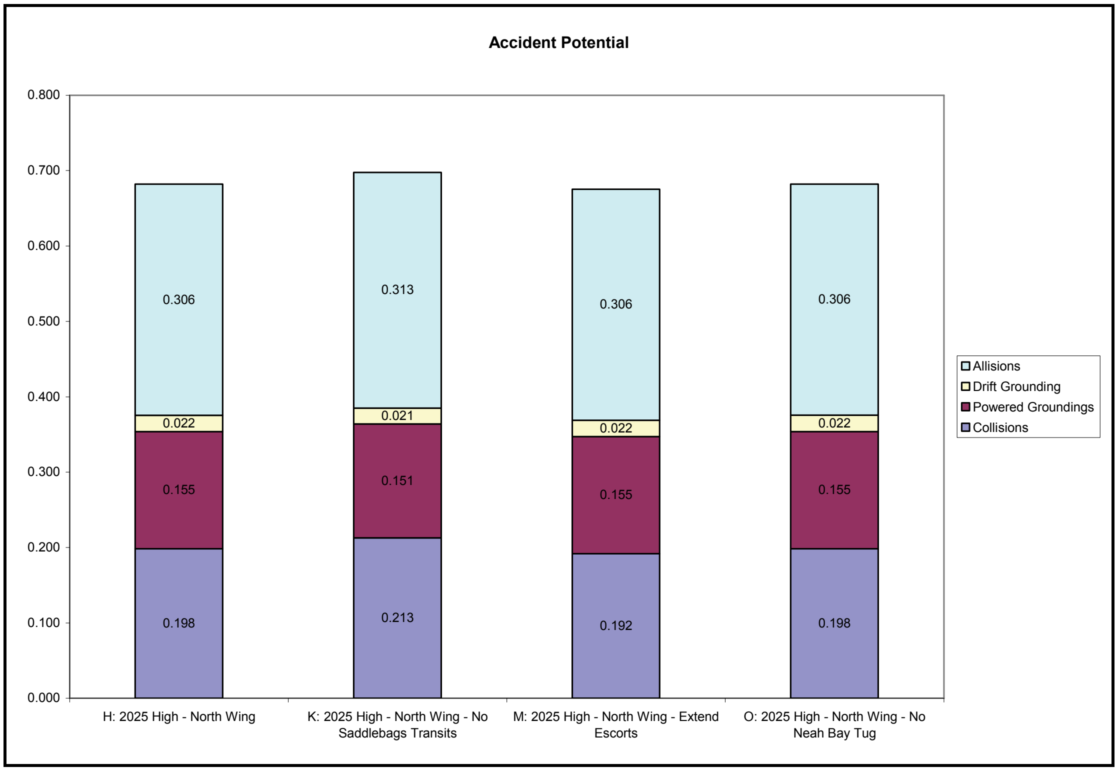


Figure 42. The accident potential by accident type for each intervention in 2025.

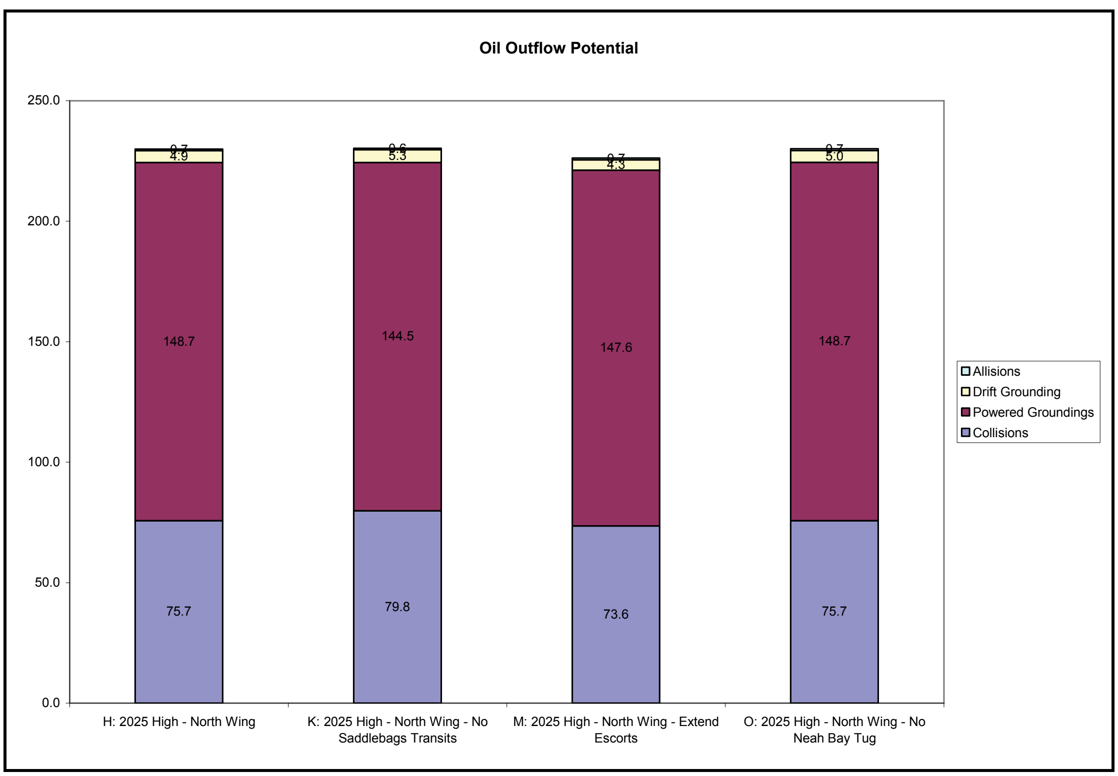


Figure 43. The oil outflow potential by accident type for each intervention in 2025.

It should be noted though that in case B a BP tanker is only diverted to Anacortes anchorage two times because the BP Cherry Point docks are full; the Cherry Point and Vendovi Island anchorages are used first. There are also very few transits between BP Cherry Point and the Shell and Tesoro docks at Anacortes. So this change does not affect many transits and, therefore, overall has a small affect. For this reason, we also assessed the affect in the high traffic scenario for 2025 to see if the difference is more pronounced at higher traffic levels. Case H is set at this traffic level and has tankers use the Saddlebags route; Case K is set at this traffic level and does not use the Saddlebags route. At these higher traffic levels, Figures 42 and 43 show that it is slightly better to use Saddlebags both in terms of accident potential and oil outflow, but this difference is still small, with a 2% decrease in accident potential and a 0.1% decrease in oil outflow. These small differences are not discernable on the geographic profiles either.

Let us now examine the concept of extending the escorting of tankers along the Straits of Juan de Fuca to Buoy J. The primary intent of a close escort by a tug is to save a tanker if it becomes disabled through total loss of either propulsion or steering. However, the crew of the escort tug also provides additional external vigilance of the tankers movements, meaning that collision and powered grounding potential is also affected. There is a limit though to the effect of extending escorts. Drift groundings only account for 3% of all accident potential in case B and drift groundings in this extended escort area only 0.1%. Oil outflow from drift groundings only account for 4% of all oil outflow potential and oil outflow from drift groundings in this extended escort area only 2%. Thus the effect that extending escort tugs through this area can have is limited in an overall sense, especially as even escort tugs do not reduce the risk of drift grounding to zero. In fact, extending the escorts reduces the drift grounding potential in the extended escort area by 17% and the oil outflow potential from drift groundings in this area by 25%. However, this corresponds to only a 1% reduction in drift grounding potential in the whole study area as there is much more drift grounding in other areas where the tankers transit closer to shore. There is a larger effect on a reduction of collision potential due to the external vigilance effect as there is more collision potential than drift grounding potential.

Overall, the effect of extending escorts is a 1.5% reduction in accident potential and a 3% reduction in oil outflow, hardly discernable in Figures 40 and 41. To give an idea of the highest reduction possible, if we assume that any time a laden tanker or ITB[9] is escorted in this area the chance of drift grounding is zero, then we would see the same 1.5% reduction in total accident potential and a 4% reduction in total oil outflow potential. At the high traffic scenario when we compare case H with the original escort system to case M with extended escorts in Figures 42 and 43, we see a 1% decrease in total accident potential and a 1.5% reduction in total oil outflow potential (with a highest possible reduction of 1% and 3%, respectively). Thus we see that drift grounding accident and oil outflow potential in the extended escorting area between Port Angeles and Buoy J is quite small compared to the total accident and oil outflow potential for the whole study area.

The final alternative to be examined is the Neah Bay tug. Before we examine this question for our cases, we must first point out the limited nature of our analysis of this problem. Firstly, the Neah Bay tug is not stationed just to assist BP tankers; it is also intended to assist non-BP tankers, bulk carriers, container vessels, and all other vessels. Secondly, our analysis has a strict geographic scope as shown in Figure 2. The Neah Bay tug can also assist vessels outside this area, along the Olympic coast and out to sea. Thus our results should be interpreted as only applying to BP vessels and only up to 8 miles outside Buoy J. Perhaps even more importantly, the VTRA geographic scope <u>includes</u> the area consisting of the approaches to and passages through the San Juan Islands and Anacortes, the Puget Sound North and South typically beyond the Neah Bay tug's operating range. Hence, if one combines this with our result that drift grounding in the extended escort area (only a portion of which is covered by the Neah Bay tug) constitutes only 0.1% of the total accident potential in case B, the effectiveness of the Neah Bay tug relative to the VTRA geographic scope and to the vessels within the VTRA scope is rather limited almost by definition.

In cases N and O, the Neah Bay tug is assumed to not be on standby. In all other cases, it is assumed to be on standby to assist a BP tanker in the case of total propulsion or steering

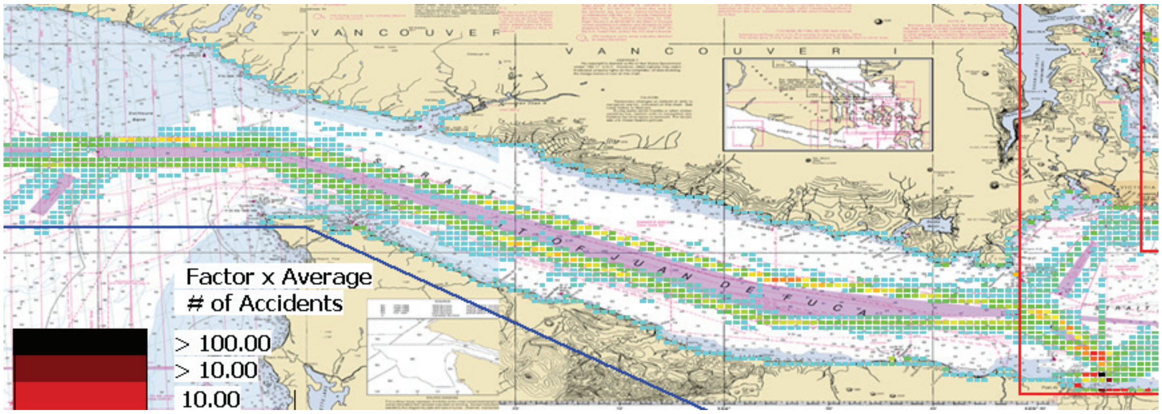
[9] Unladen tankers are not escorted, but still have some oil outflow potential as they carry fuel. ATBs are not escorted because they are smaller than the 40,000 DWT minimum for escorting.

loss. However, a standby tug is not the same as a close escort. There is no external vigilance effect as the tug is at Neah Bay, not transiting with the tanker; thus it does not affect the powered grounding or collision potential as the continuous escorting did in cases L and M. Furthermore, it cannot immediately attempt to attach a towing line to the tanker when it becomes disabled. It must transit to the tanker and has less time than a close escort to attempt a save. This also means that there is a more limited range from Neah Bay in which the tug will be able to assist a drifting vessel as compared to the extending escorting scheme of case L and M.

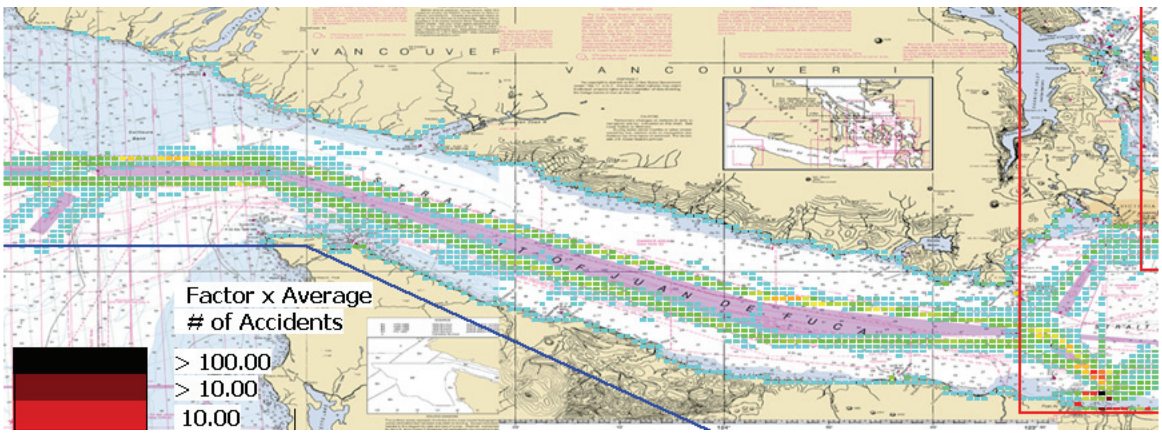
Comparing case B with the Neah Bay tug to case N without it at the 2005 traffic levels in Figures 40 and 41 and case H with the Neah Bay tug to case O without it at the high traffic scenario for 2025 in Figures 42 and 43, we see no appreciable difference in total accident potential or in total oil outflow potential for the entire VTRA study area. Again, we may test the highest possible effect of the Neah Bay tug, by assuming that if the Neah Bay tug can reach a drifting tanker, ATB, or ITB before it runs aground, then it will always save the vessel. With this assumption, we still only see a 0.03% reduction in total accident potential and a 0.8% reduction in total oil outflow potential over the entire VTRA study area. This should not be construed to mean that there could not be a significant effect of the Neah Bay tug outside our limited scope; there may be more effect for vessels outside our study area where they could drift for longer allowing the Neah Bay tug to reach them and perform a save; there could also be more effect for non-BP vessels. Finally, effectiveness analysis for the Neah Bay tug should be confined to its operating range. But for our limited scope in terms of vessels we were tasked to consider combined with the geographic scope of the VTRA study, we find that drift groundings that are within reach of the Neah Bay tug are not a major part of the total accident potential for the whole study area.

Similar conclusions can be drawn from the geographic profiles snapshots in Figure 44 and 45 where only in the middle figures an ever so slight lightening of color can be observed along the Olympia coast line and the traffic lanes.

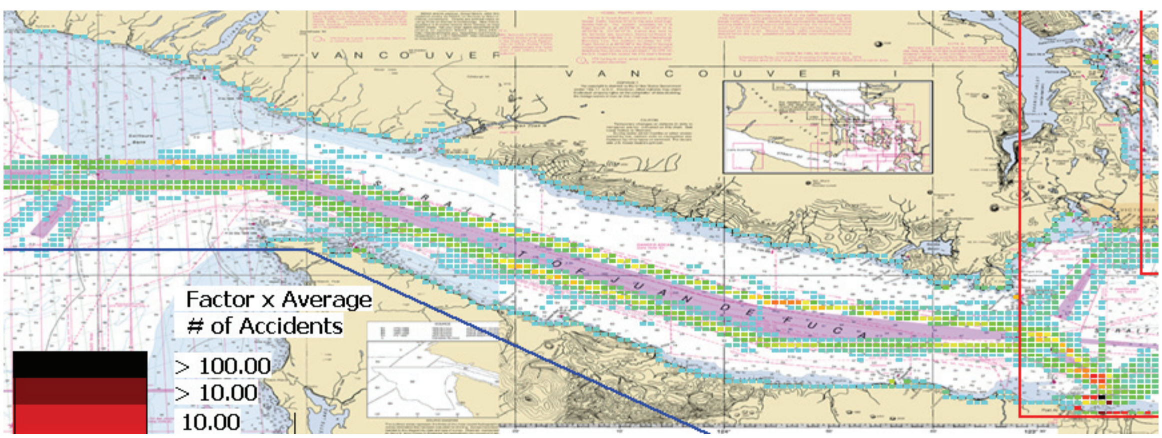
Accident Potential



Year 2005 without extended escorts, but with Neah Bay tug



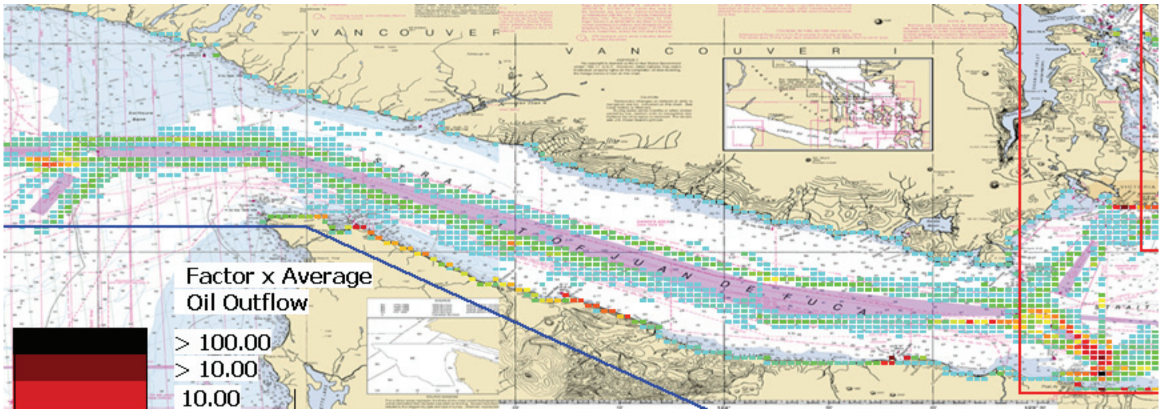
Year 2005 with extended escorts and Neah Bay tug



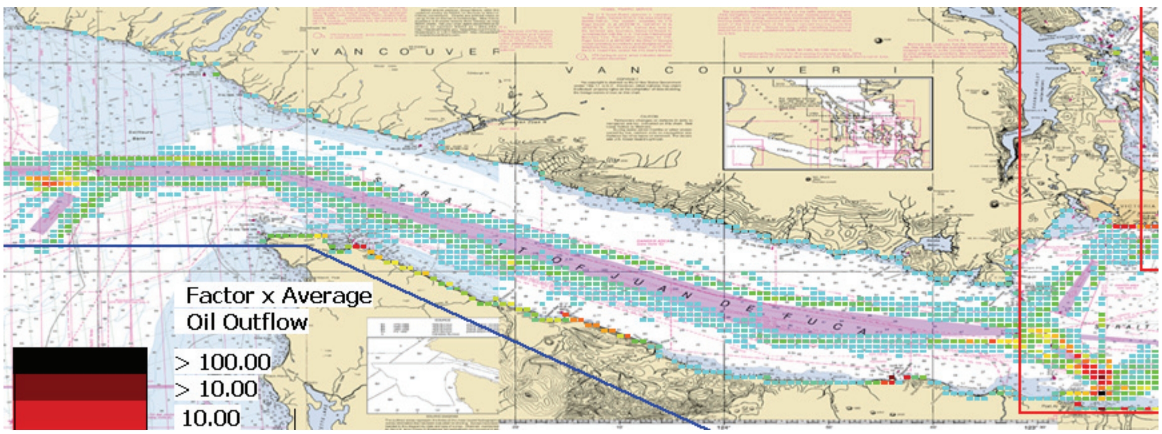
Year 2005 without extended escorts and without Neah Bay tug

Figure 44. Geographic profiles of accident potential with and without the Neah Bay tug and extended escorts.

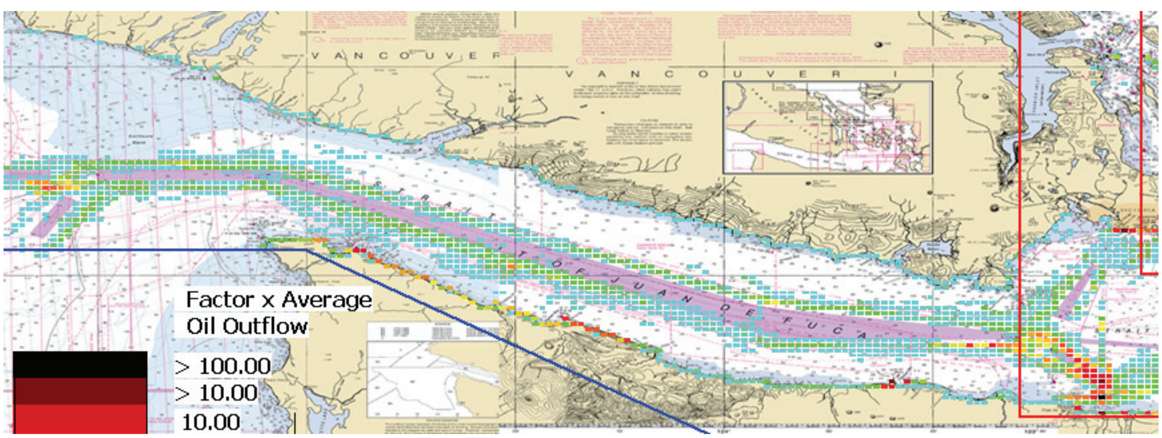
Oil Outflow Potential



Year 2005 without extended escorts, but with Neah Bay tug



Year 2005 with extended escorts and Neah Bay tug



Year 2005 without extended escorts and without Neah Bay tug

Figure 45. Geographic profiles of oil outflow potential with and without the Neah Bay tug and extended escorts.

6. Conclusions

Conclusions below pertain to accident frequencies of vessels that dock at BP Cherry Point. We refer to these vessels as BPCHPT vessels. Average annual accident frequencies have been analyzed for collisions, powered grounding, drift grounding and allisions. Conclusions are derived from the content of the main report and the technical appendices.

As per the VTRA scope, only oil outflow from BPCHPT vessels has been analyzed as well as the potential oil outflow from vessels that potentially collide with them. We refer to these later vessels as interacting vessels (IV). These oil outflows include BP persistent oil outflow (crude cargo oil and heavy fuel on board of BPCHPT vessels), BP non-persistent oil outflow (petroleum products and diesel fuel of BPCHPT vessels), IV persistent oil outflow (crude cargo oil and heavy fuel on board of IV's), IV non-persistent oil outflow (petroleum products and diesel fuel on board of IV's).

Below we provide conclusions than one may draw from our VTRA analysis results for our VTRA study area. They will be separated in four main categories. Firstly, we present conclusion regarding the risk profile of the 2005 analysis with the North Wing Dock present (which we consider to be our base case). Secondly, we discuss conclusions pertaining to a comparison of our 2000 and 2005 analysis. Thirdly, we present conclusions regarding our future scenario analysis and fourthly, we present conclusions regarding the risk intervention measures that were analyzed in this study as per our project scope.

The 2005 risk profile conclusions are further separated into system context, accident frequency and oil outflow conclusions for the entire VTRA study area. Separate accident frequency and oil outflow conclusions are also provided for an area inside and outside our largest red rectangular area as defined by, for example, Figure 28. This rectangular area approximately includes all of the areas East Strait of Juan de Fuca, Haro-Strait/Boundary pass, Rosario Strait, Cherry Point, Guemes Channel and Saddle Bag as defined by Figure 2 in this report. This rectangular area excludes approximately all of the areas West Strait of Juan de Fuca, Puget Sound North and Puget Sound South as defined by Figure 2 in this report.

2005 analysis system context conclusions - north wing operational:

- Of the total annual simulated traffic, the BP CHPT vessel traffic constitutes 1.1%. Vessels docking at the BP Cherry Point are tankers, articulated tug barges (ATB's) and integrated tug barges (ITB's).
- Of the total annual simulated traffic, all tankers, ATB's and ITB's constitute 3%.
- Of the total annual simulated deep draft traffic, the BP CHPT vessel traffic constitutes 7%.
- Of the total simulated deep draft traffic, all tankers, ATB's and ITB's constitute 16%.

2005 analysis aggregate VTRA study area conclusions - north wing operational:

- Of the total annual average accident frequency of 4/11, 1/11 (25%) to collisions, 0.079 (21.8%) can be attributed to powered groundings, 0.012 (3.2%) to drift groundings and 2/11 (50%) to allisions.
- Of the total annual average accident frequency of 4/11, about 0.320 (88%) can be attributed to the area inside our largest red square of our geographic profiles results.
- Of the total annual average oil outflow of 141.0 cubic meters analyzed, 47.0 (33.2%) to collisions, 87.3 (62%) can be attributed to powered groundings, 5.5 (3.9%) to drift groundings and 1.2 (0.9%) to allisions.
- Of the total annual average oil outflow of 141.0 cubic meters analyzed, about 92% can be attributed to the area inside our largest red square of our geographic profiles results.
- Of the total annual average oil outflow of 141.0 cubic meters analyzed, 137.4 cubic meters (97.5%) originated from BP Cherry Point vessels and 3.6 cubic meters (2.5%) from IV's. IV's include tank vessels that do not dock at Cherry Point. Hence, we may cautiously infer that of the total annual average oil outflow that we analyzed for VTRA CASE B only a small percentage can be attributed to diesel fuel of heavy fuel losses and the dominant part results from cargo losses.

2005 analysis conclusions inside largest rectangular area[10,11] - north wing operational

- Of the total annual average accident frequency of about 0.320 (88%) for this area, about 0.066 (18.2%) can be attributed to collisions, about 0.074 (20.3%) to powered groundings, about 0.011 (3.0%) cubic meters to drift groundings and about 0.169 (46.5%) to allisions.
- Of the total annual average oil outflow of about 130 (92%) cubic meters analyzed for this area, about 85.5 (61%) can be attributed to powered groundings, about 40.9 (29%) to collisions, about 3.2 (2%) to drift groundings and about 1.2 (1%) to allisions.
- Of the total annual average oil outflow of about 130 (92%) cubic meters analyzed, 127 cubic meters (90.1%) originated from BP Cherry Point vessels and 3 cubic meters (1.9%) from IV's. IV's include tank vessels that do not dock at Cherry Point.

2005 analysis conclusions outside largest rectangular area[12 13] - north wing operational

- Of the total annual average accident frequency of 0.044 (12%) for this area, 0.025 (6.7%) can be attributed to collisions, 0.006 (1.5%) to powered groundings, 0.001 (0.3%) to drift groundings and 0.013 (3.5%) to allisions.
- Of the total annual average oil outflow of about 11 (8%) cubic meters analyzed for this area, about 7 (4.8%) can be attributed collisions, 1.9 (1.4%) to powered groundings, 2.6 (1.9%) to drift groundings and almost 0 (0.0%) to allisions.
- Of the total annual average oil outflow of about 11 (8%) cubic meters analyzed for this area, 10 cubic meters (7.4%) originated from BP Cherry Point vessels and 1 cubic meters (0.6%) from IV's. IV's include tank vessels that do not dock at Cherry Point.

---

[10] Percentages of accident frequencies are of the total annual average accident frequency of 4/11 per year.

[11] Percentages of oil outflow are of the total annual average oil outflow of 141.0 cubic meters.

[12] Percentages of accident frequencies are of the total annual average accident frequency of 4/11 per year.

[13] Percentages of oil outflow are of the total annual average oil outflow of 141.0 cubic meters.

2000-2005 comparison conclusions derived from VTRA analysis results:

- If BP had restricted operations to the south wing in 2005, it could have served 96% of the BPCHPT vessels in 2005 actually served by both wings.
- In a 2005-2005 comparison, the addition of the north wing allowed the BP Cherry Point terminal to serve slightly more calling vessels, while reducing the potential for BPCHPT vessel accidents by 21% and decreasing oil outflows by 38%[14].
- With the north wing in operation in 2005 (but not in 2000), the potential for accidents involving BPCHPT vessels decreased by 10% between 2000 and 2005, and the oil outflow potential decreased by 21% between 2000 and 2005 in spite of the changes in vessel traffic during the same period.
- With only the south wing in operation in both 2000 and 2005, the potential for accidents involving BPCHPT vessels would have increased by 12% between 2000 and 2005 and the potential outflows would have increased by 18% between 2000 and 2005.

2000-2025 analysis conclusions derived from VTRA analysis results:

- At each of the low, medium, and high traffic scenarios for 2025, having the north wing leads to lower average accident potential and oil outflow potential for BPCHPT vessels than not having it.
- Assuming the north wing being operational in a 2025 analysis with medium traffic increases, results in a total annual average oil outflow of 174.4 cubic meters, which is quite similar to the 177.7 cubic meters of the previous 2000 analysis when the dock was not operational (but a reduction of 1.8%).
- Assuming the north wing being operational in a 2025 analysis with high traffic increases, results in a total annual average oil outflow of 229.9 cubic meters, compared to the 177.7 cubic meters of the same 2000 analysis when the dock was not present (an increase of 29.4%).

[14] For consistency percentages are evaluated here as percentages of 2000 levels.

- Hence, with additional traffic increases it remains possible that even with the addition of the north wing dock, oil transportation risk rises above a level previously experienced in 2000 when the north wing dock was not operational.

Risk intervention conclusions derived from VTRA analysis results:

- At the 2005 traffic levels, and not allowing the use of the Saddlebags route from BP Cherry Point to Anacortes in our maritime risk simulation model, leads to no appreciable change in either average accident potential or average oil outflow potential. In the high traffic scenario for 2025, not allowing the use of the Saddlebags route from BP Cherry Point to Anacortes in our maritime risk simulation leads to a 2% increase in average accident potential and a 0.1% increase in average oil outflow.
- At the 2005 traffic levels, extending the escorting of BP tankers and ITBs up to Buoy J in our maritime risk simulation model, leads to a decrease in both drift groundings and collisions in the extended escorting area. The overall effect is a 1.5% decrease in total average accident potential and a 3% decrease in total average oil outflow potential. In the high traffic scenario for 2025, these decreases are 1% and 1.5%, respectively.
- A restricted analysis of the risk reduction potential of the Neah Bay Tug, considering only BP tankers (about 1.1% of the total modeled traffic and about 7% of the total modeled deep draft traffic) within the VTRA study area (i.e. up to 8 miles of Buoy J where traffic separation commences and, more importantly, including the area consisting of the approaches to and passages through the San Juan Islands and Anacortes typically beyond the Neah Bay tug's operating range) our maritime risk simulation model evaluated that the Neah Bay tug has no appreciable effect on total VTRA study area average accident potential and reduces its total average oil outflow potential by 0.1%.

- In the restricted analysis performed, and assuming the Neah Bay tug has the capability to save any disabled[15] BPCHPT vessel that it could get to in time, regardless of the situational context, it was shown that the Neah Bay tug could reduce total average VTRA study area accident potential by 0.03% and total average VTRA study area oil outflow potential by 0.75%.

Quantitative results in our study are presented as average point estimates commonly used for the evaluation of alternatives in a decision analysis context. These are derived from uncertain quantities as described in each step of the analysis as described in this report and its technical appendices. As with any risk assessment model, our model too represents an abstraction of reality and its results must be interpreted with care and with awareness of scoping, data limitations and modeling assumptions. In particular, the forecasts of maritime traffic, accident frequencies, and oil outflows in 2025 must be treated with care.

One primary limitation of the VTRA study is that, due to scoping constraints, the results reflect only on a small percentage (1.1%) of the vessel traffic described in the maritime simulation. If risk interventions have an appreciable effect beyond the BPCHPT vessels analyzed in this study, they should also be tested against this larger class of vessels to determine their effects on system wide accident frequencies and oil outflows. For example, a risk intervention that reduces accident frequency and or oil outflow of BP Cherry Point vessels, but results in a larger potential increase of accident frequency and/or oil outflows from the other traffic should not be implemented. Conversely, risk mitigation measures that have little or no impact on the BP Cherry Point vessels accident frequency or oil outflow may in fact significantly reduce risk to other vessels.

As such, a full evaluation of the risk reduction potential of the Neah Bay tug was not within the scope of the VTRA, as the analysis was restricted to BPCHPT vessels in the VTRA geographic scope. A full evaluation of the risk reduction potential of the Neah Bay tug requires (1) inclusion of all non-BP vessel traffic within the VTRA study area in its

[15] Our definition of a disabled BPCHPT vessel here is one that experienced either a steering or propulsion failure.

effectiveness analysis and (2) inclusion of all vessel traffic beyond the boundaries of our VTRA study area (i.e. beyond the beginning of the Traffic Separation Scheme approximately 8 nautical miles beyond Buoy J offshore of Cape Flattery), but both limited to the tug's operating range. Neither was part of the scope of the VTRA study.

REFERENCES

M. Grabowski, J.R.W. Merrick , J.R. Harrald, T.A. Mazzuchi, and J.R. van Dorp (2000).“Risk Modeling in Distributed, Large Scale Systems”, *IEEE Transactions on Systems, Man, Cybernetics – PART A: Systems and Humans*, Vol. 30, No. 6, pp. 651-660.

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

R. Sanderson (1982). *Meteorology at Sea*. Stanford Maritime Limited.

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.

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.

SUB-APPENDIX:

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.

The Prince William Sound Risk Assessment

Jason R. W. Merrick
Department of Statistical Sciences and Operations Research, Virginia Commonwealth University, PO Box 843083, 1001 West Main Street, Richmond, Virginia 23284

J. René van Dorp • Thomas Mazzuchi • John R. Harrald • John E. Spahn
Department of Engineering Management and Systems Engineering, George Washington University, 1776 G Street NW, Suite 110, Washington, DC 20052

Martha Grabowski
Business Department, Le Moyne College, and Department of Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute, 5555 Mount Pleasant Drive, Cazenovia, New York 13035
jrmerric@vcu.edu • dorpjr@seas.gwu.edu • mazzuchi@seas.gwu.edu • harrald@seas.gwu.edu • grabowsk@lemoyne.edu

This paper was refereed.

After the grounding of the Exxon Valdez and its subsequent oil spill, all parties with interests in Prince William Sound (PWS) were eager to prevent another major pollution event. While they implemented several measures to reduce the risk of an oil spill, the stakeholders disagreed about the effectiveness of these measures and the potential effectiveness of further proposed measures. They formed a steering committee to represent all the major stakeholders in the oil industry, in the government, in local industry, and among the local citizens. The steering committee hired a consultant team, which created a detailed model of the PWS system, integrating system simulation, data analysis, and expert judgment. The model was capable of assessing the current risk of accidents involving oil tankers operating in the PWS and of evaluating measures aimed at reducing this risk. The risk model showed that actions taken prior to the study had reduced the risk of oil spill by 75 percent, and it identified measures estimated to reduce the accident frequency by an additional 68 percent, including improving the safety-management systems of the oil companies and stationing an enhanced-capability tug, called the Gulf Service, at Hinchinbrook Entrance. In all, various stakeholders made multimillion dollar investments to reduce the risk of further oil spills based on the results of the risk assessment.
(*Decision analysis: risk. Industries: petroleum, transportation. Reliability: system safety.*)

On March 24, 1989, the Exxon Valdez ran aground on Bligh Reef, spilling an estimated 11 million gallons of crude oil into Prince William Sound, Alaska. The oil spill (Figure 1) spread rapidly, affecting more than 1,500 miles of shoreline. The spill had both immediate and lingering effects on fish and wildlife resources and on the lives of people in coastal communities. The cost to Exxon Corporation for cleanup operations was estimated to be $2.2 billion (Harrald et al. 1990).

After the accident, all parties with interests in Prince William Sound (PWS) agreed to work to prevent such an event from happening again. They implemented several ideas for reducing the risk of an oil spill. They introduced weather-based closure restrictions that stopped all transits through Valdez Narrows and Hinchinbrook Entrance (Figure 2) during periods of high winds. The US Coast Guard designated Valdez Narrows a special navigation zone by restricting passage through the narrows to one way for deep-draft traffic, including oil tankers. The oil companies introduced escort tugs to accompany oil-laden tankers in their

0092-2102/02/3206/0025$05.00
1526-551X electronic ISSN

Figure 1: The stricken Exxon Valdez spilled oil into Prince William Sound, Alaska, affecting over 1,500 miles of shoreline.

transit out of PWS. These tugs were to assist a tanker if it had propulsion or steering failures, attaching lines to the disabled tanker and holding it fast, thus preventing grounding accidents. The Oil Pollution Act (1990) stated that two escort tugs should accompany each oil-laden tanker; depending on the wind conditions and the size of the tanker, three tugs were sometimes used.

In early 1995, questions arose concerning the effectiveness and benefits of existing and proposed risk-intervention measures. The PWS shipping companies (ARCO Marine Inc., BP Oil Shipping Company, USA, Chevron Shipping Company, SeaRiver Maritime Inc., and Tesoro Alaska Petroleum Company) concluded that they needed a comprehensive risk assessment to evaluate all proposals. They formed a steering committee along with the PWS Regional Citizens Advisory Committee (RCAC) ⟨http://www.pwsrcac.org⟩, the Alaska Department of Environmental Conservation (ADEC) ⟨http://www.state.ak.us/dec/⟩, and the US Coast Guard (USCG). The members consisted of presidents of oil-shipping companies, local fisherman and environmentalists representing the RCAC, senior representatives of ADEC, and the USCG captain of the port for Valdez. Although the members of the group had different perspectives on the operation of the oil-transportation system, the committee captured the substantive expertise of the PWS oil-transportation and ecosystem.

By forming the steering committee, the PWS community formalized its preference for a collaborative analysis approach rather than an adversarial one (Charnley 2000). Up to this point, the adversarial approach had prevailed in PWS risk and safety studies, pitting expert against expert. The adversarial approach often leads to a lack of trust in the decision-making process and subsequently may hamper the implementation of regulations and procedures aimed at reducing risk. Many see lack of trust as the major reason for the failure of sophisticated technological risk assessments to influence public policy in the nuclear-power arena (Slovic 1993).

The steering committee decided to fund a risk-assessment effort for the PWS oil-transportation system and engaged a consultant team from George Washington University (GWU), Rennselaer Polytechnic Institute (RPI), and Det Norske Veritas (DNV). The committee stipulated the objectives of the risk-assessment effort:

—to identify and evaluate the risks of oil transportation in PWS,

—to identify, evaluate, and rank proposed risk-reduction measures, and

—to develop a risk-management plan and risk-management tools that could be used to support a risk-management program.

In this paper, we present an overview of the modeling and analysis we used in addressing the first two objectives and discuss the effect of the analysis on the third objective and the implementation of the recommendations.

Risk Assessment and Management in Maritime Transportation

The National Research Council identified the assessment and management of risk in maritime transportation as an important problem domain (NRC 1986, 1991, 1994, 2000). In earlier work, researchers concentrated on assessing the safety of individual vessels or marine structures, such as nuclear-powered vessels (Pravda and Lightner 1966), vessels transporting liquefied natural gas (Stiehl 1977), and offshore oil and gas platforms (Paté-Cornell 1990). The USCG tried to prioritize federal spending to improve port infrastructures using a classical statistical analysis of nationwide accident data (USCG 1973, Maio et al. 1991). More recently, researchers have used probabilistic risk assessment (PRA) (US Nuclear Regulatory Commission 1975) in the maritime domain (Hara and Nakamura 1995, Roeleven et al. 1995, Kite-Powell et al. 1996, Slob 1998, Fowler and Sorgard 2000, Trbojevic and Carr 2000, Wang 2000, Guedes Soares and Teixeira 2001) by examining risk in the context of maritime transportation systems (NRC 2000).

In a maritime transportation system (MTS), traffic patterns change over time in a complex manner. Researchers have used system simulation as a modeling tool to assess MTS service levels (Andrews et al. 1996), to perform logistical analysis (Golkar et al. 1998), and to facilitate the design of ports (Ryan 1998). The dynamic nature of traffic patterns and other situational variables, such as wind, visibility, and ice conditions, mean that risk levels change over time. The PWS risk assessment differs from previous maritime risk assessments in capturing the dynamic nature of risk by integrating system simulation (Banks et al. 2000) with available techniques in the field of probabilistic risk assessment (Bedford and Cooke 2001) and expert judgment elicitation (Cooke 1991).

Defining Risk

Lowrance (1976) defined risk as a measure of the probability and severity of the consequences of undesirable events. In the PWS risk assessment, we defined the undesirable events to be accidents involving oil tankers, specifically the following:

—Collisions: An underway tanker colliding with or striking another underway vessel as a result of human error or mechanical failure and lack of vigilance (intervessel collision) or striking a floating object, for example, ice;

—Drift groundings: A drifting tanker out of control because of a propulsion or steering failure making contact with the shore or bottom;

—Powered groundings: An underway tanker under power making contact with the shore or bottom because of navigational error or steering failure and lack of vigilance;

—Foundering: A tanker sinking because of water ingress or loss of stability;

—Fire or explosion: A fire occurring in the machinery, hotel, navigational, or cargo space of a tanker or an explosion occurring in the machinery or cargo spaces; and

—Structural failure: The hull or frame cracking or eroding seriously enough to affect the structural integrity of the tanker.

The consequence of interest was oil outflow into PWS. The initial measure the steering committee wanted was the expected volume of oil outflow per year for each accident type and specified locations. However, after further discussion, it decided that any accident involving an oil tanker was an undesirable event, and thus the focus shifted to the expected number of accidents per year again broken down by accident type and location. We defined boundaries for seven locations to use in the study (Figure 2).

The basic technique used in the PWS risk assessment is probabilistic risk analysis (PRA) (Bedford and Cooke 2001). In performing a PRA, one identifies the series of events leading to an accident, estimates the probabilities of these events, and evaluates the consequences of the accident. Garrick (1984) noted that an accident is not a single event but the culmination of a series of events. A triggering incident is defined to be the immediate precursor of an accident. In the PWS risk assessment, we separated triggering incidents into mechanical failures and human errors. The mechanical failures considered to be triggering incidents were propulsion failures, steering failures, electrical power failures, and hull failures. The classifications of human errors used were diminished ability; hazardous shipboard environment; lack of knowledge, skills, or experience; poor management practices; and faulty perceptions or understanding. We based these on current USCG classifications.

We constructed an accident probability model using the relationships between the vessel's operating environment, triggering incidents, and accidents (Roeleven et al. 1995). The combination of organizational and situational factors that describes the state of the system in which an accident may occur is termed an opportunity for incident (OFI). We based our accident model on the following conditional probabilities:

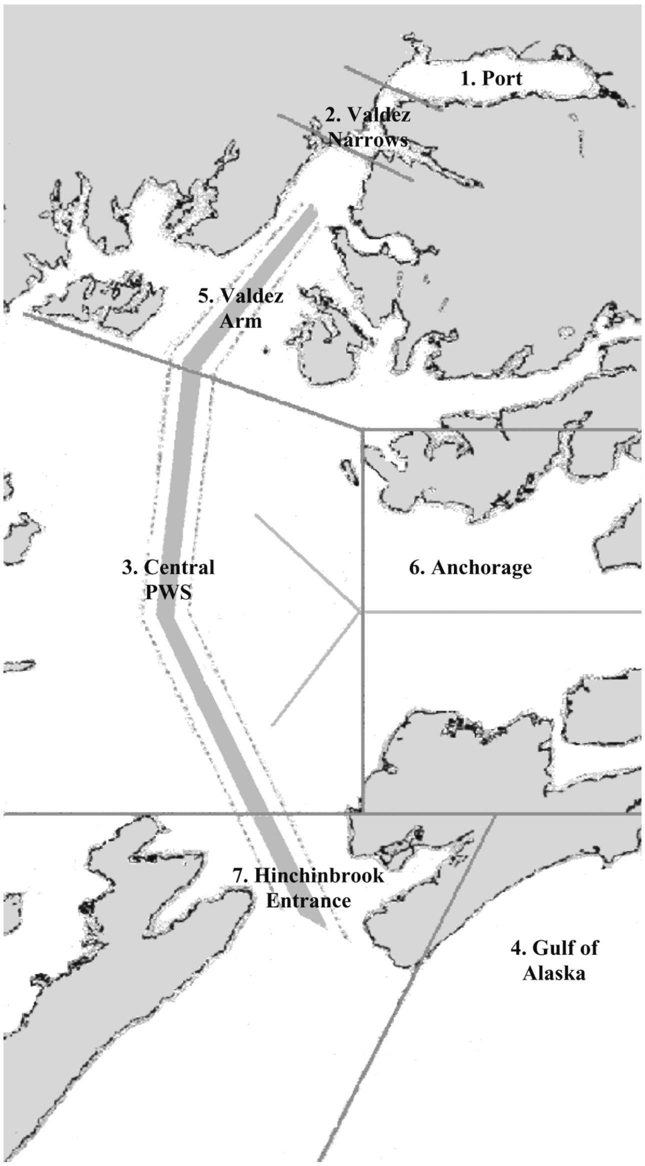


Figure 2: We divided Prince William Sound into seven locations for reporting risk.

—P(OFI): the probability that a particular system state occurs,

—P(Incident | OFI): the probability that a triggering incident occurs in this system state, and

—P(Accident | Incident, OFI): the probability that an accident occurs given that a triggering incident has occurred in this system state.

Once one has specified these probabilities, one can find the probability of an accident occurring in the system by summing the product of the conditional probabilities over all types of accidents and triggering incidents and all combinations of organizational and situational factors according to the law of total probability. Thus to perform an assessment of the risk of an accident using this model, one must determine an operational definition of an OFI and then estimate each of the terms in the probability model. Harrald et al. (1998) discuss the operational definition of an OFI in the PWS risk assessment.

The System Risk-Simulation Model

The first term to estimate is the frequency of occurrence of each combination of organizational or situational factors, that is, each OFI. Although data is collected on vessel arrivals and environmental conditions, the combinations of these events are not. Traffic rules, such as a one-way zone, mean that the movements of vessels are dependent, while weather-based closure restrictions cause dependence between vessel movements and environmental conditions. A discrete-event simulation of the system captures the complex dynamic nature of the system and accurately models the interactions between the vessels and their environment.

We created the simulation model using operational data, such as vessel-type and vessel-movement data from the USCG vessel traffic service, tanker arrival and departure information from the ship escort/response vessel system (SERVS), and publicly available data, such as meteorological data from the National Oceanographic and Atmospheric Administration weather buoys. More difficult to obtain were data on open fishing times, locations, and durations, which required local community surveys. Based on the data, we developed traffic-arrival models and weather models. In addition, because all deep-draft vessels transiting PWS must participate in the USCG vessel traffic service and follow a defined set of traffic rules, such as weather-based closure restrictions, one-way zones, the tug escort scheme, and docking procedures, we programmed these rules into the simulation.

We used the simulation as an event counter, that is,

Location	Central Sound	Likelihood of Collision		Location
Traffic proximity	Vessels 2 to 10 miles	9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9		Traffic proximity
Traffic type	Tug with tow			Traffic type
Tanker size and direction	Inbound more than 150DWT			Tanker size and direction
Escort vessels	Two or more			Escort vessels
Wind speed	More than 45			Wind speed
Wind direction	Perpendicular/On shore			Wind direction
Visibility	Greater than 1/2 mile			Visibility
Ice conditions	Bergy bits within a mile		No bergy bits in a mile	Ice conditions

Table 1: We elicited judgments from the substantive experts using pairwise comparison questionnaires in which we defined a given scenario and varied only one attribute, in this example changing whether there is ice in the traffic lanes.

we used it to count the number of occurrences of individual OFIs throughout PWS for a given time period. The simulation calculated the state of the system once every five minutes based upon the traffic arrivals, the weather, and the previous state of the system. We ran the simulation for 25-years of simulation time and, for each five-minute period, tabulated the OFIs that occurred, and thus determined OFI frequencies (Merrick et al. 2000).

We estimated the two levels of conditional probability of triggering incidents and accidents. The preferred method for estimating these probabilities is through data. The steering committee required that we use only PWS specific data in the risk assessment, rather than worldwide accident data that might not be representative. Each of the PWS shipping companies supplied proprietary mechanical-failure data. However, at the time we could obtain no reliable PWS human-error data in the maritime domain, and we could obtain very little from near-miss reports (Harrald et al. 1998). Large databases of local accident data were not available for standard statistical analysis of the organizational and situational factors that could affect risk. Cooke (1991) cites the use of expert judgment in areas as diverse as aerospace programs, military intelligence, nuclear engineering, and weather forecasting. We used expert judgment to assess relative conditional probabilities and data to calibrate these relative probabilities.

Using the log-linear accident probability model (Roeleven et al. 1995), we obtained relative conditional probabilities through a regression analysis of pairwise comparison surveys (Bradley and Terry 1952) constructed for the pilots, captains, and chief engineers with operational experience in PWS. PWS oil-shipping companies, SERVS, and regional representatives on the PWS steering committee made these substantive experts available for elicitation sessions. An example of the type of questions posed is the following taken from the expert-judgment questionnaire for collisions given that a propulsion failure has occurred (Table 1). In each situation, there is an inbound tanker, greater than 150,000 DWT in size, which has just experienced a propulsion failure. It is within two to 10 miles of a tug with tow in winds over 45 MPH blowing on shore to the closest shore point with visibility greater than half a mile in Central PWS. The only difference between the two situations is that the first situation includes an ice flow in the traffic lane, while the second does not. We ask the expert to picture the two situations, to determine which situation is more likely to result in a collision, and to indicate his or her sense of magnitude in the choice through a nine-point scale, with one indicating equally likely (Saaty 1977).

For each question, we changed only one attribute so that the experts could estimate the difference in risk between the two situations. The experts could answer a book of 120 questions in one to one-and-a-half hours. We put the questions in the books in random order and statistically tested the results to ensure nonrandom responses and to minimize response bias. All participants had very extensive knowledge with at least 20 years of experience at sea. We treated the expert responses as ratios of the probabilities of an accident in each scenario. We estimated the parameters of the accident probability model using statistical regression and calibrated the model to available data. The *Prince William Sound Risk Assessment Study Final Report* (PWS Steering Committee 1996) contains specific details of the development of the simulation model, the design and analysis of the expert-judgment questionnaires, and the integration of the simulation model and the accident probability model.

The integrated system risk-simulation model was capable of assessing the current risk of accidents involving oil tankers operating in PWS and of evaluating risk-intervention measures. We also implemented an oil-outflow model, created by DNV, in the system risk-simulation program. The program displayed risk in PWS dynamically (Figure 3) and we could interrogate it to determine the expected frequencies of accidents or the expected oil outflow per year broken down by accident type, location, and any of the organizational or situational factors.

Results of the Risk Assessment

The steering committee's first objective was to identify and evaluate the risks of oil transportation in PWS. We chose accident scenarios as the method of reporting, defining an accident scenario to be an accident type in a given location. We programmed the simulation to represent the shipping fleet, traffic rules, and operating procedures in place in 1996, the year we performed the study. We ran the simulation program for 25 years (simulation time) and estimated the expected frequency of accidents. We broke the frequencies down by location and accident type to obtain the accident-scenario results. As the primary interest was accident scenarios with the highest expected frequencies, we reported the results by sorting the accident scenarios from highest to lowest (Figure 4).

Before the risk assessment, people in PWS commonly believed that the most likely accident scenario was a drift or powered grounding in the Valdez Narrows or Hinchinbrook Entrance. However, we showed that the first seven accident scenarios accounted for 80 percent of the total expected frequency of accidents, with 60 percent coming from collisions in the port, in the Valdez Narrows, and in the Valdez Arm. We performed a further analysis to find the primary cause of these accidents. We found that the primary risk was collisions with fishing vessels that operate in large numbers in these locations during fishing openers. Although they introduce a relatively high risk of collision, few fishing vessels are large enough to penetrate the hull of a tanker. Thus the expected oil outflow from these events was low. The perceived high-risk scenarios of drift or powered groundings contributed approximately 15 percent of the expected frequency of accidents.

Integrating the oil-outflow model with the estimated frequencies of accident scenarios allowed us to estimate the expected volume of oil outflow as a measure of risk, again reported from highest to lowest (Figure 5). We discovered a surprising result using this metric. Potential collisions of outbound tankers with inbound SERVS' tugs (returning from escort duty) are a large contributor to the total expected oil outflow. Escort tugs leaving port with a tanker are intended to save the tanker in case of a propulsion or steering failure, but on their return from escort they introduce a risk of collision and can cause enough damage to tankers to spill oil. Less suprising, however, was the confirmation of the risk of drift or powered groundings in the Valdez Narrows or Hinchinbrook Entrance.

The steering committee's second objective was to identify, evaluate, and rank proposed risk-intervention measures. We developed a set of risk-intervention measures for evaluation in consultation with the PWS steering committee. We classified risk-interventions in terms of their effect on modeling parameters and analyzed them accordingly. The modeling required was extensive, but because of the level of granularity incorporated in the system risk-simulation model, we could change parameters of the accident probability model or simulation code to reflect the effects of risk-intervention measures. By stripping away previously implemented risk-intervention measures, we estimated the risk prior to the Exxon Valdez accident. Comparing this risk to the baseline case, representing the PWS system during the study period, we estimated that the accident frequency had been reduced 75 percent since the Exxon Valdez accident.

We identified further effective risk-intervention mea-

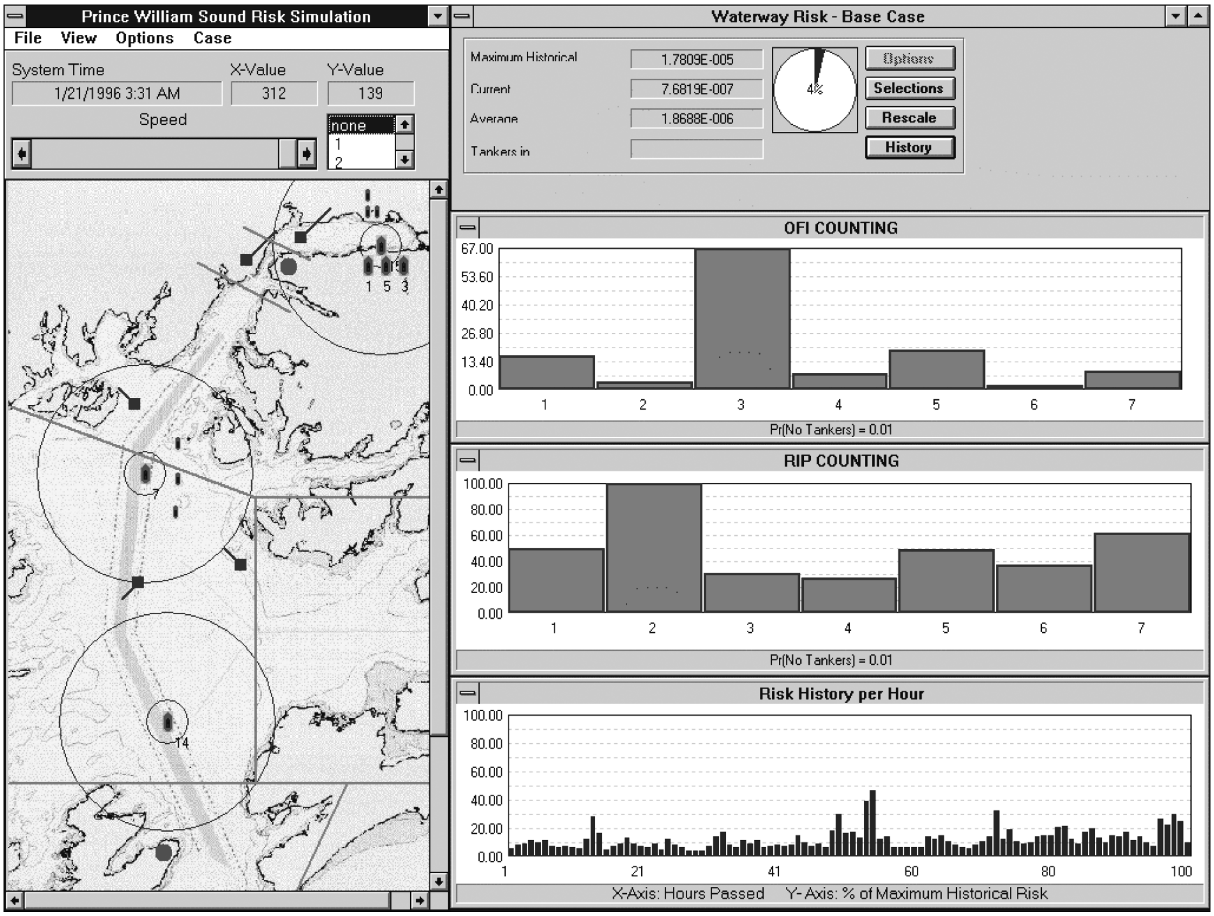


Figure 3: We created the system risk-simulation program to perform the analysis and demonstrate the results to the steering committee. On the left is a display of the dynamic behavior of the Prince William Sound marine transportation system including traffic patterns and environmental conditions, such as wind speed and direction. On the right, the analysis shown is broken into seven locations (Figure 2), with estimates of the probability of an opportunity for an incident, the probability of an accident given such an opportunity, and finally the dynamic variation in the expected frequency of accidents for the whole region.

sures (Figure 6). Under the current system, interactions with fishing vessels and escort tugs were significant contributors to the overall risk. We developed rules to reduce the number of these interactions in cooperation with the steering committee and programmed them into the simulation. We demonstrated that modifying the escort scheme to reduce interactions with tankers and managing the interactions of fishing vessels and tankers led to a major reduction in risk. The model also indicated that improving human and organizational performance through the International Safety Management (ISM) program would further reduce risk. We estimated the reduction in risk obtained by reducing the frequency of human errors in the accident probability model, with the reduction being estimated by personnel from DNV with experience in implementing the ISM program. We showed that some proposed risk-intervention measures increase risk, for example, we showed that additional weather-based closure restrictions would increase traffic congestion.

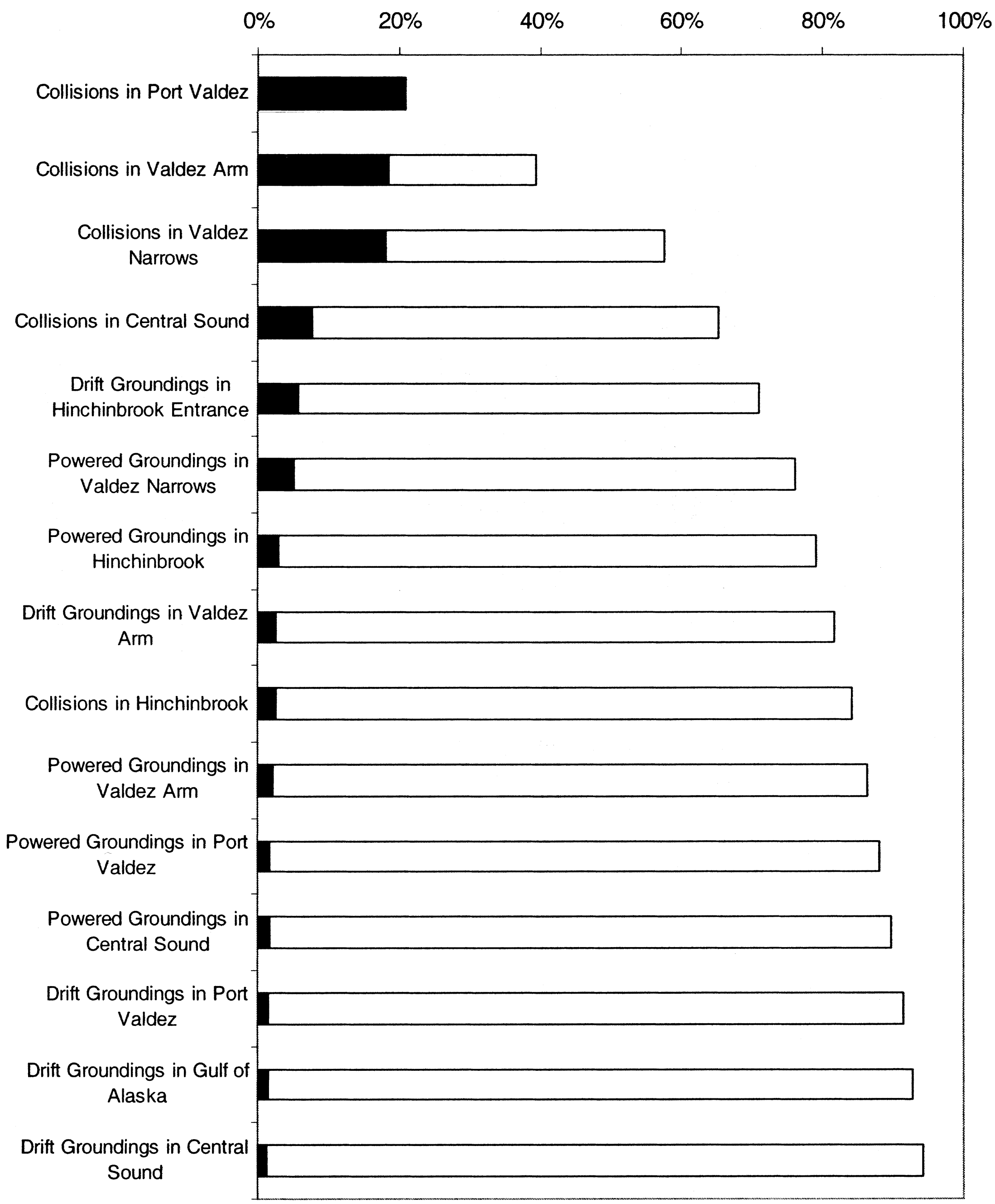


Figure 4: We sorted the combinations of accident types and locations by their expected frequency (dark bars). The cumulative percentage of the total expected frequency up to each such combination (white bars) is indicated by the total height of each bar. For example, we found that the first seven accident scenarios account for 80 percent of the total expected frequency of accidents.

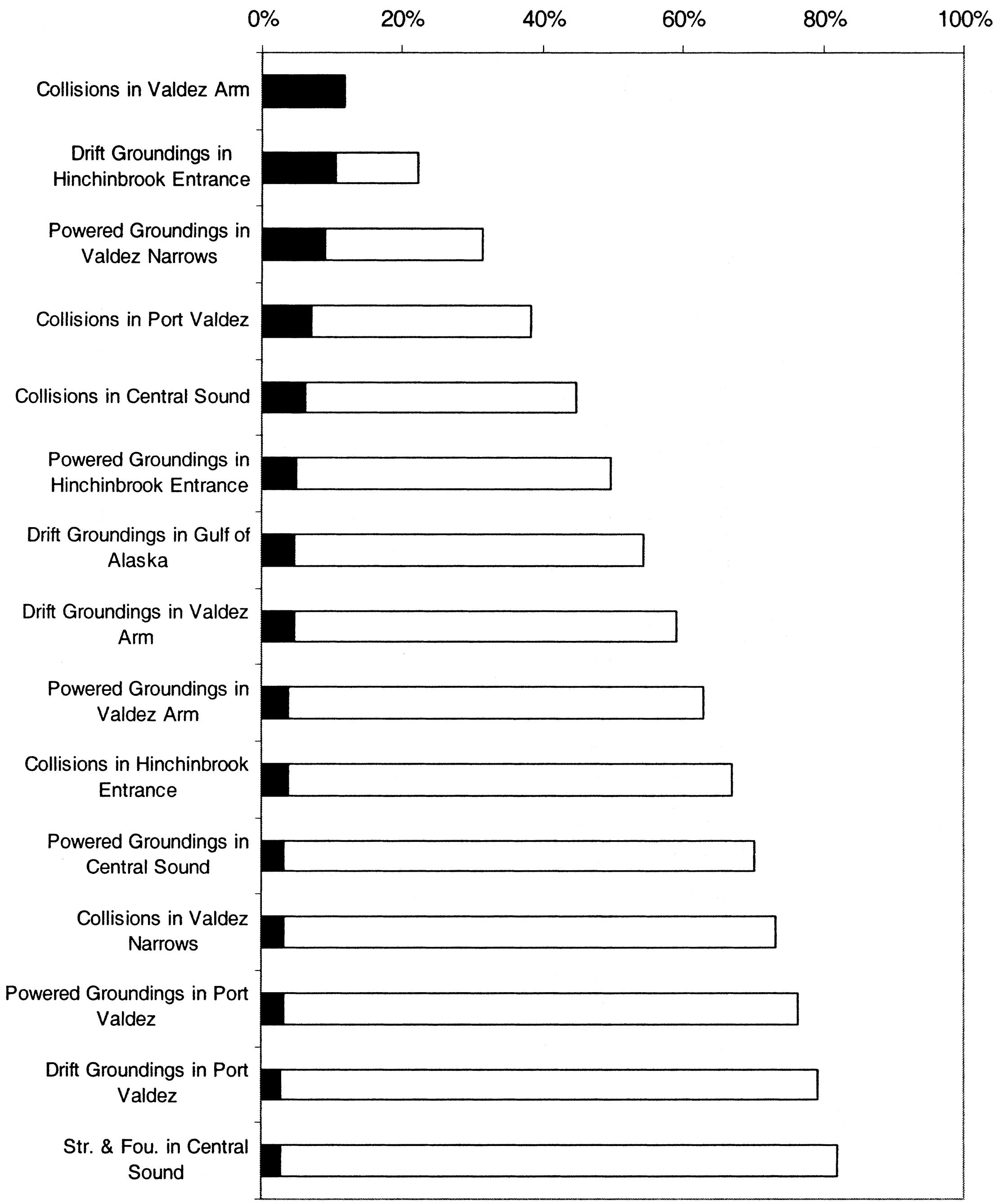


Figure 5: We sorted the combinations of accident types and locations by their expected oil outflow (dark bars). The cumulative percentage of the total expected oil outflow up to each such combination (white bars) is indicated by the total height of each bar. For example, we found that the first seven accident scenarios account for 55 percent of the total expected oil outflow.

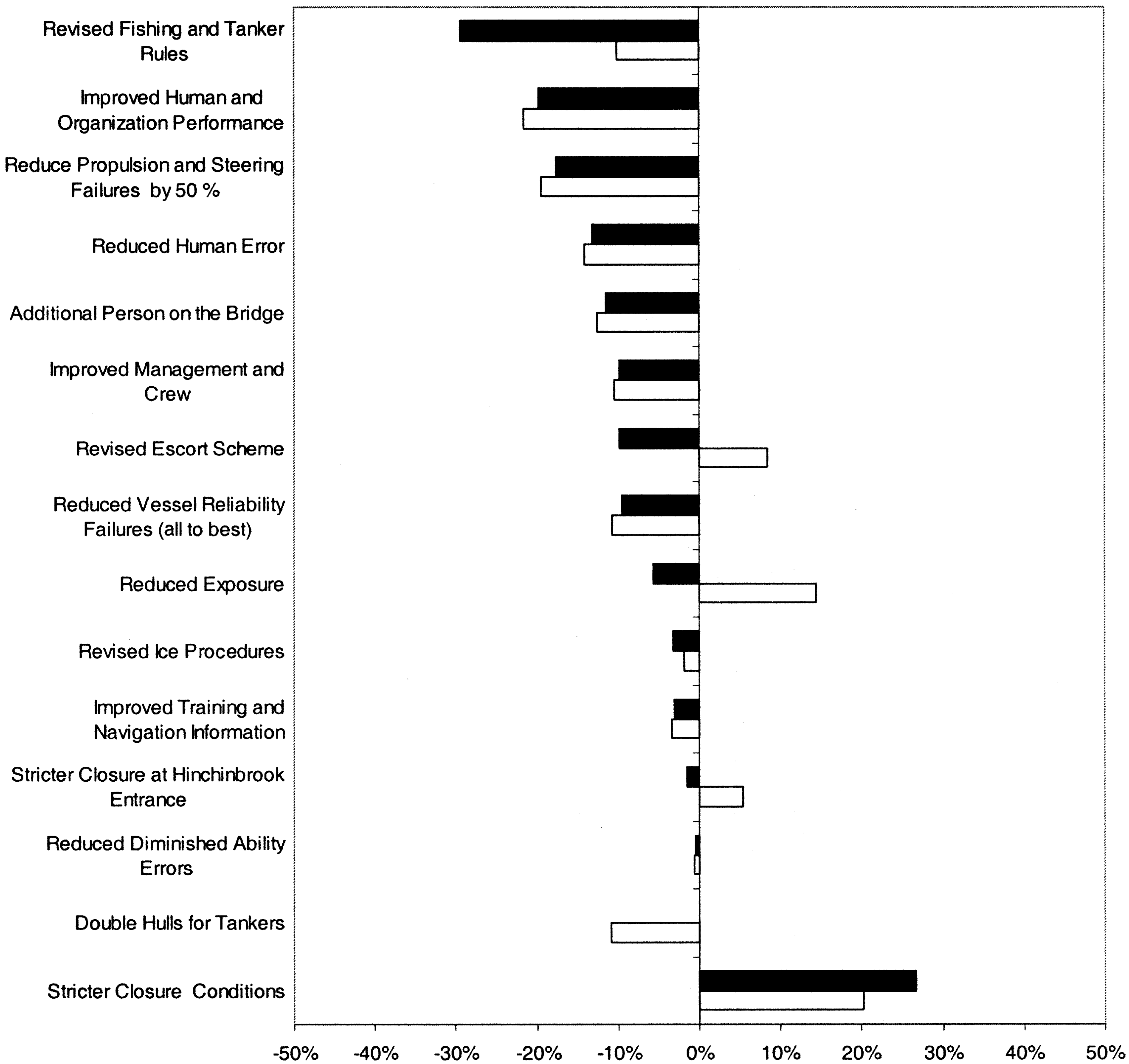


Figure 6: We tested proposed risk interventions in the system risk simulation and ranked them by percentage reduction from the study year in the expected frequency of accidents (black bars) and expected oil outflow (white bars) per year.

Estimates of expected accident frequency and expected oil outflow by accident scenario are point estimates of risk. The preferred method for reporting accident risk would be a distribution that also represents the degree of uncertainty in the results (Paté-Cornell 1996). Although we proposed an uncertainty analysis to the steering committee, time and budgetary constraints did not allow it. This was a drawback in the study, and additional research is needed to develop a technique to assess uncertainties in the system risk-simulation model. The value of an analysis, however, is not only in the precision of the results but in understanding system risk. Unlike risk assessments in more traditional areas, for example, nuclear power, our focus was the dynamic risk behavior of the system. For risk-management purposes, it is valuable to identify the peaks, patterns, unusual circumstances, and trends in system risk and in changes in system risk made by the implementation of risk-intervention measures.

Validity of the Results

In any study, it is important to validate the results. To assess the validity of our results, we need to validate both the simulation of the PWS system and the expert-judgment-based estimates of accident and incident probabilities. We used graphical comparison to the actual system and numerical comparison using summary statistics to validate the simulation part of the model. Specifically, USCG personnel from the Vessel Traffic Service (VTS) in PWS, who monitor traffic using screens resembling the graphical simulation output, verified the general behavior of traffic in the simulation regarding adherence to traffic rules, and patterns of vessel arrivals and departures. In addition, we compared summary statistics from the simulation, such as the average number of trips to the anchorage area as a result of weather-based closure conditions, the average number of tanker diversions due to ice in tanker lanes and the average number of closed waterways at separate locations due to weather restrictions, to those observed in the VTS system.

However, estimates of accident and incident probabilities based on expert judgments are more difficult to validate. While the use of proper procedures, such as structured and proven elicitation methods, can reduce uncertainty and bias in an analysis, they cannot eliminate them. As one referee noted, our use of mariners with experience in PWS could introduce a group bias. For example, had the Exxon Valdez not run aground, the opinions of the experts might have been quite different. The bias the referee refers to is availability bias (Cooke 1991), that is, people make assessments in accordance with the ease with which they can retrieve similar events. In the case of the Exxon Valdez accident, the effect of the availability bias would be to increase perceived levels of accident risk. However, each question in the PWS questionnaires required the comparison of two carefully defined scenarios. One could argue that both scenarios would be affected by the availability bias in a similar manner. As a result, the effect of the availability bias would be reduced. The Exxon Valdez accident scenario (a powered grounding of a tanker in the Valdez Arm) received only a modest ranking of 10 out of 17 accident scenario's that contribute to approximately 95 percent of total accident risk (Figure 4).

Risk assessments typically deal with low probability, high consequence events, and thus statistical validation of their results is difficult even when using nationwide or global accident databases. Using nationwide or global accident data in localized risk assessments is also questionable in terms of validity, prompting the PWS steering committee to require our use of only PWS specific data. This requirement meant we could not validate our risk assessment in the traditional sense. In the case of the probability of triggering incidents, such as mechanical failures, where available data and expert judgments overlapped, we observed good correspondence. Such correspondence could add to the validity of the other expert-based estimates, where such comparisons could not be made.

In the PWS risk assessment we followed a collaborative analysis approach (Charnley 2000). This included educating the steering committee in the language and modeling of risk. As we developed a common framework for analyzing risk, we discussed proposed risk-intervention measures at the level of their detailed effect on the whole system, rather than their gross effects on one part. We discussed the assumptions behind the model with the steering committee. The members of the steering committee were able to challenge the assumptions upon which they based their own opinions concerning the operation of the oil-transportation system in PWS.

We presented all our results to the steering committee in monthly meetings. The members questioned various results and often required more detailed analysis to reach a deeper understanding. The simulation model allowed us to demonstrate many results graphically, giving the steering committee a better intuition and trust in their validity. Members challenged certain results and often identified problems with the analysis, such as incorrect implementation of vessel traffic rules in the simulation, which we corrected. The committee put no pressure on us to change results merely because members disagreed. In the end, the steering committee unanimously accepted the results we obtained with the system risk-simulation model despite members' diverse perspectives at the onset of the study. Using the collaborative analysis approach, we built on the substantive knowledge represented in the steering committee and instilled trust in our results and recommendations, normally acquired through the use of classical statistical validation procedures.

Actions Taken

At the conclusion of the study, our contract team delivered a final report to the steering committee (PWS Steering Committee 1996). This report included technical documentation of the methodology used in the study, the results of the modeling, and recommendations based on these results. Following the risk-assessment project, the steering committee split up into risk-management teams charged with implementing the recommendations in specific areas.

One of the key questions the steering committee asked at the start of the study was whether the current escort system was capable of stopping drift groundings in the Valdez Narrows. The study showed that the current escort tugs were capable of saving a disabled tanker in the environmental conditions experienced in the Valdez Narrows. However, because of other considerations, the PWS shipping companies decided to accept proposals for two tractor-tugs. The designers used our result extensively in the design process. Crowley Maritime Services have invested $30 million to build the tugs Nanuq (Figure 7) and Tan'erliq to fulfill the requirements developed.

To date the various organizations comprising the risk-management teams have taken the following actions based on our results:

—The oil companies have introduced an enhanced-capability tug called the Gulf Service (Figure 8) to escort oil-laden tankers through Hinchinbrook Entrance, which is being replaced by new azimuthing stern-drive escort vessels designed for higher transit speed/open water assist scenarios that include the Hinchinbrook Entrance transit.

—We have completed a further project to find an improved escort scheme, which SERVS have adopted, minimizing interactions between oil tankers and escort tugs, while maintaining the ability to save disabled tankers.

Figure 7: The 153-foot, 10,000 horsepower, state-of-the-art tractor-tug Nanuq has been put in service to escort tankers through Valdez Narrows.

Figure 8: The enhanced capability tug Gulf Service has been stationed at Hinchinbrook Entrance to save disabled tankers even in extreme environmental conditions.

—The Coast Guard VTS manage interactions between fishing vessels and tankers.

—SERVS has increased the minimum required bridge crew on board escort tugs from one to two to add additional error-capture capability.

—The International Maritime Organization has approved a change to the tanker route through central PWS, reducing the number of course changes required.

—The shipping companies have made long-term plans for quality-assurance and safety-management programs.

The Benefits of the Risk-Assessment Process

It is difficult to compare this project with other more traditional projects in operations research and management science, whose benefits are typically measured in terms of reduced operating costs or increased profits. The benefits of risk assessments are less tangible as the objective is to reduce the occurrence of future accidents. However, because clean-up operations for the Exxon Valdez accident cost over $2 billion, the benefits of preventing a single such accident would be of similar magnitude. We can only estimate the reduction in the frequency of accidents using our models and can only estimate the benefits of the study in terms of clean-up cost. Using our risk models, we estimated that accident frequency had been reduced by 75 percent since the Exxon Valdez accident. According to our risk models, the further reduction in accident frequency from all measures taken as a result of the PWS risk assessment is 68 percent, with a 51 percent reduction in the expected oil outflow. This means that, since the Exxon Valdez accident, the accident frequency has been reduced by an estimated total of 92 percent. The costs of the risk assessment, roughly $2 million over a two-year period, pale in comparison to the potential clean-up costs for a single major oil spill resulting from a tanker accident. However, the benefits go beyond clean-up costs and include the protection of pristine environments, and the prevention of loss of life and injury to vessel crews. In addition, the shipping companies have used the results of the PWS model in making decisions to invest in multimillion dollar equipment.

While the stakeholders in PWS all recognized the need for a rational method to evaluate the merits of risk-intervention measures, to improve the allocation of resources, and to avoid implementing measures that would adversely affect system risk, they did not trust each other at the beginning of the project. The steering committee wanted to use the project as a forum to build trust amongst stakeholders, to educate all interested parties, and to provide a common understanding of oil-transportation risk. The PWS risk assessment fostered a cooperative risk-management atmosphere involving all stakeholders.

At the end of the project, the stakeholders published the final report as their document, not just as a report from the consultant team. Members of the steering committee from environmental groups, the fishing industry, and the oil companies wrote joint press briefings and formed risk-management teams to manage implementation of the model results. The unified acceptance and presentation of the results of the study by all stakeholders and the level of implementation of the results can be primarily considered a benefit of the collaborative analysis process. All stakeholders finished the project convinced that they had reduced risk of further multibillion dollar accidents and, with the cooperation fostered by the collaborative analysis process, the stage has been set for further improvements in managing risk.

The success of the PWS risk assessment has not gone unnoticed, and the National Science Foundation has awarded other researchers funding (for example, NSF SBR-9520194, NSF SBR-9710522) to study the risk-assessment process we followed. Our study is described as an example of collaborative analysis by Busenberg (2000) and Charnley (2000). Busenberg (1999) commented as follows:

> "All ten of the participants who were interviewed agreed that this process allowed the steering committee to gain a better understanding of the technical dimensions of maritime risk assessment . . . The results of the risk assessment were released in late 1996, and were unanimously accepted as valid by the RCAC, oil industry, and government agencies involved in this issue. The participating groups agreed that the study showed the need for an ocean rescue tug vessel in the Sound. In 1997, the oil industry responded by deploying a vessel of this class in the Sound."

Acknowledgments

The authors are indebted to the editor and associate editor of *Interfaces* and the referees for their valuable comments and suggestions that substantially improved the first version of this paper.

References

Andrews, S., F. H. Murphy, X. P. Wang, S. Welch. 1996. Modeling crude oil lightering in Delaware Bay. *Interfaces* **26**(6) 68–78.

Banks, J., J. S. Carson, B. L. Nelson, D. M. Nicol. 2000. *Discrete-Event System Simulation*. Prentice Hall, Upper Saddle River, NJ.

Bedford, T. M., R. M. Cooke. 2001. *Probabilistic Risk Analysis: Foundations and Method*. Cambridge University Press, Cambridge, U.K.

Bradley, R., M. Terry. 1952. Rank analysis of incomplete block designs. *Biometrica* **39** 324–345.

Busenberg, G. 1999. Collaborative and adversarial analysis in environmental policy. *Policy Sci.* **32**(1) 1–11. Supported under NSF SBR-9520194.

——. 2000. Innovation, learning, and policy evolution in hazardous systems. *Amer. Behavioral Sci.* **44**(4) 1–11. Supported under NSF SBR-9520194, NSF SBR-9710522.

Charnley, G. 2000. *Enhancing the Role of Science in Stakeholders-Based Risk Management Decision-Making*. HealthRisk Strategies, Washington, DC.

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

Fowler, T. G., E. Sorgard. 2000. Modeling ship transportation risk. *Risk Anal.* **20**(2) 225–244.

Garrick, G. J. 1984. Recent case studies and advancements in probabilistic risk assessment. *Risk Anal.* **4**(4) 267–279.

Golkar, J., A. Shekhar, S. Buddhavarapu. 1998. Panama Canal simulation model. D. J. Medeiros, E. F. Watson, J. S. Carson, M. S. Manivannan, eds. *Proc. 1998 Winter Simulation Conf.* 1229–1237.

Guedes Soares, C., A. P. Teixeira. 2001. Risk assessment in maritime transportation. *Reliability Engrg. System Safety* **74**(3) 299–309.

Hara, K., S. Nakamura. 1995. A comprehensive assessment system for the maritime traffic environment. *Safety Sci.* **19**(2–3) 203–215.

Harrald, J., H. Marcus, W. Wallace. 1990. The Exxon Valdez: An assessment of crisis prevention and management systems. *Interfaces* **20**(5) 14–30.

——, T. Mazzuchi, J. Merrick, R. van Dorp, J. Spahn. 1998. Using system simulation to model the impact of human error in a maritime system. *Safety Sci.* **30**(1–2) 235–247.

Kite-Powell, H. L., D. Jin, N. M. Patrikalis, J. Jebsen, V. Papakonstantinou. 1996. Formulation of a model for ship transit risk. MIT Sea Grant Technical Report, Massachusetts Institute of Technology, Cambridge, MA, 96–119.

Lowrance, W. W. 1976. *Of Acceptable Risk*. William Kaufman, Los Altos, CA.

Maio, D., R. Ricci, M. Rossetti, J. Schwenk, T. Liu. 1991. Port needs study. Report No. DOT-CG-N-01–91–1.2. Prepared by John A. Volpe, National Transportation Systems Center, U.S. Coast Guard, Washington, DC.

Merrick, J., J. R. van Dorp, J. Harrald, T. Mazzuchi, J. Spahn, M. Grabowski. 2000. A systems approach to managing oil transportation risk in Prince William Sound. *Systems Engrg.* **3**(3) 128–142.

National Research Council. 1986. *Crew Size and Maritime Safety*. National Academy Press, Washington, DC.

——. 1991. *Tanker Spills: Prevention by Design*. National Academy Press, Washington, DC.

——. 1994. *Minding the Helm: Marine Navigation and Piloting*. National Academy Press, Washington, DC.

——. 2000. *Risk Management in the Marine Transportation System*. National Academy Press, Washington, DC.

Paté-Cornell, M. E. 1990. Organizational aspects of engineering system safety: The case of offshore platforms. *Science* **250**(4985) 1210–1217.

——. 1996. Uncertainties in risk analysis: Six levels of treatment. *Reliability Engrg. System Safety* **54**(2–3) 95–111.

Pravda, M. F., R. G. Lightner. 1966. Conceptual study of a supercritical reactor plant for merchant ships. *Marine Tech.* **4** 230–238.

Prince William Sound Steering Committee. 1996. Prince William Sound risk assessment study final report.

Roeleven, D., M. Kok, H. L. Stipdonk, W. A. de Vries. 1995. Inland waterway transport: Modeling the probabilities of accidents. *Safety Sci.* **19**(2–3) 191–202.

Ryan, N. K. 1998. The future of maritime facility designs and operations. D. J. Medeiros, E. F. Watson, J. S. Carson, M. S. Manivannan, eds. *Proc. 1998 Winter Simulation Conf.* 1223–1227.

Saaty, T. 1977. A scaling method for priorities in hierarchical structures. *J. Math. Psych.* **15**(3) 234–281.

Slob, W. 1998. Determination of risks on inland waterways. *J. Hazardous Materials* **61**(1–3) 363–370.

Slovic, P. 1993. Perceived risk, trust and democracy. *Risk Anal.* **13**(6) 675–682.

Stiehl, G. L. 1977. Prospects for shipping liquefied natural gas. *Marine Tech.* **14**(4) 351–378.

Trbojevic, V. M., B. J. Carr. 2000. Risk based methodology for safety improvements in ports. *J. Hazardous Materials* **71**(1–3) 467–480.

U.S. Coast Guard. 1973. Vessel traffic systems: Analysis of port needs. Report No. AD-770 710. U.S. Coast Guard, Washington, DC.

U.S. Nuclear Regulatory Commission. 1975. Reactor safety study: An assessment of accident risks in U.S. commercial nuclear power plants. WASH-1400 (NUREG-75/014).

Wang, J. 2000. A subjective modeling tool applied to formal ship safety assessment. *Ocean Engrg.* **27**(10) 1019–1035.

Richard L. Ranger, Manager, Operational Integrity, Polar Tankers, Inc., 300 Oceangate, 11th Floor, Long Beach, California 90802-4341, writes: "During the period from September 1995 through December 1996, I was one of the representatives of ARCO Marine, Inc.

on the multi-stakeholder Steering Committee established to oversee the work of the consultant team on Prince William Sound Risk Assessment project. In the period that followed I represented ARCO Marine (now Polar Tankers, Inc.) in a succession of multi-stakeholder discussions which considered implementation of risk mitigation measures identified during the PWS Risk Assessment.

"In its review of the system then in place for marine transportation of crude oil in Prince William Sound, Alaska, the PWS Risk Assessment tested the capabilities of current methods of probabilistic risk analysis, and established some new benchmarks for use of certain analytical methods in combination. To the participating stakeholders, who use, regulate, or benefit from the PWS marine transportation system, the principal value of the PWS Risk Assessment was the fact that it undertook quantitative risk characterization in the context of the values, norms, and expectations of our diverse group. Science and method were tested against assumptions based upon policy and perception. In turn, science and method tested and challenged these other means of decision making. Researchers learned from stakeholders, and vice versa. The outcome was not simply a detailed project report but a deepened understanding by all stakeholders regarding where improvements in the system might be possible, of realistic expectations for those improvements, and of the nature and significance of uncertainties about both.

"The years since the publication of the report from the PWS Risk Assessment have not been free from disagreement among the stakeholders, but they have been years of a substantially improved quality of dialogue, and of more informed decision making. They have also been years marked by steady incremental improvement in the capability of the PWS marine transportation system to prevent vessel casualties and pollution incidents from occurring. The PWS Risk Assessment was clearly a catalyst in achieving these outcomes. It marks a unique convergence of technical inquiry and stakeholder dialogue that balanced analysis appropriate to the problem with deliberation over the needs and interests of affected parties.

"Like many pathbreaking efforts, the PWS Risk Assessment did not reach such results easily, nor necessarily within the original budget and schedule expectations of any of the participants. Still, it represents an important reference point for future projects that involve assessment of operational risk in the context of public dialogue about such risk, its components, its acceptability, and its potential consequences."

A. Elmer III, President, SeaRiver Maritime, Inc., PO Box 1512, Houston, Texas 77251-1512, writes: "The PWS Risk Assessment was proposed by PWS Shipping Companies to foster an environment in which the often misunderstood and complex concept of maritime risk could be discussed and reviewed by all stakeholder parties concerned with the safety of marine transportation in Prince William Sound. To facilitate the process, the consultant team was asked to join with the PWS Steering Committee in studying and evaluating the risks associated with the transporting of Alaskan North Slope crude oil from Valdez through Prince William Sound, Alaska.

"The consultant team developed a framework that described, qualitatively, the risks and built models based upon this framework. The PWS Steering Committee was first educated in the concept and language of risk and risk management and the framework in which to study risk. The PWS Steering Committee then participated in the development of the modeling assumptions upon which the models were based. This process fostered continual open discussion and dialogue on the detailed and specific effect of proposed changes to the marine transportation system.

"The close coordination of the risk model development through the PWS Steering Committee led to a high level of trust in the results and consensus on changes to be made to the system. Following the project, results of the risk assessment study have been implemented, including the following:

• The stationing of an enhanced-capability tug at Hinchinbrook Entrance.

• A redesigning of the tanker escort system to ensure that tankers are escorted by suitable escort tugs in each area of Prince William Sound.

• Establishing improved coordination between tankers and escort tugs and maintaining the ability to respond to a disabled tanker.

• The implementation of close coordination of tanker movement with other PWS activities (e.g., commercial fishing season openings) to ensure safety of transit.

• Continual improvement of shipping companies' Safety Management Systems and training programs.

"The PWS Risk Assessment project consultants brought industry, industry service groups, state and federal regulators, and public stakeholders together to work through the defining and assessment of marine transportation risk and the development of risk-reduction measures for the PWS Marine Transportation System."

J. P. High, Acting Assistant Commandant for Marine Safety and Environmental Protection, United States Coast Guard, 2100 Second Street SW, Washington, DC 20593-0001, writes: "The U.S. Coast Guard was one of the sponsors of the Prince William Sound Risk Assessment and remains heavily involved in past and ongoing efforts to manage risks associated with commercial shipping in Prince William Sound and elsewhere.

"The submitted risk assessment was the first such assessment of its size and was groundbreaking relative to both the scope of the effort and the large number of diverse stakeholders. The results of the assessment were used to directly support decisions made by the stakeholders that have reduced risks in the area. Additionally, as the first of its size, this study has been a very useful benchmark for other similar risk assessments.

"The U.S. Coast Guard strongly supports efforts to improve maritime safety, especially those like this one that focused on risk identification, evaluation, and management."

SUB-APPENDIX:

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.

Risk Analysis, Vol. 21, No. 1, 2001

A Risk Management Procedure for the Washington State Ferries

Johan R. van Dorp,[1*] Jason R. W. Merrick,[2] John R. Harrald,[1] Thomas A. Mazzuchi,[1] and Martha Grabowski[3]

The state of Washington operates the largest passenger vessel ferry system in the United States. In part due to the introduction of high-speed ferries, the state of Washington established an independent blue-ribbon panel to assess the adequacy of requirements for passenger and crew safety aboard the Washington state ferries. On July 9, 1998, the Blue Ribbon Panel on Washington State Ferry Safety engaged a consultant team from The George Washington University and Rensselaer Polytechnic Institute/Le Moyne College to assess the adequacy of passenger and crew safety in the Washington state ferry (WSF) system, to evaluate the level of risk present in the WSF system, and to develop recommendations for prioritized risk reduction measures, which, once implemented, can improve the level of safety in the WSF system. The probability of ferry collisions in the WSF system was assessed using a dynamic simulation methodology that extends the scope of available data with expert judgment. The potential consequences of collisions were modeled in order to determine the requirements for onboard and external emergency response procedures and equipment. The methodology was used to evaluate potential risk reduction measures and to make detailed risk management recommendations to the blue-ribbon panel and the Washington State Transportation Commission.

KEY WORDS: Maritime risk assessment; system simulation; expert judgment

1. INTRODUCTION

The Washington state ferry (henceforth WSF) system is the largest ferry system in the United States. In 1997, total ridership for the ferries serving the central Puget Sound region was nearly 23 million, a 4% increase over 1996 ridership, and more passengers than Amtrak, the U.S. passenger rail carrier, handles in a year. The state of Washington instituted the ferry system in 1951 to connect King and Snohomish Counties with Kitsap County, saving travelers the long drive around Puget Sound via the Tacoma Narrows Bridge, and to provide mainland access to Vashon Island and Whidbey Island. Prior to 1951 private ferry system(s) offered these services. Figure 1 shows the current ferry routes for the central Puget Sound region. This map illustrates the ferry system's role in linking together the Washington state highway system in the Puget Sound region.[(1)]

Though to date the Washington state ferries have had an exceptional safety record, the WSF system is facing a number of important changes. First, its regulatory environment, which has been relatively inactive, has changed significantly with the implementation of 46 C.F.R. 199, Subchapter W, of the Code of Federal Regulations, Lifesaving Systems for Certain Inspected Vessels.[(2)] The WSF system is required by these regulations to address the response to cata-

[1] The George Washington University, Washington, DC.
[2] Virginia Commonwealth University, Richmond, VA.
[3] Rensselaer Polytechnic Institute, Troy, NY.
*Address correspondence to: Johan René van Dorp, Engineering Management and Systems Engineering Department, The George Washington University, 707 22nd Street N.W., Washington, DC 20052; dorpjr@seas.gwu.edu.

0272-4332/01/0200-0127$16.00/1

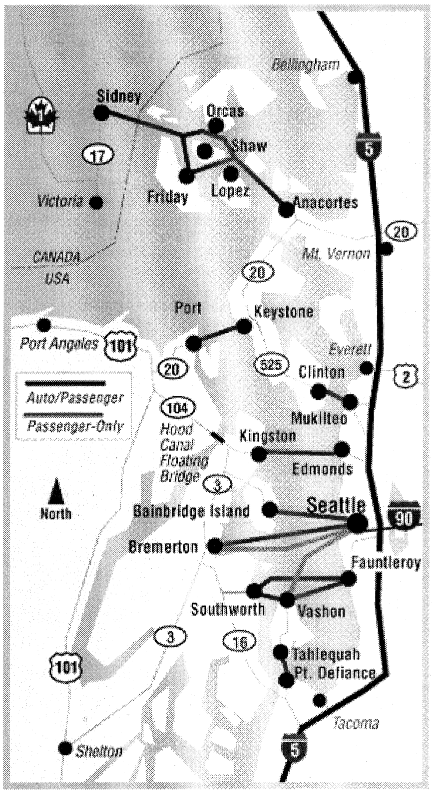


Fig. 1. Washington state ferry system map.

strophic accidents and the requirements for ensuring that passengers could survive such accidents. Specifically, the regulations require the WSF system, within 5 years, either to equip all ferries with adequate survival craft or to provide a safety assessment, a comprehensive shipboard safety management system, and shipboard contingency plans approved by the U.S. Coast Guard (USCG), the U.S. regulatory body for maritime affairs.

A second set of changes in the WSF system stems from pressures to develop a seamless, intermodal transportation system in Washington state in the face of simultaneous increases in the volume and mix of riders on the ferries. Because increasing numbers of Washington state residents are riding the ferries to work, and because connections to other transportation modes (bus, bicycle, car) from the ferries are critical to the success of such an intermodal system, the WSF system is under increased pressure to perform in ways different from those of the past, to measure and report its performance in different ways, and to increase the fluidity with which connections to other transportation modes are made from the ferries.

A third set of changes in the WSF system stems from new technology, for example, high-speed ferries, being introduced into the system to address some pressures for faster transport–passenger-only ferries. These new technologies are being introduced into an aging fleet with some consideration given for how best to mix new and old vessels, new and old technology, new and old operational dynamics, and varying degrees of sophisticated automation. In addition, the International Maritime Organization (IMO) has enacted implementation of the Standards for Training and Certification of Watchkeeping (STCW)[3] for all vessels above 200 gross tons (GT) and has begun the process of developing a high-speed code for vessels. To date the WSF has been exempt from STCW requirements and is in full compliance with all prevention regulations. The focus on high-speed ferries could change this status.

In light of these changes, the state of Washington established the independent Blue Ribbon Panel on Washington State Ferry Safety to assess the adequacy of requirements for passenger and crew safety aboard the Washington state ferries. On July 9, 1998, the panel engaged a consultant team from The George Washington University and Rensselaer Polytechnic Institute/Le Moyne College to assess the adequacy of passenger and crew safety in the WSF system, to evaluate the level of risk present in the WSF system, and to develop recommendations for prioritized risk reduction measures, which, once implemented, can improve the level of safety in the WSF system.

This article provides a discussion of (1) a framework for risk assessment and risk management of maritime transportation systems, (2) an overview of the modeling approach used in the WSF risk assessment, (3) an overview of WSF baseline risk assessment results, (4) WSF risk intervention evaluation results, and (5) recommendations to the panel and the Washington State Transportation Commission.

2. A FRAMEWORK FOR RISK ASSESSMENT AND MANAGEMENT

In order to evaluate proposed risk interventions, one must first define a measure of risk. Risk is often defined by combining the likelihood of an undesirable event and relevant consequences in a single quantitative measure. For example, consequences may include injury, loss of life, or economic losses. It is also possible to define some surrogate measure of risk that indirectly accounts for such attributes. Next, one needs to understand the events and situations that lead to

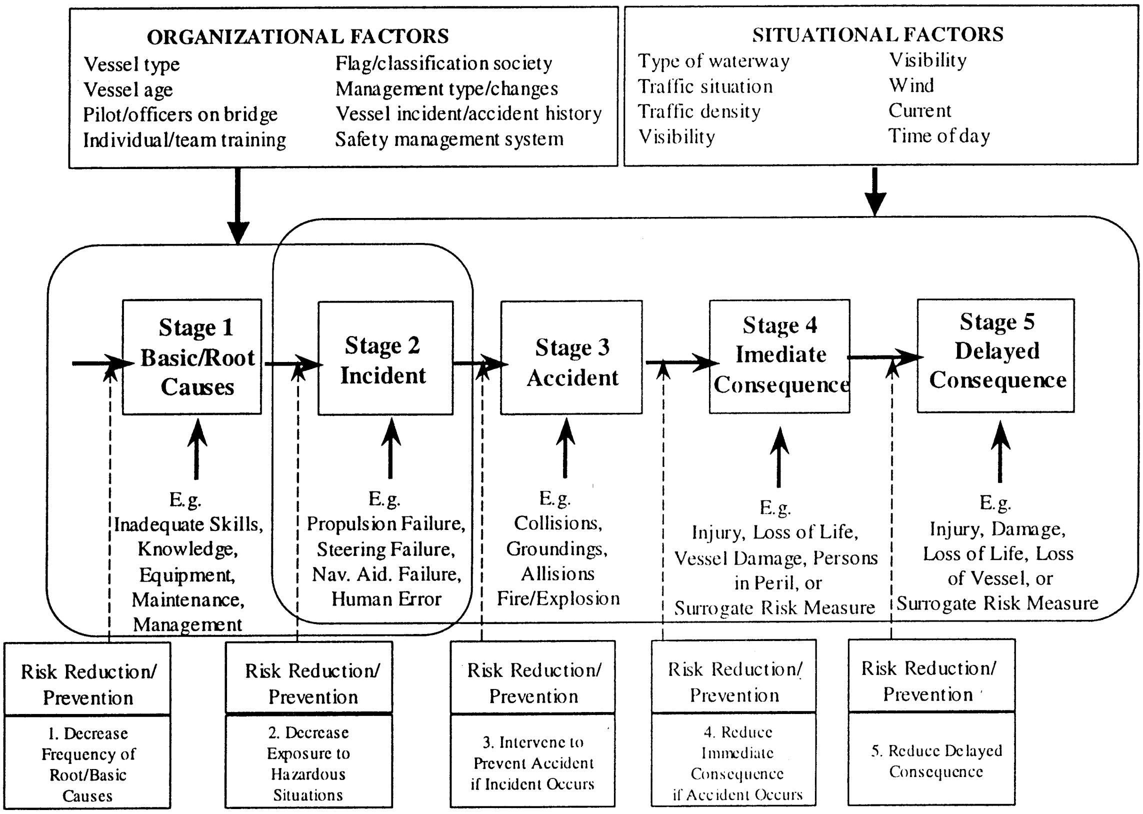


Fig. 2. The maritime accident event chain.

the undesirable event and the impact of proposed risk interventions on these events and situations. Figure 2 shows the maritime risk taxonomy used by the study team and illustrates the importance of organizational and situational factors in both the occurrence and severity of an accident.

In addition, Fig. 2 identifies five categories of risk interventions based on intended impact on the accident event chain. Three categories of impact intend to reduce the likelihood of occurrence of accidents and two categories of impact intend to reduce the consequences of accidents that could occur. Note that a single risk intervention may belong to multiple impact categories.

The objective of risk management is to structure, evaluate, rank, and implement policies and procedures that reduce the threat to life, property, the environment or all of the above posed by hazards. The structuring and evaluation of risk management alternatives/risk interventions herein is based on a multistep process. The first step is to define a quantitative measure of risk. In this study a surrogate consequence measure was defined focusing on response time alternatives as required by Subchapter W while addressing risk communication concerns of the blue-ribbon panel in terms of providing the results to the public. This surrogate measure will be introduced in Section 3.1. The second step is to identify potential risk interventions and determine their impact on the accident event chain (see, for example, Fig. 2). The third step is to develop a comprehensive quantitative model for comparing the risk interventions in a meaningful manner. The fourth step is to establish a baseline level of risk by defining a baseline scenario and using the developed model to quantify its risk. Additional risk intervention measures may be identified by focusing on high-risk contributors to the baseline level of risk. The fifth step is to model the effect of all the risk interventions in terms of changes to model parameters. The final step is to implement these changes to the model and evaluate the risk interventions relative to the established baseline level of risk.

The ranking and implementation of risk interventions involves assessment of tradeoffs of risk reduction with respect to other measures of interest, such as cost, implementation time, and political acceptability. While this was an important part of the

WSF risk assessment, the ranking and implementation is not a topic discussed further in this article. Rather, the focus is on the assessment of baseline risk and the evaluation of risk interventions.

3. RISK INTERVENTION MODELING IN THE WSF SYSTEM

The six-step process used for structuring and evaluating risk interventions in the WSF risk assessment will be discussed in the sections below.

3.1. Defining Risk for the WSF System

The focus of this study was on passenger safety, including consideration of both the probability of occurrence and the severity of consequence of accidents. Accident types that are a potential threat to the Washington state ferries include collisions (or striking of another vessel), fires or explosions, allisions (or striking of a fixed object), and groundings (or strandings*)*. The potential vulnerability to these accidents is determined by the internal factors previously described and by factors external to the system, such as high levels of traffic congestion, the emergency coordination and response capabilities of external organizations, and the intentional or unintentional presence of hazardous materials on board.

The consequence evaluation focused on defining the appropriate accident response alternatives as required by Subchapter W. Hence, the risk analysis focused solely on WSF passengers. Accidents with vessels not putting WSF passengers in peril were not considered in the study. A measure termed "Maximum required response time" (MRRT) was developed as a surrogate measure for the potential accident impact. The MRRT was defined as the maximum allowable time for response to avoid additional (post-accident) injuries or fatalities due to a failure to respond in time. Three categories of MRRT were deemed appropriate: less than 1 hr, between 1 and 6 hr, and greater than 6 hr. In conjunction with the consulting team, the blue-ribbon panel judged that accidents in the first category primarily require an effective external emergency response, for example, other ferries or vessels, to prevent additional injuries or fatalities since the time would probably not permit in-time launching of survival craft. For accidents in the second category, time is available for evacuation to a safe haven. In order to meet Subchapter W requirements, the WSF system must demonstrate the ability to mobilize evacuation vessels or plan to provide survival craft adequate for all passengers. For accidents in the third category, adequate response in all cases can be provided without evacuating the passengers from the ferry. Of course, in any accident it is desirable to respond in the shortest amount of time possible. The MRRT measure merely provides an upper bound on the desirable response time.

Historical records for all accident events involving Washington state ferries were collected for an 11-year period and analyzed. Fire and explosions were limited, historically, to stack fires that were contained while under way. Allisions were incidents occurring at the dock and led primarily to property damage and not casualties or injuries as the impact speeds were low. Groundings occurred at shallow areas with small tide fluctuations. In each case, the ferry involved remained a stable, safe platform for the passengers until an orderly evacuation was performed. There were two collisions in an 11-year period of accident data. In each collision, the ferry was able to return to dock and safely disembark the passengers. Summarizing, the Washington state ferries have a commendable safety record in terms of casualties and injuries, with no fatalities.

Potential accident scenarios that could lead to high consequences in injuries and fatalities were, however, developed in conjunction with the Blue Ribbon Panel on Washington State Ferry Safety. Specifically, collisions involving high-speed ferries, collisions between ferries and deep-draft vessels, and acts of intentional fire/explosion were deemed to be events that could possibly fall in the 1–6 hr MRRT and less than 1 hr MRRT categories. Due to the sensitivity of acts of intentional fire/explosion, the panel decided that it was not appropriate to discuss the vulnerability to these acts in the open public forum of the WSF risk assessment. Based on the characteristics of the WSF system, allisions and groundings were judged by the project team, in conjunction with maritime experts, to fall in the more than 6 hr MRRT category. The blue-ribbon panel accepted this assumption. Hence, the main focus was the development of models for collision risk estimating the frequency of collisions and their associated consequences in terms of the three MRRT categories identified.

3.2. Identification and Structuring of Risk Interventions

In the WSF risk assessment, the project team collected a total of 40 risk reduction measures that had been proposed for this system and for other maritime systems, and structured the measures. The

Table I. Summary of Risk Interventions Classes Tested

Risk reduction class	Intervention
1	Adopt international safety management standard fleetwide
2	Implement all mechanical failure reduction measures fleetwide
3	Implement high-speed ferry rules and procedures
4	Implement weather, visibility restrictions
5	Implement traffic separation for high-speed ferries
6	Implement traffic control for deep-draft traffic
7	Increase time available for response

sources of these measures were (1) interviews with ferry system and U.S. Coast Guard personnel, (2) the Revision of the HSC Code, Formal Safety Assessment of High Speed Catamaran (HSC) Ferries Submitted by the United Kingdom,[4] (3) the *Final Report: Prince William Sound Risk Assessment*,[5] (4) Scoping Risk Assessment, Protection against Oil Spills in the Maritime Waters of Northwest Washington State,[6] and (5) alternatives specified in 46 CFR 199, Subchapter W. The 40 risk reduction measures were synthesized to seven classes of risk reduction measures, listed in Table I. The intended impact of these classes on the causal chain of Fig. 2 is displayed in Fig. 3. Note that some classes intervene at multiple points in the accident event chain.

3.3. An Overview of the Modeling Approach for WSF System Collision Risk

The situational and organizational factors, indicated in Fig. 2, that influence the probability of occurrence of events in the causal chain lead to dynamic fluctuations in system collision risk. Identifying how and when these risk spikes occur is a fundamental objective of the use of dynamic system simulation as a risk assessment methodology. As an example of the contribution of situational factors to collision risk, it is clear that a ferry traveling on a clear day with no other traffic nearby is at lower risk than a ferry in foggy conditions with many other vessels nearby. Modeling the contribution of risk factors asks for a quantitative evaluation of collision risk in both situations, that is, how much more risky the first situation is compared to the other. In the WSF risk assessment, a constructive modeling approach combining system simulation, expert judgement, and available data was used to allow for estimation of the contribution of these situational and organizational factors to collision risk.

A specific combination of situational and organizational factors in a given time point for a specific ferry is an opportunity for incident (OFI). Thus each OFI consists of variables that may be considered contributing risk factors. The risk factors considered in the WSF risk assessment are listed in Table II. Modeling the system in terms of the factors in Table II, re-

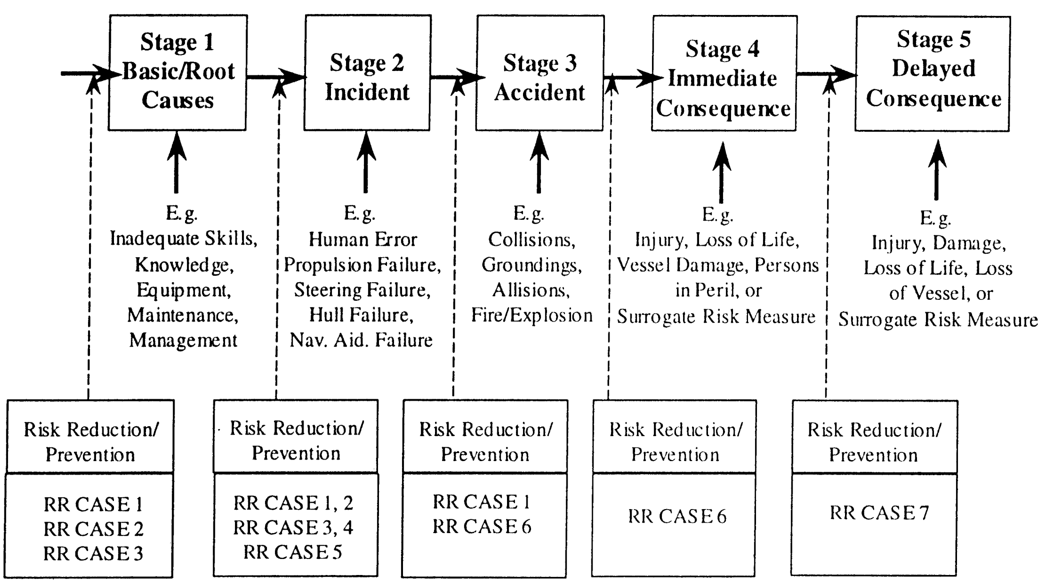


Fig. 3. Impact of risk reduction classes on the causal chain. RR = risk reduction.

Table II. The Variables Considered in the Collision Risk Model

Variable name	Possible values
Ferry route	Seattle-Bremerton, Anacortes-Sidney, etc.
Ferry class	Issaquah, Jumbo, Chinook, etc.
Interacting vessel type	Container, bulk carriers, other ferries, etc.
Type of interaction	Crossing, meeting, overtaking
Proximity of interacting vessel	Less than 1 mile, from 1 to 5 miles
Wind speed	0 knots, 10 knots, 20 knots
Wind direction	Perpendicular to ferry, along ferry
Visibility	Less than 0.5 mile, more than 0.5 mile

quires extensive collection of traffic and weather data. Traffic data are available from the USCG logging arrivals of deep-draft vessels to the Puget Sound area. Ferry schedules are published by the Washington State Ferry Service. Weather data was obtained from the National Oceanic and Atmospheric Administration (NOAA) and local airport data. A visibility model was created using a land visibility model developed with local airport data and a sea visibility model using dew point temperature data and water temperature data from NOAA weather buoys.

Traffic data in terms of annual statistics alone cannot be used to infer how often interactions between these vessels occur and in what conditions. Thus, a simulation of the WSF system was built to represent the movement of the Washington state ferries, the movement of other vessels in the area, and the environmental conditions at any given time. Figure 4 gives a screen capture of the WSF system simulation capturing the southern Puget Sound area and central Puget Sound Area. Figure 4 displays (1) ferry routes in central Puget Sound, (2) two wind fans modeling direction and strength in the central Puget Sound and southern Puget Sound regions, (3) bad-visibility conditions (less than 0.5 miles) in southern Puget Sound, and (4) good visibility in central Puget Sound.

Using this simulation, a counting model was developed that observed and recorded snapshots of the study area at regular intervals and counted the occurrences of the various OFIs in terms of the variables displayed in Table II. The simulation is called the OFI generator and the counting model is called the OFI counter. Using the OFI counter, summary statistics on, for example, the number of OFIs involving crossing situations of a high-speed ferry and a container vessel on the Seattle Bremerton route in bad visibility conditions can be analyzed. The next step is to assess the likelihood of triggering incidents and collisions given the risk factors in Table II.

The preferred method for estimating these probabilities is through data. Accident database information is typically limited, however, to accident and immediate-consequence data, as indicated by Fig. 5. For evaluation of the risk intervention measures impacting early on in the causal chain, the assessment of probabilities in the beginning of the causal chain is required. The assessment of incident probabilities leading to an accident, however, is often not supported by available data in accident and consequence databases. Cooke[7] cites the use of expert judgment in areas as diverse as aerospace programs, military intelligence, nuclear engineering, evaluation of seismic risk, weather forecasting, economic and business forecasting, and policy analysis. Paté-Cornell[8] discusses the necessity of using expert judgment when sufficient data are not available, and Harrald, Mazzuchi, and Stone[9] proposed the use of expert judgment in the analysis of risk in maritime environments.

In the WSF risk assessment, the average likelihood of system events along the maritime accident event chain was estimated using both historical data and expert judgment. A database containing 11 years of incident, accident, and transit data for Puget Sound and the inland waters of the state of Washington was created for this project, reconciling USCG, state of Washington, Marine Exchange, U.S. Army Corps of Engineers, and ferry system databases through rigorous data selection and cross validation. Expert judgment was obtained from WSF captains, USCG personnel, and members of the Puget Sound Pilots Association using elicitation methods based on pairwise comparisons of OFIs. The expert judgment was combined with and calibrated to the accident and incident data available and was used to model the effect of the variables in Table II on the accident and incident probabilities. Figure 6 summarizes the use of the different modeling techniques to establish collision frequencies.

The final step in modeling the maritime accident event chain is consequence modeling. Engineering models of collision impact damage scenarios were used to assess the damage to each ferry class in various collision scenarios. The damage model follows the method of Minorsky.[10] The Minorsky method determines damage size as a function of the collision energy, the colliding-vessel bow angle, and the effective deck thickness of the Washington state ferries. The collision energy is calculated using the masses of both the struck ship (ferry) and the striking ship. The dam-

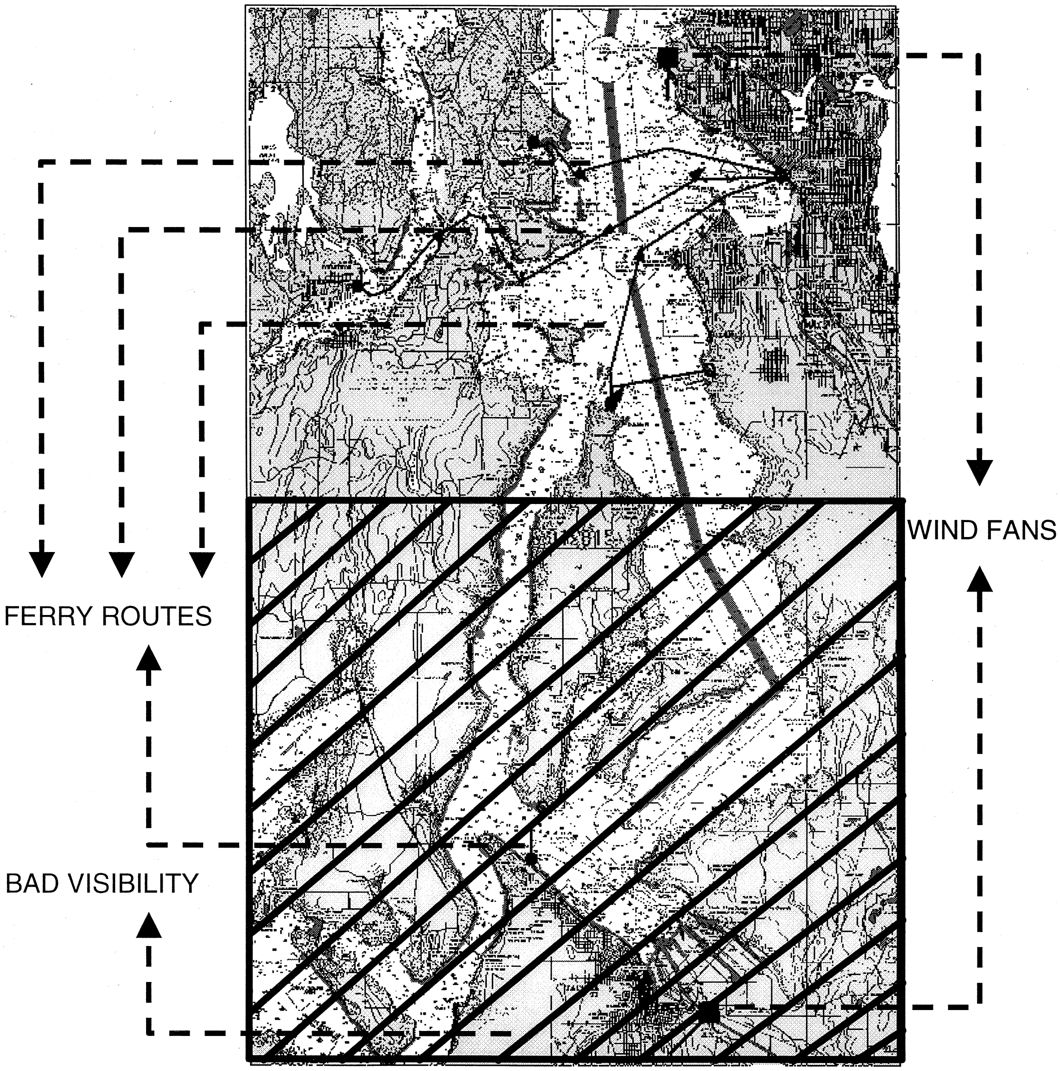


Fig. 4. Screen capture of the Washington state ferry system simulation.

age calculation results in a damage penetration along the waterline (DP_w) and damage width (DW) for every collision scenario. Figure 7 illustrates the importance of location of impact, angle of impact, and horizontal bow angle (α) in these calculations.

To establish the distribution over the three MRRT categories given calculated damage, a response time model was developed. Structural plans of the ferries were used to estimate the damage to bulkheads given calculated damage width and penetration. In case of damage below the waterline of the ferry and damage of enough bulkheads, flooding of multiple compartments of the ferry is possible.

To help address the response time question given the potential flooding of multiple compartments, the concept of MRRT is used. In the event that the possible number of flooded compartments is lower than the design limit of the ferry, the MRRT is judged to be long. If the possible number of flooded compartments is higher than the design limit, the MRRT may be judged to be short. The analysis was conducted for each possible class of striking vessel and each possible class of ferry in order to determine MRRT categories for each possible collision scenario.

Readers interested in a more in-depth discussion of the modeling approach—for example, the treatment of the expert-judgment elicitation procedure and subsequent analysis—are referred to Technical Appendix III of Harrald, van Dorp, Mazzuchi, Merrick, and Grabowski.[11]

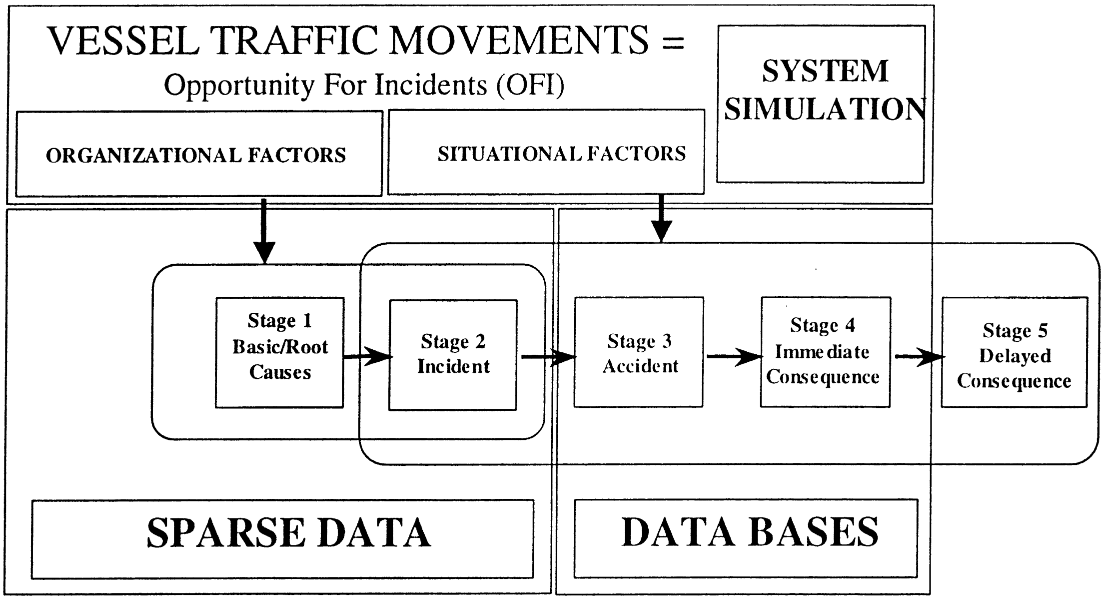


Fig. 5. Typical data availability relative to the maritime accident event chain.

3.4. Defining a Baseline Scenario

A representative simulation scenario was developed for the 11-year period for which historical data were collected. This simulation scenario (referred to as the calibration scenario) was developed for calibration purposes of the accident probability model to the historical data collected. The fall, spring, and summer sailing schedules in the last year (1997) of this 11-year period were used for the calibration scenario. These schedules are published by the WSF and comprise a full year of service. The WSF ferry schedules had remained fairly stable during this 11-year period. The WSF supplied the assignments of ferry classes to routes for the year 1997. The assignments of ferry classes to routes had remained fairly stable as well over this 11-year period. The blue-ribbon panel and WSF scheduling staff approved the use of the fall 1997, spring 1997, and summer 1997 sailing schedules and 1997 ferry class assignments for the calibration scenario.

To evaluate the risk reduction measures in Table I, a baseline level of risk needed to be established and thus a baseline scenario needed to be defined. The Washington state ferry risk assessment project started in July 1998. At this time, one high-speed ferry, the *Chinook*, had been delivered and was operating on the Seattle to Bremerton route. Two Jumbo Mark II class ferries also had started service or would start service on the Seattle to Bainbridge

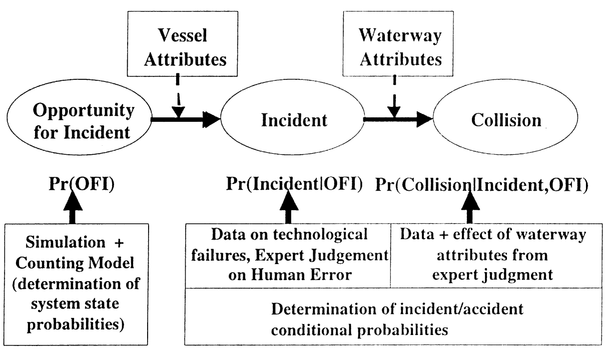


Fig. 6. Summary of modeling methodologies to establish collision risk.

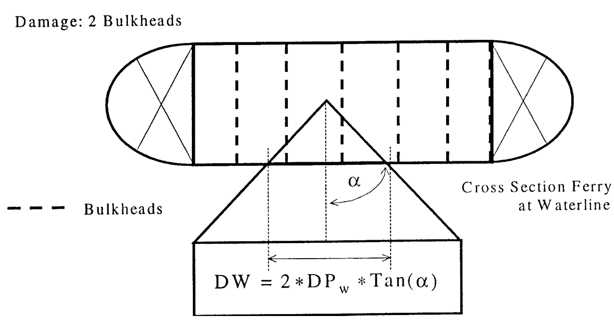


Fig. 7. Illustration of damage model calculations. DW = damage width, DP_w = damage penetration along the waterline.

Island route during 1998. The WSF schedule after the introduction of these ferries was considered the basis for the baseline scenario. Therefore, the calibration scenario was modified using 1998 schedules to represent a WSF schedule and assignments of ferries to routes after the introduction of these two new ferry classes: one high-speed ferry, the *Chinook*, and two Jumbo Mark II class ferries. The modified calibration scenario was defined as the baseline simulation scenario. The baseline simulation scenario was used to establish the baseline level of risk for risk intervention evaluation.

3.5. Modeling the Effect of Risk Interventions

The seven intervention classes described in Table I reduce accident probabilities, consequences, or both by intervening in the causal chain. The effect of a risk intervention measure may be modeled by changing model parameters from the baseline scenario. As shown in Fig. 3, some measures have an impact early on in the maritime accident event chain. Therefore, to model the effect of these risk interventions in a meaningful way, it is important that the system risk model represents events that far back in the causal chain. Rather than making worst case or best case assumptions concerning the effect of risk interventions on model parameters, the approach of reasonable assumptions following data analysis on human error in other transportation modes and mechanical-failure data of the WSF was taken, followed by sensitivity analysis.[12] The assumptions made to represent the seven intervention classes are listed in Table III. These assumptions were made in cooperation with maritime experts and were presented to and accepted by the Blue-Ribbon Panel on Ferry Safety.

4. BASELINE RISK AND RISK INTERVENTION EVALUATION RESULTS

In this section, a detailed discussion of baseline risk will be given in terms of the distribution of annual collision frequencies per year over the three MRRT categories by (1) ferry route and (2) ferry route and interacting vessel. Following the discussion of baseline risk, the effectiveness of risk intervention measures will be evaluated and presented. Results on the sensitivity analysis will be discussed as well.

4.1. Baseline Risk Results

Table IV presents the evaluated expected annual frequency of collisions per year over the three MRRT categories for the baseline scenario defined in Section 3. The average time between consecutive collisions in Table IV is the reciprocal of the statistical expected number of collisions per year.

Table IV summarizes the level of collision risk in the WSF system as a whole. The baseline statistical frequency of collisions per year, calculated using the baseline simulation, is 0.223 per year. The calibration statistical frequency of collisions per year, calculated using the calibration simulation, is 0.182 per year (equals two collisions over an 11-year period). Further analysis showed that this 22.7% increase in statistical frequency of collisions was mainly a result of replacing one of the older, slower passenger-only ferries on the Seattle–Bremerton route by the high-speed passenger-only ferry, the *Chinook*. It should be noted that the increase in statistical frequency of collisions is primarily of the 0–1 hr MRRT category due to the impact resulting from a high-speed collision with another vessel.

Table IV does not provide insight into which ferry route contributes most to the baseline level of

Table III. Summary of Modeling Effect of Risk Interventions Classes Tested

Class	Intervention	Assumed impact
1	Adopt ISM (International Safety Management) standard fleetwide	Reduce human error incidents by 30%, reduce mechanical failures by 3.7%, reduce consequences by 10%
2	Implement all mechanical-failure reduction measures fleetwide	Reduce mechanical-failure incidents by 50%
3	Implement high-speed ferry rules and procedures	Reduce human error incidents on high-speed ferries by 30%, reduce mechanical-failure incidents on high-speed ferries by 3.7%
4	Implement weather, visibility restrictions	Reduce the interactions with other vessels in bad visibility conditions by 10%
5	Implement traffic separation for high-speed ferries	Reduce interactions with high-speed ferries within 1 mile by 50%
6	Implement traffic control for deep-draft traffic	Set maximum allowable traveling speed in Admiral Inlet, north Puget Sound, central Puget Sound, and south Puget Sound at 15 knots
7	Increase time available for response	Improve response time in the 1–6 hr MRRT category by 50%

Note: MRRT = maximum required response time.

Table IV. Baseline Risk

Category	Statistical expected number of collisions per year per category	Average time between consecutive collisions per category (years)
0–1 hr MRRT	0.055	18.1
1–6 hr MRRT	0.015	67.5
>6 hr MRRT	0.152	6.6
Total	0.223	4.5

Note: MRRT = maximum required response time.

system collision risk. To further the understanding of the baseline collision risk levels, Fig. 8 shows the contribution to collision risk by ferry route. Table V gives the abbreviations used for the 13 different ferry routes displayed in Fig. 8.

In Fig. 8, the annual frequency of collisions for each route is further broken down into the three MRRT categories. Figure 8 shows that the six routes that contribute most to the level of system collision risk are (1) the Seattle to Bremerton car ferries, (2) the Seattle to Bremerton passenger ferries, (3) the Seattle to Bainbridge Island ferries, (4) the Edmonds to Kingston ferries, (5) the Fauntlerory to Vashon Island ferries, and (6) the Seattle to Vashon ferries. These routes are geographically centered around the main Seattle metropolitan area.

It cannot be concluded from the information in Fig. 8 whether the risk levels for the ferry routes are driven by (1) high numbers of interactions with other vessels, that is, traffic congestion relative to the other ferry route, (2) high collision risk per interaction, or (3) both. Hence, the next step in understanding baseline risk is to further decompose the collision risk levels by the type of vessels that the ferries interact with on a particular ferry route. The type of interacting vessel contributes both to the collision probability for each interaction and the MRRT categorization of each interaction.

The results will be presented in three-dimensional graphs displaying the collision risk levels by ferry route and interacting vessel type. The keys for these graphs are given in Table V and Table VI. Figure 9 shows the number of interactions per year by ferry route and by interacting vessel type. The higher bars to the right of the Vessel Class Index axis shows that the number of interactions is much higher with Washington state ferries (Keys 13 to 22 in Table VI) than with non-WSF vessels (Keys 1 to 12). For the Ferry Route Index axis, the highest bars are on Route indices 1 through 3. These are the Seattle to Bremerton routes and the Seattle to Bainbridge route.

Figure 10 shows the average collision probability per interaction by ferry route and interacting vessel type. The higher bars to the left of the Vessel Class Index axis (Keys 1 to 12) show that the interactions with

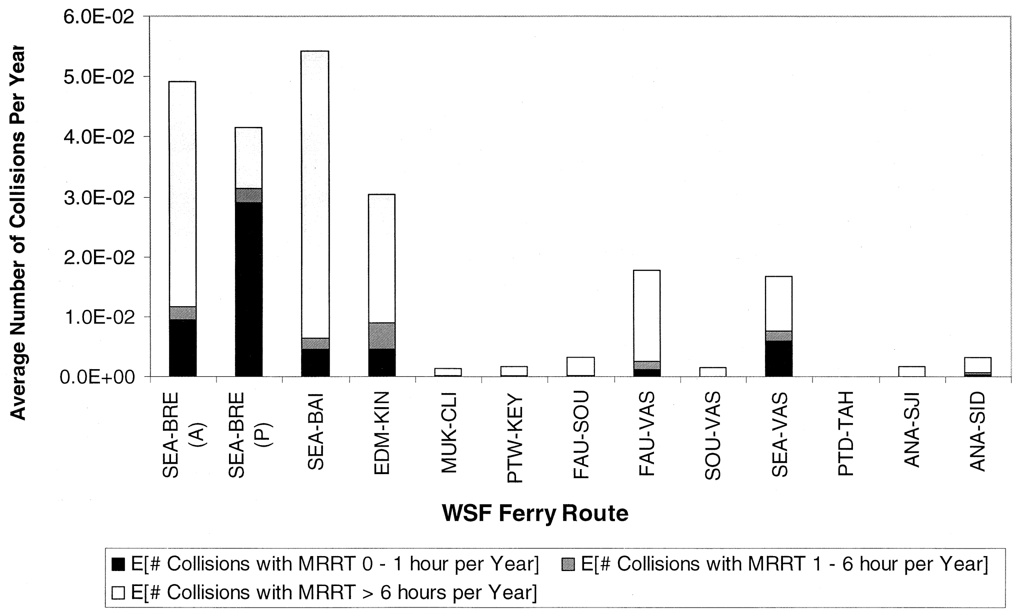


Fig. 8. Statistical expected number of collisions per year by ferry route. See Table V for abbreviations. WSF = Washington state ferries. MRRT = maximum required response time.

Table V. Numbering Keys and Abbreviations for Ferry Routes

Route index	Ferry route	Abbreviation
1	Seattle–Bremerton car ferries	SEA-BRE (A)
2	Seattle–Bremerton passenger ferries	SEA-BRE (P)
3	Seattle–Bainbridge	SEA-BAI
4	Edmonds–Kingston	EDM-KIN
5	Mukilteo–Clinton	MUK-CLI
6	Port Townsend–Keystone	PTW-KEY
7	Fauntleroy–Southworth	FAU-SOU
8	Fauntleroy–Vashon	FAU-VAS
9	Southworth–Vashon	SOU-VAS
10	Seattle–Vashon	SEA-VAS
11	Port Defiance–Tahlequah	PTD-TAH
12	Anacortes–San Juan Islands	ANA-SJI
13	Anacortes–Sidney	ANA-SID

non-WSF vessels are more likely to lead to a collision than interactions with Washington state ferries (Keys 13 to 22). Figure 11 shows the annual collision frequency by ferry route and type of interacting vessel and is a combination of the information in Figs. 9 and 10. The highest bars are on Routes 1 to 3, the Seattle–Bremerton routes and the Seattle–Bainbridge route. Overall, there are relatively high bars for the annual collision frequency for interactions with both other WSF ferries and non-WSF vessels on these routes.

From Fig. 10 it can be observed that the annual frequency of collisions with non-WSF vessels is driven by the collision probability for each interaction. From Fig. 9 it can be observed that the annual frequency of collisions with WSF ferries are driven by the number of interactions per year.

In terms of emergency response, accidents that fall in the less than 1 hr MRRT category are of particular concern. Using the damage model and the response time model, the annual collision frequencies in Fig. 11 can be filtered to include only those in the less than 1 hr MRRT category. The results are shown in Fig. 12. It can be concluded from Figs. 11 and 12 that the Seattle–Bremerton passenger-only route (Ferry Route Index Key 2) and the vessels that interact with it have a larger statistical expected number of collisions with an MRRT of less than 1 hr. The Seattle to Vashon passenger-only route (ferry Route Index Key 10) also has a relatively high annual frequency of collisions in the less than 1 hr MRRT category. The new high-speed passenger-only ferry is solely assigned to the Seattle–Bremerton passenger-only route. Collisions involving the high-speed passenger-only ferries are always assessed to require a maximum response time of less than 1 hr. The older passenger-only ferries are used for both the Seattle to Bremerton and the Seattle to Vashon passenger-only routes and interact with both large car ferries and deep-draft non-WSF vessels, as shown in Fig. 9. A proportion of the collisions of the older passenger-only ferries with large car ferries and deep-draft non-WSF vessel fall in the less than 1 hr MRRT category.

The information in Fig. 12 may be summarized in the form of a ranked cumulative risk contribution chart, as presented in Fig. 13. The ferry route and interacting vessel combinations are ordered from left to right by the percentage contribution to the statistical expected number of collisions per year. The dark part of each bar in Fig. 14 indicates the percentage contribution to the statistical expected number of collisions in the less than 1 hr MRRT category for that collision scenario. The total height of the bar indicates the cumulative percentage including all colli-

Table VI. Numbering Keys for Interacting Vessels

Vessel index	Vessel class	Vessel index	Vessel class
1	Passenger	12	Misc.
2	Tug/barge	13	Jumbo Mark II
3	Freight ship	14	Jumbo
4	Container	15	Super
5	Bulk carrier	16	Issaquah
6	Refrigerated cargo	17	Evergreen
7	Tanker	18	Steel electric
8	Product tanker	19	Rhododendron
9	Other	20	Hiyu
10	Roll-on, roll-off	21	Passenger-only vessel
11	Naval	22	Chinook

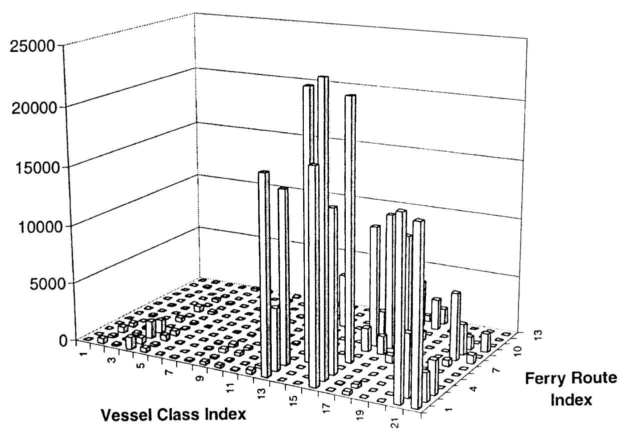


Fig. 9. Number of interactions per year by ferry route and vessel class.

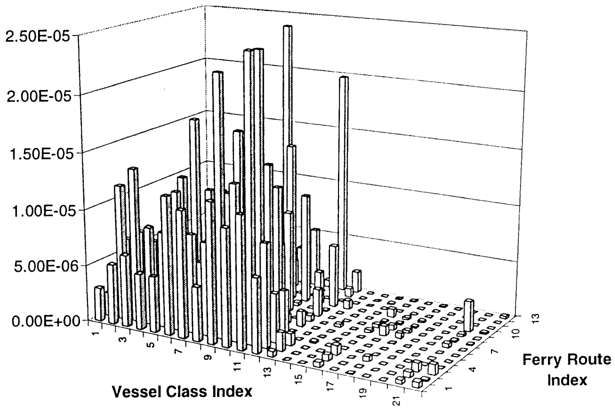


Fig. 10. Average collision probability per interaction by ferry route and vessel class.

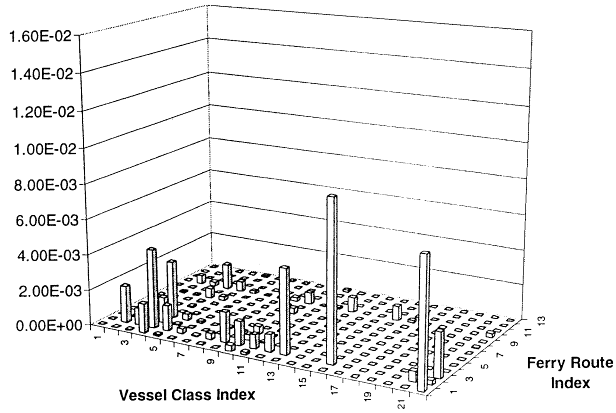


Fig. 12. Statistical expected number of collisions per year with a maximum required response time of less than 1 hr by ferry route and vessel class.

sion scenarios to the left. In other words, Fig. 13 contains the top collision scenarios that accumulate to 88% of the statistical expected number of collisions per year in the less than 1 hr MRRT category.

4.2. Evaluation of Risk Interventions

All cases were tested to evaluate their effect on the annual frequency of collisions and on the annual frequency of collisions in each of the MRRT categories. The results of these analyses are represented in Fig. 14. For each risk intervention class, the total percentage reduction in the statistical frequency of collisions is comprised of the percentage reduction in the statistical frequency of collisions in each of the three MRRT categories relative to the baseline scenario in Table IV.

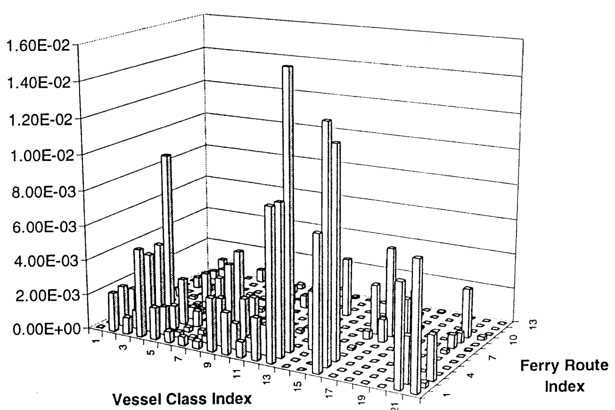


Fig. 11. Statistical expected number of collisions per year by ferry route and vessel class.

Case 1 has the largest risk reduction at 16% and reflects the effect of the fleetwide implementation of the International Safety Management (ISM) code. Noted is a large reduction for both the less than 1 hr and the more than 6 hr MRRT categories. Case 2, the implementation of mechanical-failure-reducing measures, is the next most effective at 11%. Of note is a large reduction in each MRRT category as well as the large reduction predicted for collisions with a MRRT of 1 to 6 hr. The implementation of traffic separation rules for the high-speed ferries, Case 5, causes a 6% reduction in the total statistical expected number of collisions. As this reduces the statistical expected number of collisions involving high-speed ferries, all this reduction is for collisions with an MRRT of less than 1 hr. A 5% reduction in the total statistical expected number of collisions is predicted for the implementation of visibility restrictions, Case 4. The implementation of high-speed ferry rules (ISM restricted to high-speed ferry routes), Case 3, decreases the total statistical expected number of collisions by 2%, with all the reduction being for collisions with an MRRT of less than 1 hr. Case 7 is aimed at reducing the consequences if a collision occurs, not the probability of occurrence. This case reflects the implementation of procedures to evacuate passengers to a safe haven in the event of collision with an MRRT of 1 to 6 hr—survival craft. Reducing the speed of commercial vessels in Puget Sound, Case 6, also does not reduce the total statistical expected number of collisions. The statistical expected numbers of collisions with an MRRT of less than 1 hr and an MRRT of 1 to 6 hr are both reduced, however, while the statistical expected number of collisions with an MRRT of more than 6 hr increased by the same amount.

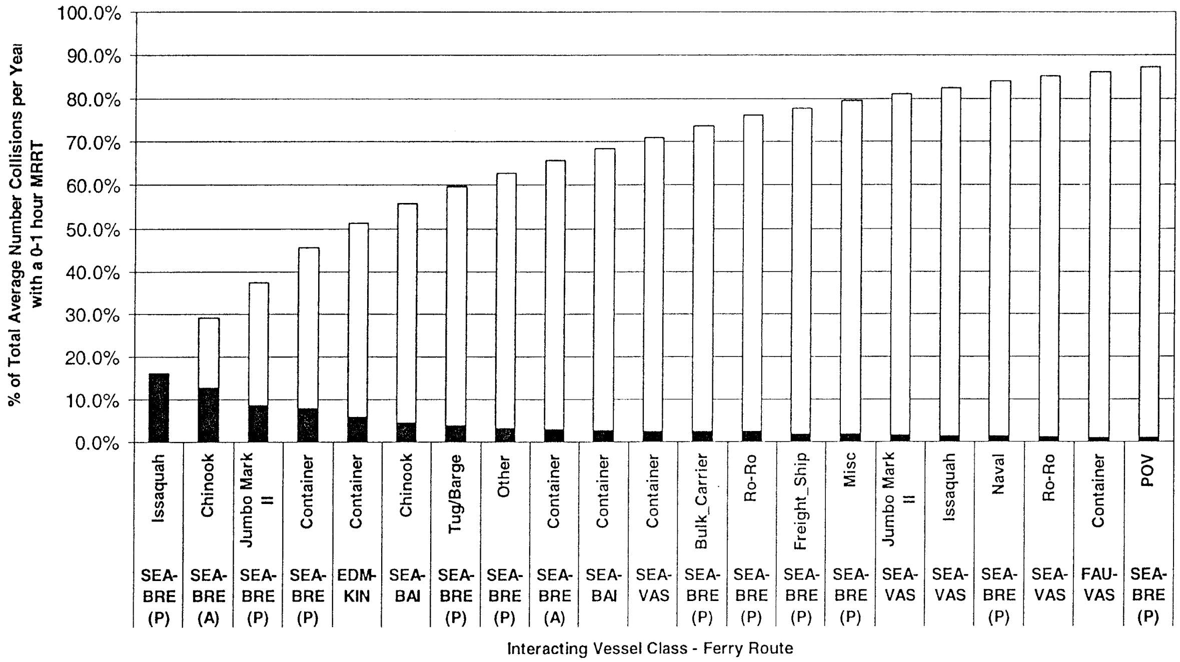


Fig. 13. Distribution of the statistical expected number of collisions per year with a maximum required response time (MRRT) of less than 1 hr by ferry route and vessel class. See Table V for abbreviations.

5. SENSITIVITY ANALYSIS RESULTS

The analysis of the WSF risk assessment provides the basis for determining how the risk in the system could be reduced to even lower levels. The findings of a quantitative study must be interpreted with care, however, as uncertainty is introduced at various levels of the analysis. Sources of this uncertainty include incomplete or inaccurate data, biased or uninformed expert judgment, modeling error, and computational error. Testing for the level of uncertainty in an analysis requires accounting for both parameter uncertainty and model uncertainty and their impact on the results and conclusions. This is referred to as an "uncertainty analysis."[12]

While the use of proper procedures such as rigorous data selection and cross validation—structured and proven elicitation methods for expert judgment and use of accepted models—can reduce uncertainty and bias in an analysis, it can never be fully eliminated. The reader should recognize that the value of an analysis is not only in the precision of the results, but also in the understanding of the system. Of great value is the identification of peaks, patterns, unusual circumstances and trends in system risk, and changes in system risk through risk mitigation measure implementation.

The methodology in this study has been reviewed for rigor and tested in operational settings.[13] The methodology thus provides many safeguards to remove bias and to detect error. The general approach toward modeling assumptions in the WSF risk assessment was that of reasonableness rather than pursuing one worst case assumption after the other. The latter approach may lead to risk assessment results related to highly unlikely scenarios and therefore less-useful results. The approach of using reasonable assumptions rather than worst case assumptions is supported by scientists in the field of risk analysis.[12]

Although a formal uncertainty analysis has not been presented with these results, sensitivity of the results to some of the more contentious modeling assumptions has been tested. The assumptions tested/challenged through the sensitivity cases were

1. All collisions involving a high-speed ferry fall in the category of collision with an MRRT of 0–1 hr
2. The vertical bow angle reduces the damage penetration below the waterline
3. The horizontal bow angle for vessels in the WSF system is, on average, 66°

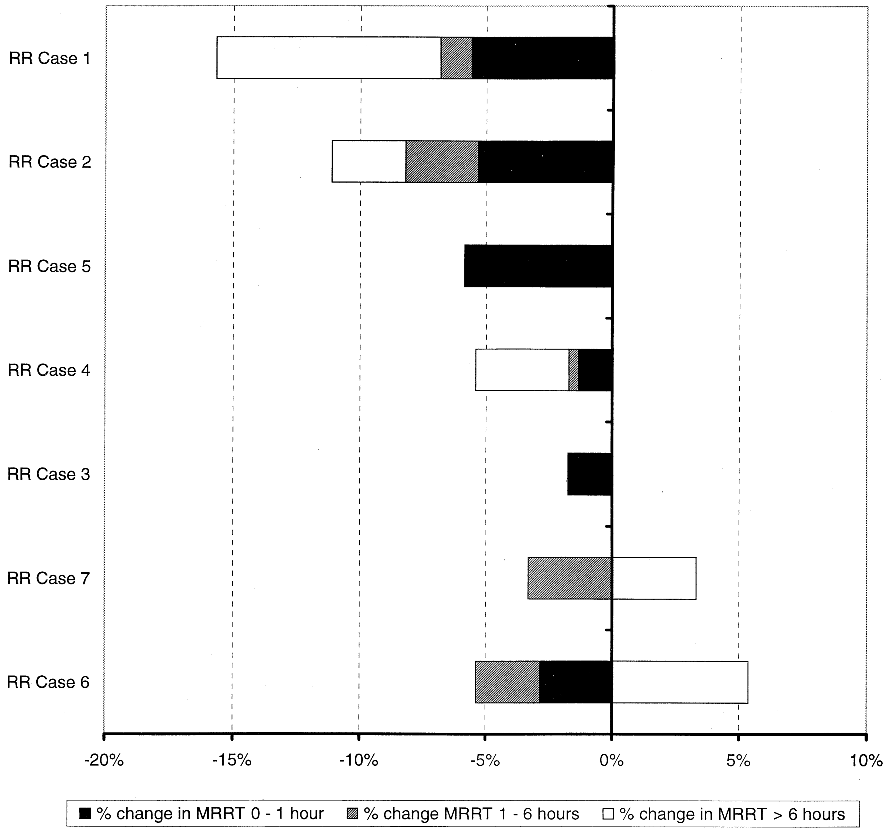


Fig. 14. Estimated risk reduction (RR) for the seven tested cases. MRRT = maximum required response time.

4. The collision speed for non-WSF vessels is 80% of the traveling speed, and the collision speed of WSF vessels is 50% of the traveling speed
5. The relative depth penetration (RDP = percentage damage penetration relative to the beam of the WSF-ferry) threshold beyond which the RDP determines the distribution of collisions over the three MRRT categories is 50%
6. The steel electric vessel has parts that satisfy one-compartment vessel characteristics and two compartment vessel characteristics

To test these six assumptions, nine sensitivity cases were developed and analyzed. For demonstrative purposes, the first listed assumption (Assumption 1) is that all collisions involving the new high-speed passenger-only ferries fall in the less than 1-hr MRRT category. This assumption was modified so that all three MRRT categories are equally likely in case of a collision involving the high-speed passenger-only ferry and is henceforth referred to as Sensitivity Case 1. This assumption is more optimistic than Assumption 1. The results of the sensitivity analysis are shown in Fig. 15.

Figure 15 shows that the statistical frequency of collisions in the less than 1 hr MRRT category reduces by 9% in Sensitivity Case 1. Also of note is that the combined percentage increase in statistical frequency of collisions in the 1–6 hr MRRT category and more than 6 hr MRRT category equals the percentage reduction in the less than 1 hr MRRT category. In other words, the effect of the modified assumption is a redistribution of the total statistical frequency of collisions over the three different MRRT categories. The same observation can be made for all the other sensitivity cases tested as well.

Figure 16 summarizes the collision analysis by ferry route under Sensitivity Case 1. Comparing Figs. 8 and 16, it can be observed that by altering Assumption 1 the statistical frequency of collisions in the less than 1 hr MRRT category has primarily been reduced on the Seattle Bremerton passenger ferries, Seattle

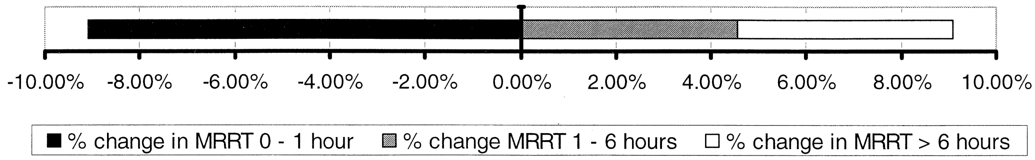


Fig. 15. Percent change in the annual collision frequency in each maximum required response time (MRRT) category under Sensitivity Case 1.

Bremerton car ferries, and the Seattle Bainbridge ferries. The predominant WSF ferry routes in terms of the statistical frequency of collisions in the less than 1 hr MRRT category, however, are the same under the original assumption and the modified assumption for high-speed passenger-only ferries. Similar conclusions can be drawn when analyzing these results for the other sensitivity cases as well.

6. GENERAL CONCLUSIONS

Sixteen specific risk reduction recommendations are cited in Harrald *et al.*[11] Recommendations derived from the analysis were divided into three categories: (1) general risk management recommendations for the Washington state ferries to manage risk in the system, (2) recommendations for reducing the likelihood of accidents, and (3) recommendations for minimizing the potential consequences of accidents. Interested readers are referred to Harrald *et al.*[11] for the specific recommendations. Below are general conclusions in terms of the previous three categories of risk management recommendations.

In terms of general risk management, it was recommended that the Washington state ferries should improve their capabilities to detect and manage risk and to prepare for potential emergencies. This requires a continuing set of systems, capabilities, and structures in order to be effective. Maintaining and enhancing safety in the WSF system requires management and resources devoted to risk prevention, accident response, and consequence management. The WSF risk assessment report supports the currently planned and funded fleetwide implementation of the ISM system.

In terms of reducing the likelihood of accidents, it was recommended that the WSF should continue to implement safety management and training programs, provide adequate relief crews as necessary to accomplish training, and coordinate with the USCG to minimize the likelihood of an accident. It was

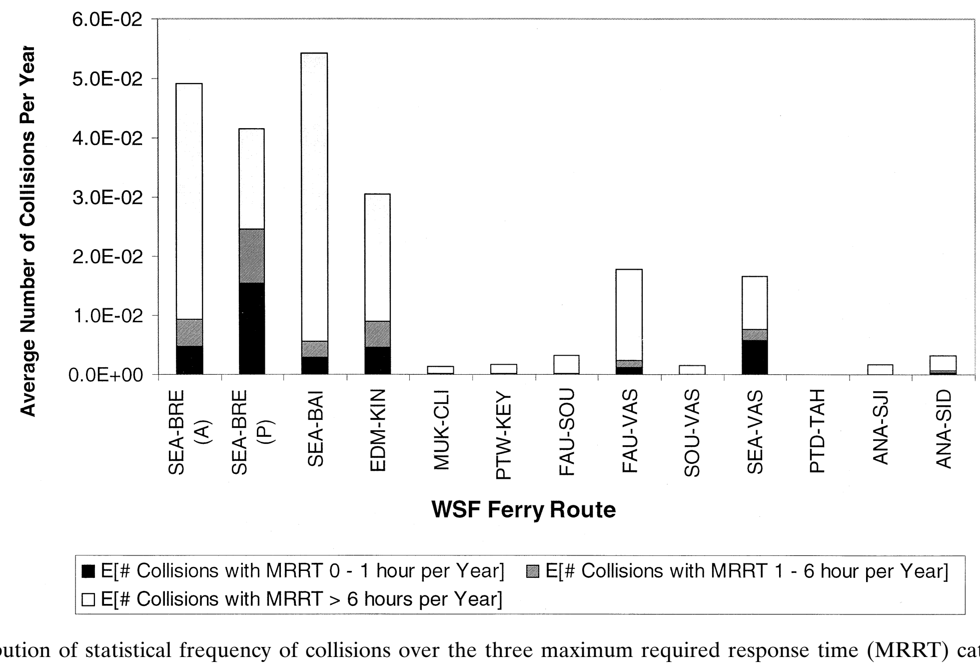


Fig. 16. Distribution of statistical frequency of collisions over the three maximum required response time (MRRT) categories by ferry route—Sensitivity Case 1. See Table V for abbreviations.

noted that since the consequences of an intentional act of destruction (sabotage or attack) aboard a ferry could be severe, the WSF should work with the Washington State Patrol and federal agencies to determine the need for additional security measures to combat the threat of intentional acts of destruction aboard ferries.

In terms of minimizing the potential consequences of accidents, it was recommended that the WSF, the USCG, and other response organizations should work collaboratively to ensure that consequences will be minimized for any accident that does occur. Specifically, it strongly recommends that the WSF and the USCG and other public safety agencies address the problem of minimizing injury and loss of life from very low-probability but potentially high-consequence accidents through planning, implementing, and exercising adequate response plans and procedures. It recognizes that the skills of the ferry crew will be crucial in any emergency situation and strongly recommends enhancing these emergency skills through training, certification, drills, and exercises.

The report finally concludes that the most cost-effective way to minimize the risk of potential accidents is to invest in WSF people and systems and to make improvements and changes to WSF policies, procedures, and management systems—rather than to merely invest in capital equipment such as survival craft. The creation of a safety culture that will enable these recommendations to be realized will require the support and leadership of WSF management; shoreside operations; and fleet deck officers, engineers, and other shipboard personnel.

The conclusions and recommendations made to the WSF were driven by the total statistical frequency of collisions and by the distribution of the total statistical frequency of collisions over the three MRRT categories. Based on the results of the sensitivity analysis performed, it was concluded that the conclusions and recommendations made were robust relative to the modified assumptions tested.

As a closing note, it might be of interest to mention that it is impossible for any risk analysis performed in a dynamic public arena to foresee changes as a result of political processes. An example is the passage of Initiative 695, which eliminated the state motor vehicle excise tax. The effect for the WSF is a disproportional loss in operating and capital budget potentially impacting the level at which recommendations from this study will be implemented. Loss of operating budget already temporarily interrupted the service of two high-speed ferries, the *Chinook* and the *Snohomish*. The current legislative plan, includes funding to maintain the operations of the *Chinook* and *Snohomish*. A simulation scenario including two high-speed ferries in the WSF schedules was analyzed in the WSF risk assessment report as well. For detailed results interested readers are referred to the WSF risk assessment report in Harrald *et al.*[11]

ACKNOWLEDGMENTS

The authors would like to thank the associate editor and the referees for their most helpful comments, which benefited this article in quality of content and readability.

REFERENCES

1. Washington State Ferries. (2000). http://www.wsdot.wa.gov/ferries/info_desk/route-maps/. (9 January 2001).
2. Lifesaving Systems for Certain Inspected Vessels, 46 C.F.R. 199, Subchapter W (1 October 1998).
3. Implementation of the 1995 Amendments to the International Convention on Standards of Training, Certification and Watchkeeping for Seafarers, 1978 (STCW). 62 Fed. Reg. 123, 34506–34541.
4. International Maritime Organization, Subcommittee on Ship Design and Equipment. (1998). Revision of the HSC Code, Formal Safety Assessment of High Speed Catamaran (HSC) Ferries Submitted by the United Kingdom (41st Session, Agenda Item 5), E/DE/41/5–6.
5. Harrald, J. R., Mazzuchi, T. A., Merrick, J. R. W., Shrestha, S. K., Spahn, J. E., & van Dorp, J. R. (1996). *Final report: Prince William Sound risk assessment.* Submitted to the Prince William Sound Risk Assessment Steering Committee, Anchorage, AK.
6. Dyer, M. D., Schwenk, J., Watros, G., & Boniface, D. (1997). *Scoping risk assessment, protection against oil spills in the maritime waters of northwest Washington State.* Cambridge, MA: John Volpe National Transportation Systems Center.
7. Cooke, R. M. (1991). *Experts in uncertainty: Expert opinion and subjective probability in science.* New York: Oxford University Press.
8. Paté-Cornell, M. E. (1996). Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering and System Safety, 54*(2–3), 95–111.
9. Harrald, J. R., Mazzuchi, T. A., & Stone, S. M. (1992). Risky business: Can we believe port risk assessments? In D. Torseth (Ed.), *Ports '92 Conference Proceedings* (pp. 657–669). New York: American Society of Civil Engineers.
10. Minorsky, V. U. (1959). An analysis of ship collisions with reference to protection of nuclear power plants. *Journal of Ship Research, 3*, 1–4.
11. Harrald, J. R., van Dorp, J. R., Mazzuchi, T. A., Merrick J. R. W., & Grabowski, M. (1999). Final report: Washington state ferry risk assessment. Submitted to the Washington State Transportation Commission, Seattle, WA.
12. Hattis, D., & Anderson, E. L. (1999). What should be the implications of uncertainty, variability, and inherent "biases"/"conservatism" for risk management decision making. *Risk Analysis, 19*, 95–107.
13. Marine Board, National Research Council. (1998). *Review of the Prince William Sound, Alaska, risk assessment study.* Washington, DC: Committee on Risk Assessment and Management of Marine Systems, National Academy Press.

SUB-APPENDIX:

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.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Reliability Engineering and System Safety 81 (2003) 119–132
RELIABILITY ENGINEERING & SYSTEM SAFETY
www.elsevier.com/locate/ress

A traffic density analysis of proposed ferry service expansion in San Francisco Bay using a maritime simulation model

Jason R.W. Merrick[a,*], J. Rene van Dorp[b], Joseph P. Blackford[b], Gregory L. Shaw[b], Jack Harrald[b], Thomas A. Mazzuchi[b]

[a]*Department of Statistical Sciences and Operations Research, Virginia Commonwealth University, P.O. Box 843083, 1001 West Main St, Richmond, VA 23284, USA*

[b]*Department of Engineering Management and Systems Engineering, The George Washington University, 1776 G St NW, Suite 110, Washington, DC 20052, USA*

Received 24 November 2002; accepted 12 February 2003

Abstract

A proposal has been made to the California legislature to dramatically increase the frequency and coverage of ferry service in the San Francisco Bay area. A major question in the approval process is the effect of this expansion on the level of congestion on the waterway and the effect this will have on the safety of vessels in the area. A simulation model was created to estimate the number of vessel interactions in the current system and their increases caused by three alternative expansion plans. The output of the simulation model is a geographic profile showing the frequency of vessel interactions across the study area, thus representing the level of congestion under each alternative. Comparing these geographic interaction profiles to a similar one generated for the current ferry service in the San Francisco Bay allows evaluation of the increase in exposure of ferries to adverse conditions, such as, for example, the interaction of high-speed ferries in restricted visibility conditions. This analysis has been submitted to the legislature as part of the overall assessment of the proposal and will be used in the expansion decision.

Keywords: Maritime transportation; Simulation; Safety; Accident prevention

1. Introduction

In an effort to relieve congestion on freeways, the state of California is proposing to expand ferry operations on San Francisco (SF) Bay by (1) phasing in up to 100 ferries in addition to the 14 currently operating, (2) extending the hours of operation of the ferries, (3) increasing the number of crossings, and (4) employing some high-speed vessels. The state of California has directed the SF Bay Area Water Transit Authority (WTA) to produce an *Implementation and Operations Plan*, part of which requires working with the US Coast Guard (USCG), the California Maritime Academy, and SF Bay Area ferry operators in preparing a 'plan for ensuring safety of vessel operations traveling on the SF Bay.' The purpose of this plan is to realistically evaluate the levels of safety relative to various aspects of ferry operation.

In the process of developing the safety plan the WTA used data from the Federal Transit Administration National Transit database to describe the current safety level. Federal databases describe the past safety performance of the existing ferry services. Between 1996 and 2000, ferry service appeared to be the safest federally subsidized transit mode in the SF Bay Area. The WTA's comparison showed that ferry transportation had: (1) no fatalities for patrons, employees, or others (i.e. bystanders). The average for the rail and roadway transit modes was 0.004 fatalities per 1,000,000 passenger miles; (2) less than one-fourth the patron injury rate of the rail and roadway transit modes. Ferry operations averaged 0.28 injuries per 1,000,000 passenger miles; (3) about two-thirds the bystander injury rate of the rail and roadway transit modes. Ferry operations averaged 1.5 injuries per 1,000,000 vehicle miles; (4) on average 5.6 reported accidents per 100,000 transits, or 3.8 reported accidents per year for the 10-year period from 1992 to 2001; this is in line with the rates for similar marine transportation systems.

* Corresponding author. Tel.: +1-804-828-1301x136; fax: +1-804-828-8785.

E-mail address: jrmerric@vcu.edu (J.R.W. Merrick).

0951-8320/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0951-8320(03)00054-1

The WTA safety plan further documents a wide range of risks and associated risk controls. For risks and necessary risk controls that are already documented in codes, standards, and regulations, the plan provides a very brief overview. In conclusion, the safety plan indicates that analysis of the existing ferry services show that those services provide safe transit and are currently effectively managing risks. However, the question remains whether this 'safe' operation can continue with the new pressures of aggressive service expansion. The three proposed expansion scenarios are: (1) Alternative 3: Enhanced Existing System; (2) Alternative 2: Robust Water Transit System and (3) Alternative 1: Aggressive Water Transit System. From these, Alternative 3 is the least aggressive expansion scenario and Alternative 1 is the most aggressive one. The WTA tasked the author's to investigate the impact of ferry service expansion on maritime traffic congestion in the SF Bay area by developing a maritime simulation model of the SF Bay. Due to time and budget constraints a full-scale risk assessment, such as the authors' previous work in the Prince William Sound Risk Assessment [1–3] or the Washington State Ferries Risk Assessment [4,5], was not feasible. In these studies, a simulation of the traffic and weather patterns was used to count interactions between the vessels and an expert judgment based accident probability model was used to estimate the likelihood of a collision if such an interaction occurs. Instead, to assess the impact of aggressive ferry expansion, the scope of the SF Bay study was limited to the simulation part of the model, leaving the accident probability part to a later project if the expansion proposal is approved.

Limiting the scope of the analysis to interactions, however, will still allow meaningful conclusions regarding potential effect of the ferry service expansions on observed collision rates. In fact, interactions are known to be one of the drivers in collision risk [5]; an increase in interactions will typically result in an increase in collision risk if additional risk interventions are not put in place. The purpose of the simulation is to assess the interactions of vessels in the current ferry system and to compare their geographic profile to the interactions seen under the proposed scenarios. For instance, if the daily volume of ferry transits increases 10-fold does the number of interactions increase 10-fold? Is it possible that, since the proposed alternatives include new routes to new areas of the SF Bay, the additional interactions are distributed in such a manner that no additional high-traffic density areas occur that could indicate safety problems? Due to its unique visibility conditions, one of the main safety concerns in the SF Bay is transiting through restricted visibility. If there are additional high-traffic density areas, do they perhaps occur in restricted visibility conditions? The simulation study in this paper attempted to answer such critical safety questions.

An outline of the paper is as follows. Previous work in maritime risk assessment and simulation are discussed in Section 2. Sections 3–5 discuss the construction of the simulation, specifically the interaction-counting model in Section 3, vessel movements in Section 4 and restricted visibility modeling in Section 5. The results of the study are outlined in Section 6. Conclusions and recommendations are presented in Section 7.

2. Literature review

The National Research Council has repeatedly identified the assessment and management of risk in maritime transportation as an important problem domain [6–9]. In earlier work, researchers concentrated on assessing the safety of individual vessels or marine structures, such as nuclear powered vessels [10], vessels transporting liquefied natural gas [11], and offshore oil and gas platforms [12]. The USCG has used a classical statistical analysis of nationwide accident data to prioritize federal spending to improve port infrastructures [13,14]. More recently, researchers have used probabilistic risk assessment (PRA) [15] in the maritime domain [16–23] by examining risk in the context of maritime transportation systems (MTS) [9].

In a MTS, traffic patterns change over time in a complex manner. Researchers have used system simulation as a modeling tool to assess MTS service levels [24], to perform logistical analysis [25], and to facilitate the design of ports [26]. The dynamic nature of traffic patterns and other situational variables, such as wind, visibility, and ice conditions, mean that risk levels change over time. Recent PRAs [27] in the maritime domain have used simulation to model the dynamic nature of the transportation system.

The Prince William Sound Risk Assessment [1–3] used a simulation of the oil transportation system to evaluate changes in the dynamic pattern of traffic caused by proposed risk intervention measures, such as weather-based closure conditions for certain parts of the transit and modifications to the tug escort service put in place to save disabled tankers from running aground. Accident and incident data was augmented using expert judgment to take the simulations interaction counts and arrive at estimates of accident frequency and the expected volume of oil outflow. The Washington State Ferries Risk Assessment [4,5] used an improved version of the technique, but with the consequence of interest being passenger safety rather than environmental damage.

As mentioned previously, the study in this paper used the simulation part of this approach to only assess the impact of ferry expansion on the level of vessel interactions in the Bay. If the expansion proposal is approved, the simulation analysis can be extended to a full PRA through an accident probability model based on available accident, incident data and expert judgments.

3. The simulation: interaction counting model

In the simulation program, a snapshot of the simulation is taken every minute; counts of the interactions are taken and

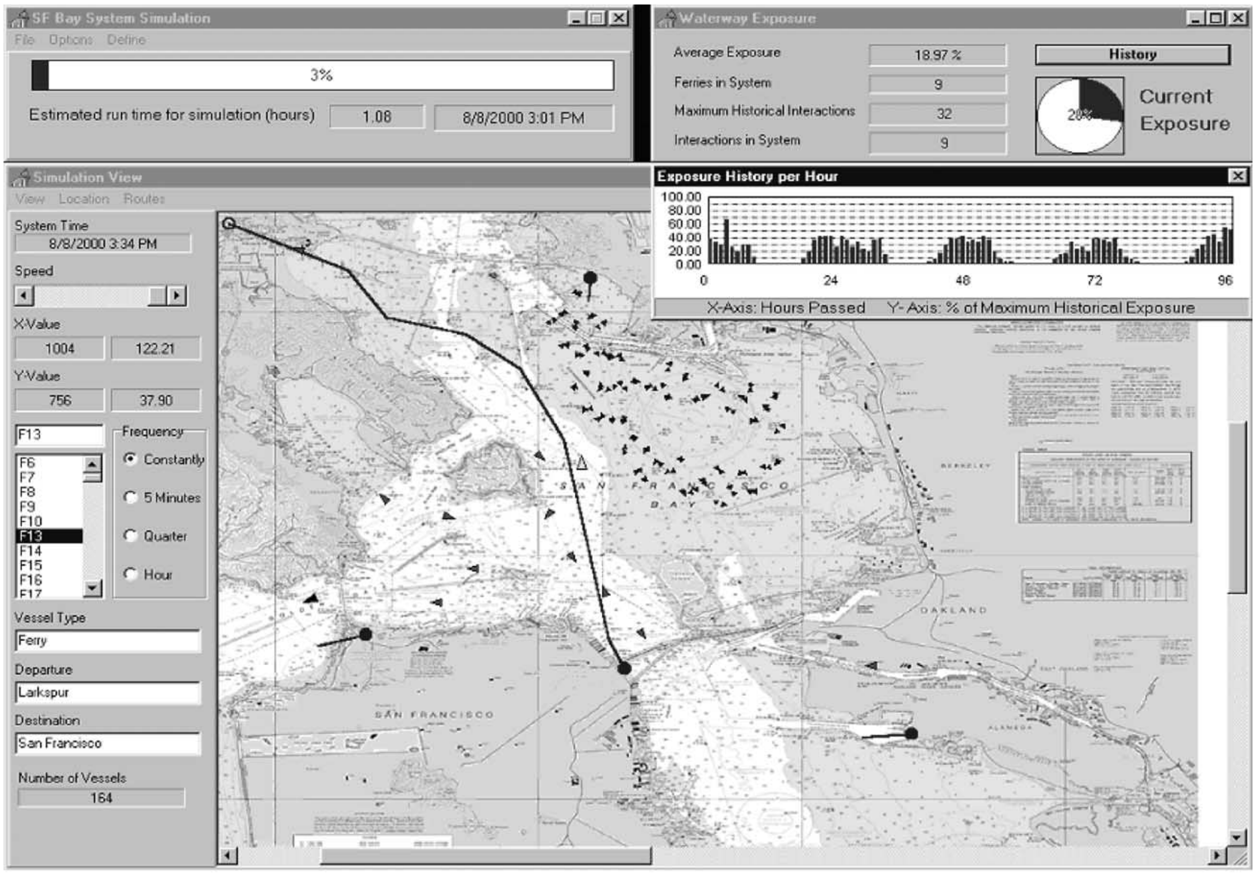


Fig. 1. A snapshot of the SF Bay maritime simulation model.

recorded in an event database. Fig. 1 shows such a snapshot of the SF Bay maritime simulation. Moving boats are represented by the triangles. Which pairs of vessels are interacting? This depends on both the distance between the vessels the time until the vessels meet.

The interaction model is based on Closest Point of Approach type arguments and stems from the considerations that a ferry captain will make when considering interactions with other vessels. For example, vessels close in at different speeds, thus in evaluating a situation involving other vessels, a captain is interested in which will arrive first, not necessarily which is closest.

Consider a ferry transiting through the system. As a default, any other vessel within a half a nautical mile[1] of the ferry is counted as interacting; half a nautical mile is too close for comfort to most professional mariners. If another vessel is more than half a mile away and in addition is more than five minutes away from crossing the track of the ferry, it is not counted as an interaction. If a vessel is within five minutes of crossing the ferry track and in addition this crossing will occur within one nautical mile in front of the ferry or within half a mile behind the ferry, the vessel is counted as interacting with ferry. Experts with maritime experience outside the ferry service and a group of ferry captains from the Washington State Ferry Service provided input for this methodology [5,28].

The snapshot of the simulation at a specific time is analyzed to determine whether the ferries in the system are interacting with other vessels (including other ferries) using the interaction model above. For each interaction found, the information about the type of the other vessel, the type of interaction (crossing, meeting or passing), the visibility conditions and the coordinates of the vessels are recorded and written to an interaction database. This database is then used to find the number of interactions occurring in a simulation run in each of a grid of cells across the SF Bay.

This information can then be represented in the form of a colored map, with the colors representing the number of interactions in each cell of our grid. This map may be interpreted as a geographic profile of ferry interactions. The color gradient for the grid cells is established using a simulation of the current ferry service on the SF Bay (to be referred to as the Base Case). The Base Case analysis allows existing trouble spots to be identified, thereby not attributing these to the planned ferry service expansions. Next, using the Base Case color scale, similar geographic profiles can be generated for these expansions. Emerging hot spots resulting from the expansions can be visually observed by

[1] One nautical mile equals approximately 1.15 miles.

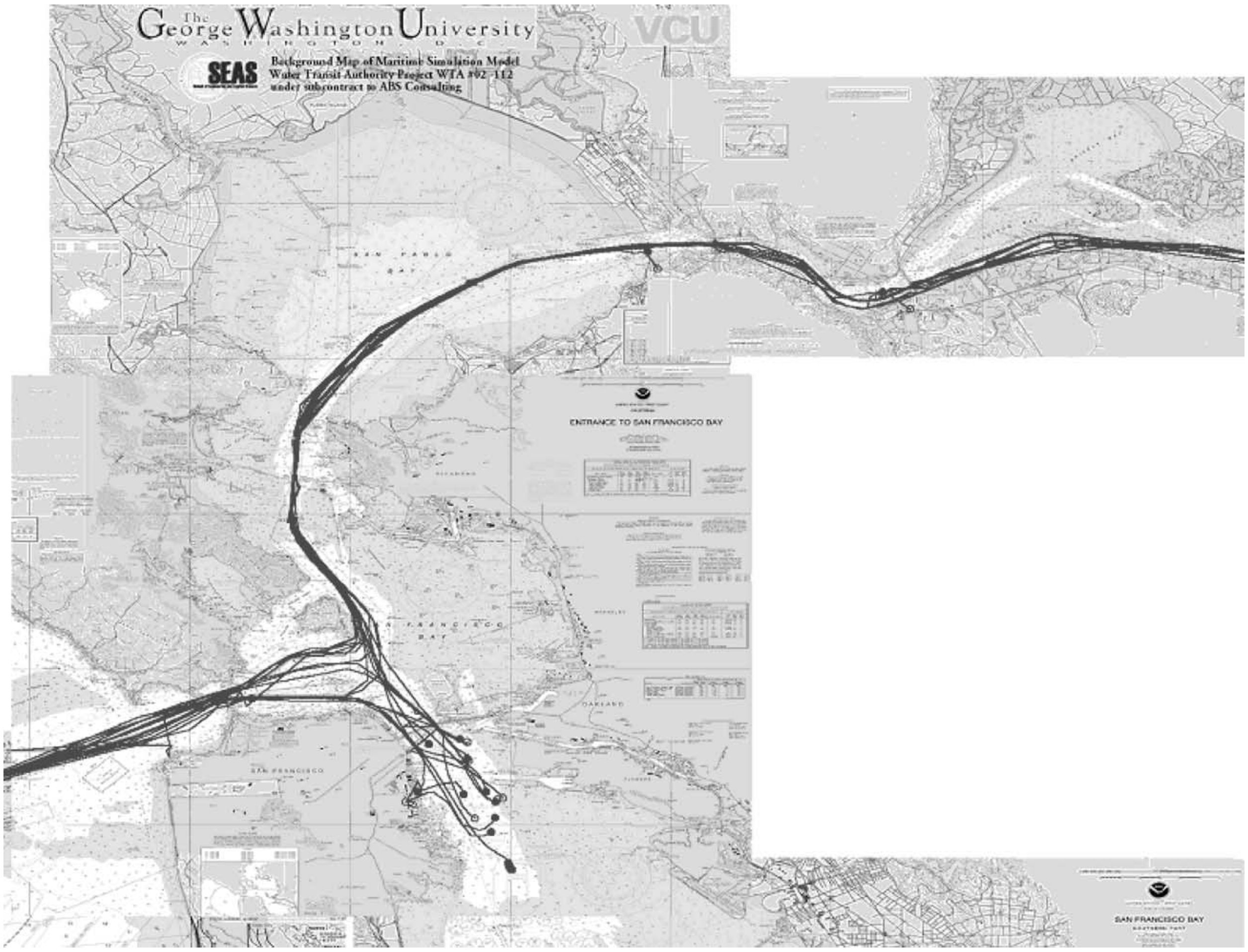


Fig. 2. Vessel routes for LPG vessels in the SF Bay maritime simulation model.

comparing their geographic profile to that of the Base Case. For further discussion of the interaction-counting model, see Ref. [27].

4. The simulation: vessel movement

To achieve an accurate count of the number of interactions, we must have an accurate simulation of the vessel movements. This means we need an accurate background map of the Bay, an accurate representation of the movement of the ferries themselves and an accurate representation of the movements of the other vessels in the Bay. The background map of the maritime simulation model for the SF Bay area (Fig. 1) was constructed from NOAA electronic charts, which were converted to bitmaps for use with the simulation program. This allowed accurate representation of the vessel coordinates and speed.

Ferry movements for the base case simulation were obtained from ferry schedules collected from ferry operators for the years 1998–2001. Each proposal for expansion of the ferry service included the number of transits per day, the time between transits, and the start time. At the current stage of the proposed expansions, the schedules are simply defined by operations starting at 6 a.m. and running every 15, 30, or 60 min depending on the route.

The ferry routes configurations for the base case simulation and proposed expansions were obtained from GIS maps created by the URS Corporation for the WTA. In all, 18 ferry routes were considered for the base case simulation and up to 64 ferry routes for the proposed expansion alternatives. The cruising speed of each ferry class along their route is a known, constant speed when underway. The ferries slow down when leaving and entering dock. Ferries also slow in restricted visibility. Ferries that usually maintain between 25 and 35 knots will reduce speed to 12 knots. Slower excursion ferries will slow to 10 knots. These speeds were determined in discussions with ferry captains and were confirmed by the ferry companies. To reflect this behavior in the simulation model, restricted visibility needs to be represented adequately. The modeling of visibility conditions in the simulation is discussed in Section 5.

In building maritime simulation models, non-ferry traffic is usually modeled by analyzing traffic arrival/departure data to construct probability distributions for vessel inter-arrival times. These distributions are then used to simulate vessel arrivals and transits in the system [27]. However, the presence of the San Francisco Vessel Traffic System (SF VTS) eliminated the need for this approach. Data on date, time, and transits for 6000 routes for up to 26 different vessel types were obtained from the VTS for the 1998–2001 period. Waypoint data obtained from the SF VTS was used in conjunction with the bitmap of the SF Bay area to produce the total vessel transit picture. Fig. 2 shows an example of the routes of a particular class of vessels. Again average vessel speeds for each class are maintained during transits with the exception of vessels slowing down in restricted visibility. Average vessel speed information was obtained through personal communication with SF Bar Pilots. In restricted visibility, deep-draft traffic slows to about 70% of its usual transit speed. This rule was determined by discussions with members of the SF Bay Pilot's Association and operators from the VTS. These databases of traffic arrivals and routes were read in to the simulation program, removing the problem of validation of arrivals models [28].

Unfortunately, the SF VTS does not routinely record the movements of small vessels such as recreational yachts. As at certain times this can be the most numerous type of traffic on the Bay, special events, such as regattas, were modeled in the simulation as well. The USCG supplied their Marine Event List for over 1000 special events for the year 2001. Due to time and budget constraints only the main type of special events were modeled in the maritime simulation, i.e. 828 scheduled regattas in 2001. The data on regatta times and areas were obtained from the USCG data. Through discussions with the SF VTS, 13 locations were defined for these regatta events. Regattas were modeled by blocking the defined areas (Fig. 3) during their times and dates and then randomly moving the assigned number of participating vessels within each area.

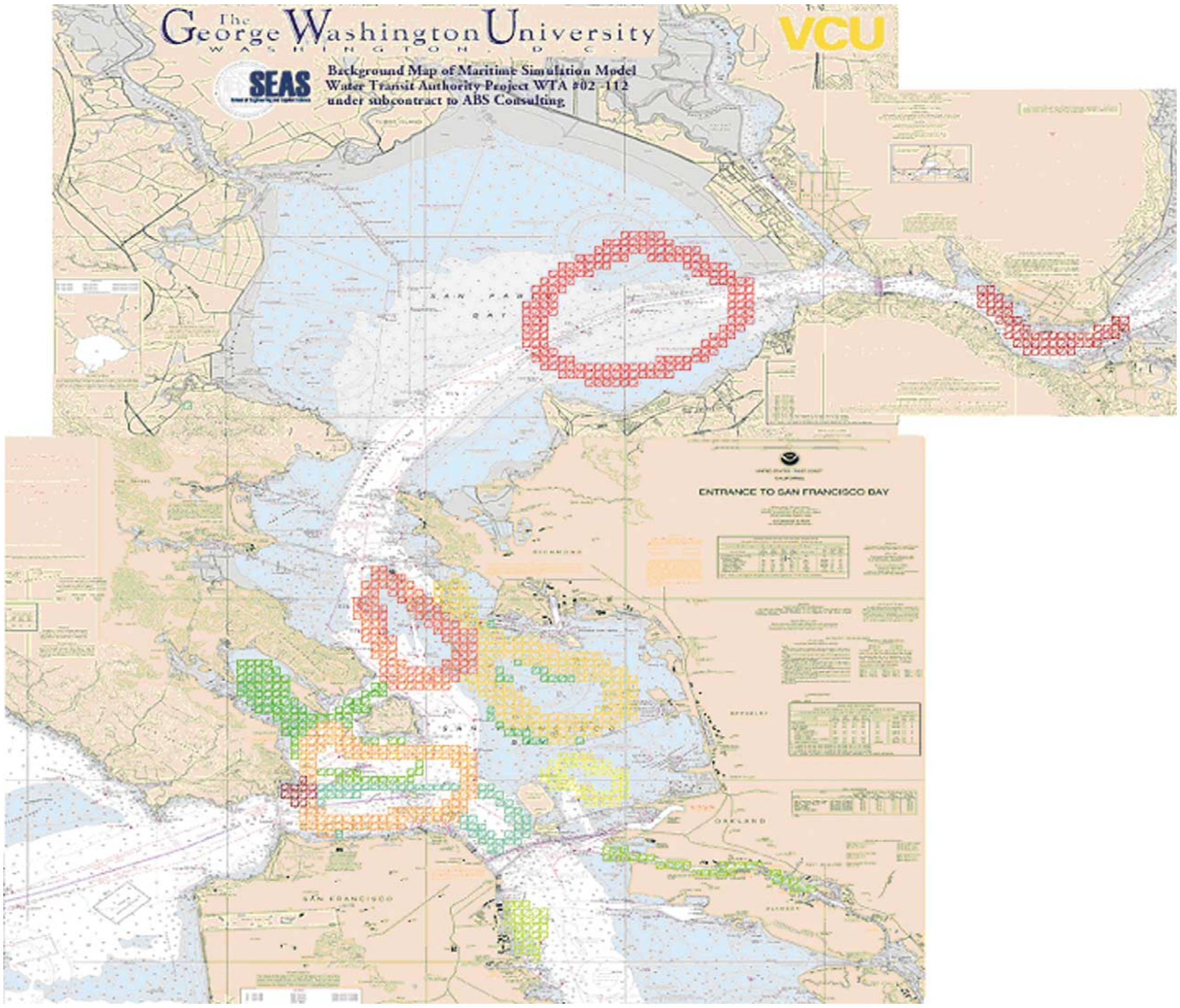


Fig. 3. Definition of regatta locations in the SF Bay maritime simulation model.

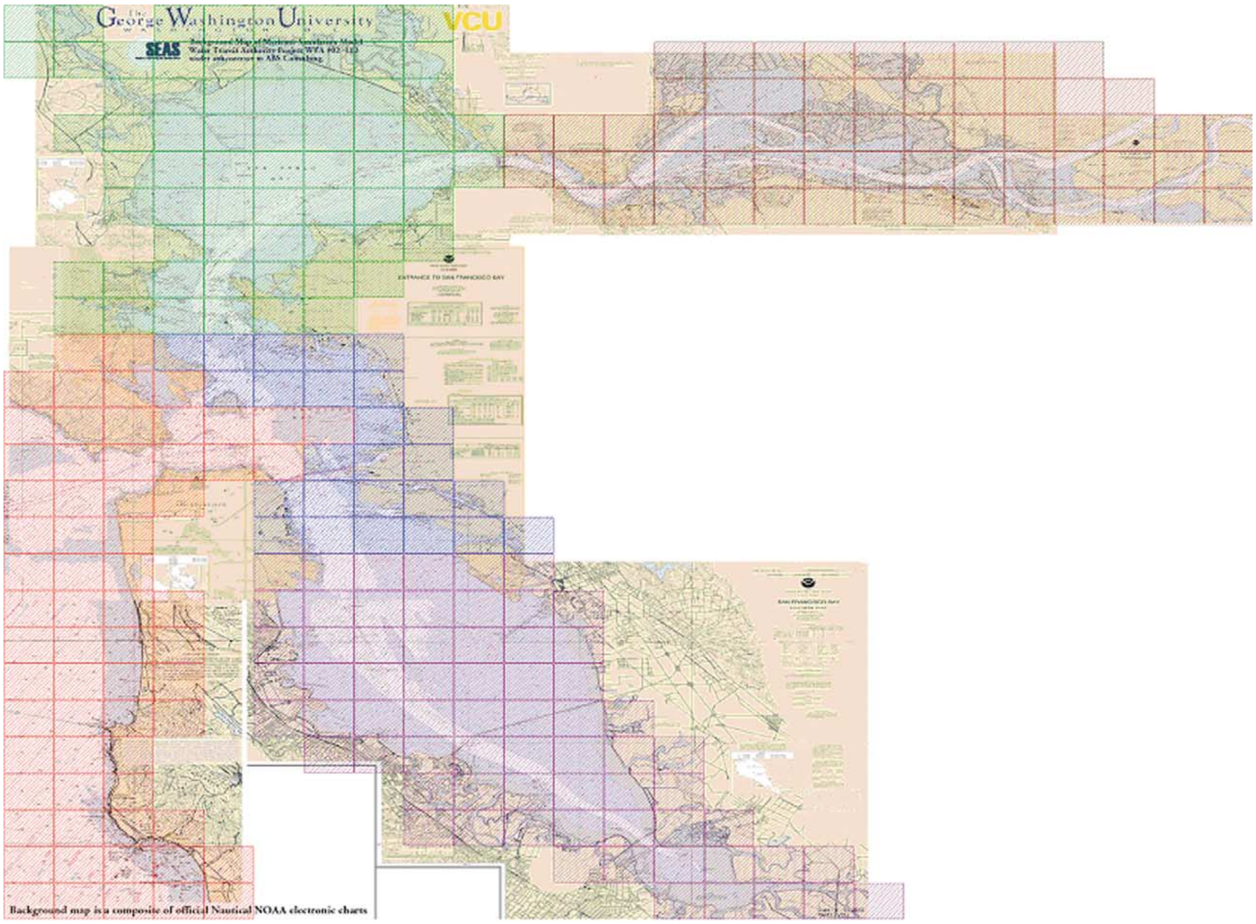


Fig. 4. Definition of visibility locations in the SF Bay maritime simulation model: Golden Gate (Red), San Pablo Bay (Green), Alameda (Blue), South Bay (Purple) and Grizzly Bay (Maroon).

5. The simulation: restricted visibility

Restricted visibility conditions have a significant impact on the pattern of traffic in the SF Bay in part due to the channel fog phenomenon at the Golden Gate Bridge during the third quarter of the year. To model these traffic patterns, visibility conditions were modeled in the simulation and, as mentioned previously, the movements of vessels were modified depending on these conditions. For the purposes of visibility modeling, the SF Bay area was divided into five regions; Golden Gate, San Pablo Bay, Alameda, South Bay and Grizzly Bay. The locations for visibility were defined using a square-grid breakdown of the study area. Fig. 4 identifies the different visibility locations used in the maritime simulation model. The location definitions

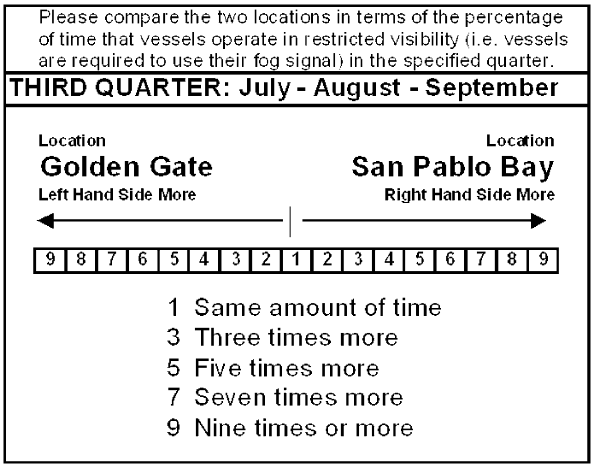


Fig. 5. Example pair wise comparison question for the location Golden Gate.

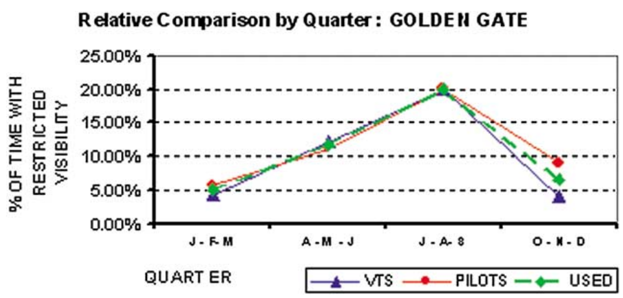


Fig. 6. Restricted visibility analysis results for the location Golden Gate for the first quarter of the year (J–F–M), second quarter (A–M–J), third quarter (J–A–S) and fourth quarter (O–N–D).

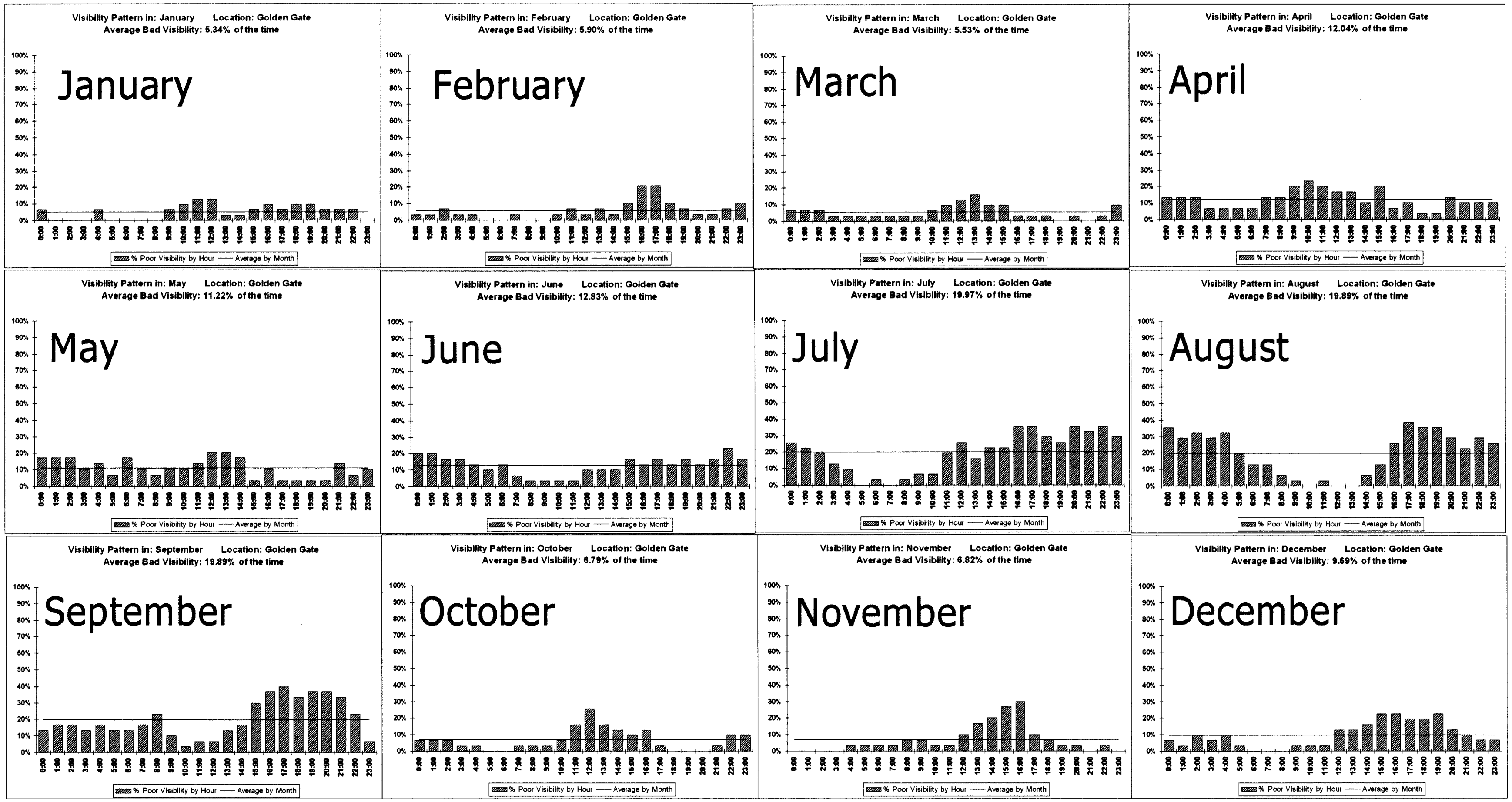


Fig. 7. Hourly percentages of restricted visibility for the location Golden Gate by month.

displayed in Fig. 4 were in part used to model the phenomenon of channel fog observed at the Golden Gate Location. Hourly wind speed and direction data is recorded via NOAA buoys for the period 1998–2001 at the five locations as well as dew point and water temperature data. Visibility data, however, is not gathered and thus a visibility model had to be developed.

The visibility model used in the simulation is based on a model described in Ref. [29]. The model stated that if the dew point is above the water temperature, then visibility will be restricted, otherwise the visibility will be good. In such a model, visibility is defined as good if it is greater than or equal to 0.6 miles and bad otherwise. Dew point and water temperature are recorded by the NOAA buoys, making such modeling of visibility possible. Rather than using this definition, we adhere to the rules of the road definition of restricted visibility (i.e. vessel operators are required to use their fog signals). A calibration constant was introduced into the visibility model to allow for this disparity, requiring the difference between the dew point and the water temperature to be above the calibration constant for such restricted visibility conditions to occur.

The calibration constant for the Golden Gate location for the third quarter of the year (July, August and September) was calculated from the US Coast Pilot's [30] data. The US Coast Pilot [30] states that restricted visibility conditions occur at Golden Gate approximately 20% of the time during the third quarter, the worst quarter for visibility in the Golden Gate location. However, no percentages are provided in the US Coast Pilot for the remaining quarters of the year; only anecdotal data is provided. Expert judgment was used to determine the calibration constants for restricted visibility conditions in the remaining three quarters at Golden Gate by comparing them to the third quarter. The experts involved were 7 operators from the SF VTS and 5 SF Bar Pilots with extensive experience throughout the SF Bay Area.

The process followed to elicit the remaining calibration constants utilizes the well-known Analytical Hierarchy Process [31,32]. Fig. 5 provides an example pair wise comparison question used in this process. Each expert is asked to assess whether restricted visibility is more likely in the quarter on the left-hand side or that on the right-hand side and by how much. The experts' assessments are used to calculate a relative multiplier for each quarter. By simple averaging of each expert's assessed values, for example, the resulting relative multiplier for the first quarter of the year was 0.258. This means that the experts indicated that the percentage of time that restricted visibility conditions occur in the first quarter of the year at Golden Gate should be 0.258 times the 20% of the third quarter (for which data was available) or 5.17%. Fig. 6 provides the results for the location Golden Gate. Note the (perhaps remarkable) agreement between the USCG VTS experts and SF Bar Pilots displayed in Fig. 6 for the remaining quarters of the year.

Table 1
Estimated percentages of time that restricted visibility occurs by quarter and by location

	First quarter, J–F–M	Second quarter, A–M–J	Third quarter, J–A–S	Fourth quarter, O–N–D
Golden Gate	5.17%	11.66%	20.00%	6.69%
San Pablo Bay	12.38%	6.17%	6.30%	9.62%
Alameda	7.49%	7.61%	10.61%	7.02%
South Bay	4.92%	5.00%	5.53%	4.74%
Grizzly Bay	14.40%	5.17%	5.34%	11.06%

The green line in Fig. 6 indicates the percentages that were used for calibration of the modified visibility model [29] for the Golden Gate location. Fig. 7 provides the monthly model results for this location for the year 2000. Note that, in the third quarter (July, August and September) the model reflects early morning fog that burns off during the late morning hours and early afternoon hours and reestablishes itself during the late afternoon. The latter daily pattern is typical for the channel fog phenomenon for this quarter at the Golden Gate location [30].

No visibility data, in terms of percentage of time that restricted visibility occurs, was available for the remaining locations San Pablo Bay, Alameda, South Bay and Grizzly. Hence, we had to rely once again on expert judgment to determine calibration constants for restricted visibility conditions. We followed the same process as above, comparing these four locations by quarter to the previously established percentage of time that restricted visibility occurs in Golden Gate (Fig. 6). For example, a multiplicative factor of 2.397 was assessed for the location San Pablo Bay during the first quarter of the year when compared to the Golden Gate location. Utilizing the previously established 5.17% for restricted visibility in Golden Gate during this quarter, the percentage of time that restricted visibility occurs in San Pablo Bay was set at 2.397 times 5.17% or 12.38%. Table 1 provides the estimated percentages of time that restricted visibility occurs by

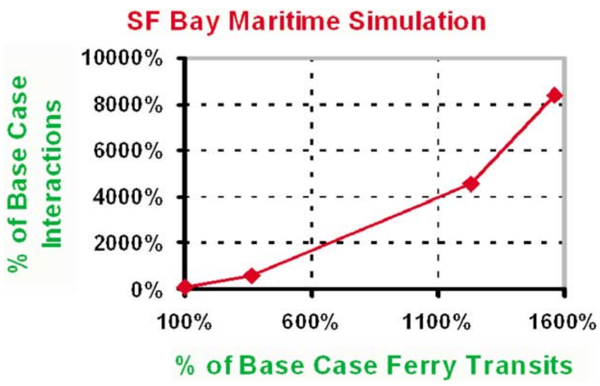


Fig. 8. Exponential growth in interactions due to ferry service expansion.

Table 2
Percentage comparisons to the Base Case under various criteria

	Base Case ferry transits (%)	Base Case grid cells covered (%)	# Base Case total interactions (%)
Base Case	100	100	100
Alternative 3	365	116	624
Alternative 2	1228	233	4620
Alternative 1	1559	240	8359
Alternative 3-BVI	–	91	110

quarter of the year and by location. The information in Table 1 was used to calculate the calibration constants for the visibility model for the remaining locations, San Pablo Bay, Alameda, South Bay and Grizzly.

6. Results

Fig. 1 shows a screen shot of the simulation program created to perform the vessel interaction analysis. For a more detailed look, movies of the simulation for each of the cases can be viewed at http://www.people.vcu.edu/~jrmerric/SFBayMovies/. Recall that the simulation was intended to answer certain specific questions. For the defined scenarios, what is the increase in the number of interactions involving ferries? What is the increase in the area in which such interactions occur? Are there any high-density areas that could be a cause of concern, either in the current ferry system or in any of the proposed scenarios? As interactions in restricted visibility are of particular concern, what is the affect of the proposed scenarios on frequency and density of such interactions?

We will start our discussion of the results of the simulation analysis with some basic comparisons to current ferry operations. The current ferry operations, or the Base Case, are used as a reference point to compare the proposed alternatives and to give an understanding of the traffic patterns currently seen by ferries in the study area. Fig. 8 summarizes the analysis findings. Observe from Fig. 8 that the number of ferry to vessels interactions grows exponentially with the number of ferry transits, not linearly. This result was somewhat of a revelation for the WTA. Table 2 gives the detail of

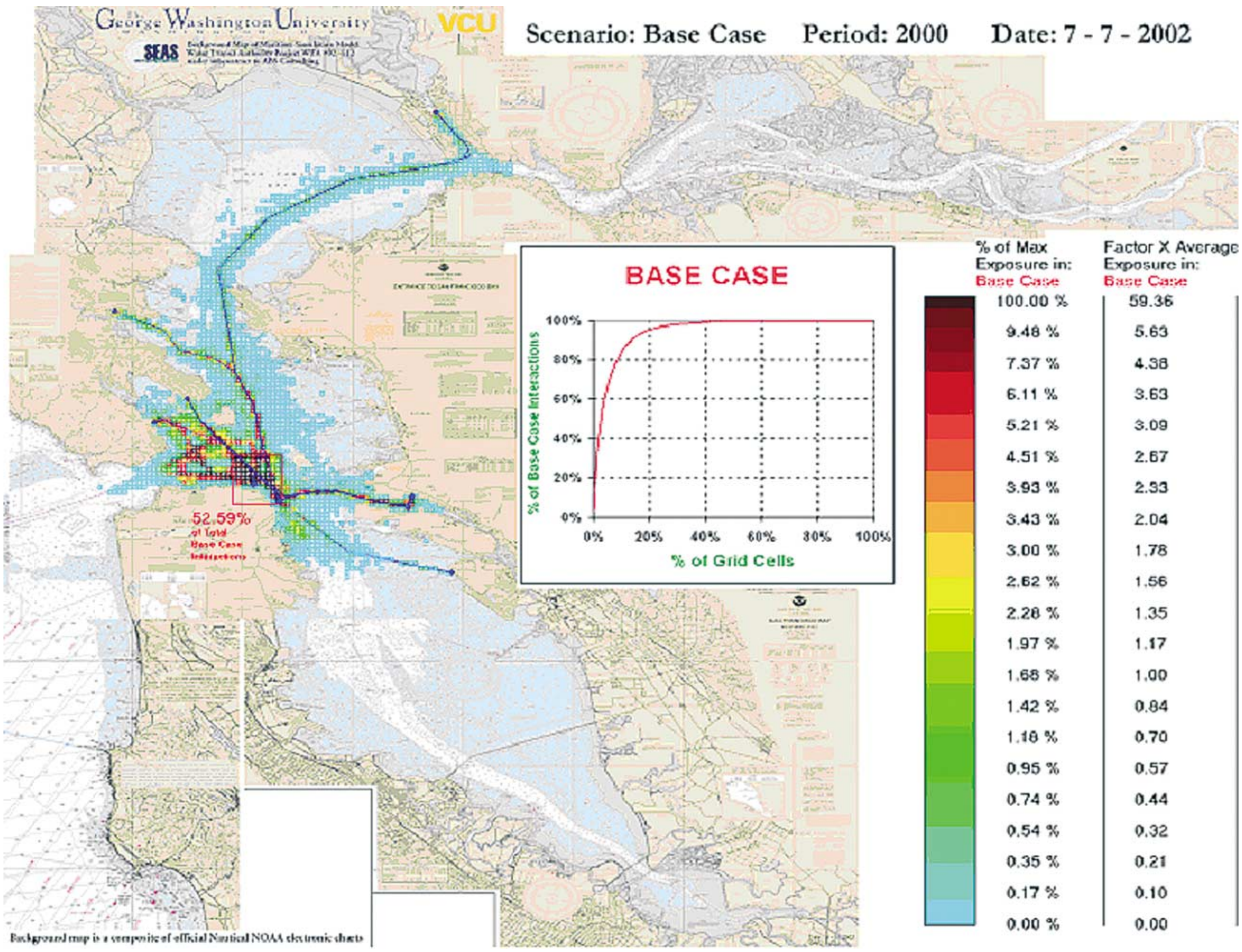


Fig. 9. The full base case simulation results.

the comparison of the three alternative cases to the Base Case.

Alternative 3 (the least aggressive expansion) has 3.65 times as many transits as the Base Case, but covers only a little larger area, with 16% more grid cells having at least one interaction in them in the simulation. In all over 6 times as many interactions occur in Alternative 3 than occurred in the Base Case, while the coverage area of these interactions only increases by a factor of 1.16. Thus Alternative 3 makes the current operating area more congested with more interactions. In addition, the fourth row in Table 2 displays results for Alternative 3 counting only those interactions that occur in restricted visibility. Note that, 1.10 times as many interactions occur in Alternative 3 in restricted visibility than the whole Base Case (regardless of visibility). Moreover, these interactions cover only 91% of the coverage area in the Base Case and are thus more concentrated. We will return to this important observation.

Alternative 2 has 12.28 times as many transits as the Base Case, but covers a much larger area, with 2.33 times as many grid cells having at least one interaction. In all over 46 times as many interactions occur in Alternative 2 than occurred in the Base Case. Thus Alternative 2 increases the operating area from the Base Case and leaves the system much more congested with many more interactions. Finally, Alternative 1 (the most aggressive expansion) has 15.59 times as many transits as the Base Case, but covers only a little larger area than Alternative 2, with 2.4 times as many grid cells having at least one interaction than in the Base Case. In all over 83 times as many interactions occur in Alternative 1 than occurred in the Base Case. Thus Alternative 1 increases the operating area by about the same factor as Alternative 2, but significantly increases congestion with many more interactions compared to Alternative 2.

Fig. 9 shows the geographic interaction profile for the Base Case. The Base Case ferry routes are shown in color. Fig. 9 is quite complex, as it attempts to convey all the Base Case results in one figure. We will examine the pieces of Fig. 9 one by one. The analysis is broken down across a grid of approximately 1/4 mile by 1/4 mile cells. The cells are

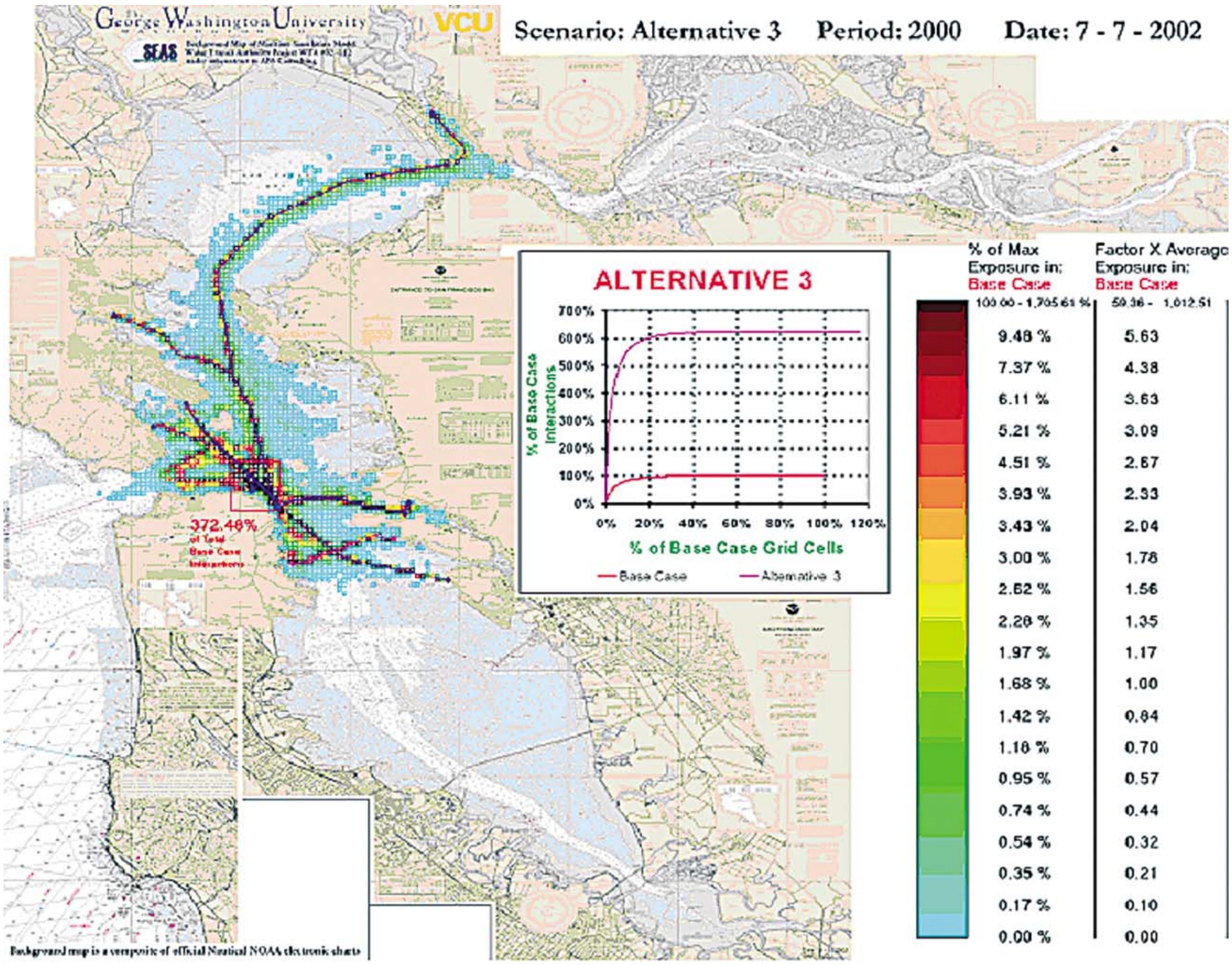


Fig. 10. The full Alternative 3 simulation results.

color coded in Fig. 9 to represent the number of interactions that occur in that cell over the 1-year simulation time. Both the cell containing the ferry and the cell containing the interacting vessel are recorded; hence the colored cells away from the ferry routes.

To the right of Fig. 9, the legend gives an interpretation for the color-coding of the cells. The scale goes from blue, with the fewest interactions, to black with the most interactions. The solid black cell has the most interactions of any cells in the base case simulation. This Base Case maximum is used as a reference point for the legend. The percentages shown in the legend are calculated as a percentage of this maximum number of interactions. For example, an orange cell has an interaction count that is only 3% of the maximum number of interactions observed in a grid cell in the Base Case. Another reference scale is also provided. The average number of interaction per cell in the Base Case has 1.68% of the maximum number of interactions in a cell observed in the Base Case. Returning to our example, an orange cell has 1.78 times the number of interactions seen in the average cell in the Base Case. A solid black cell, with the most interactions, has over 60 times as many interactions as the average in the Base Case, indicating that some cells are highly congested when compared to the average cell. One can also see that the legend is not numerically linear. Since some of the cells are much more congested than others, we have had to develop a color gradient following a power curve to highlight their differences.

What can we learn about the current ferry operations, or Base Case, from Fig. 9? The majority of the dark colored grid cells are in the Central Bay area, particularly close to the Ferry Building. In fact, if we take the red square around the Ferry Building, almost 53% of all the interactions in the Base Case occur in this area. This is the area with most ferries, a great deal of other VTS Traffic and organized recreational events operating, combined with the worst visibility for a large part of the year (especially in the third quarter of each year).

Figs. 10 and 11 examine Alternative 3 (the least aggressive expansion) and Alternative 1 (the most aggressive expansion) and compare their results to the Base Case in the same figures. A similar geographic interaction profile was generated for Alternative 2 (the future ferry expansion

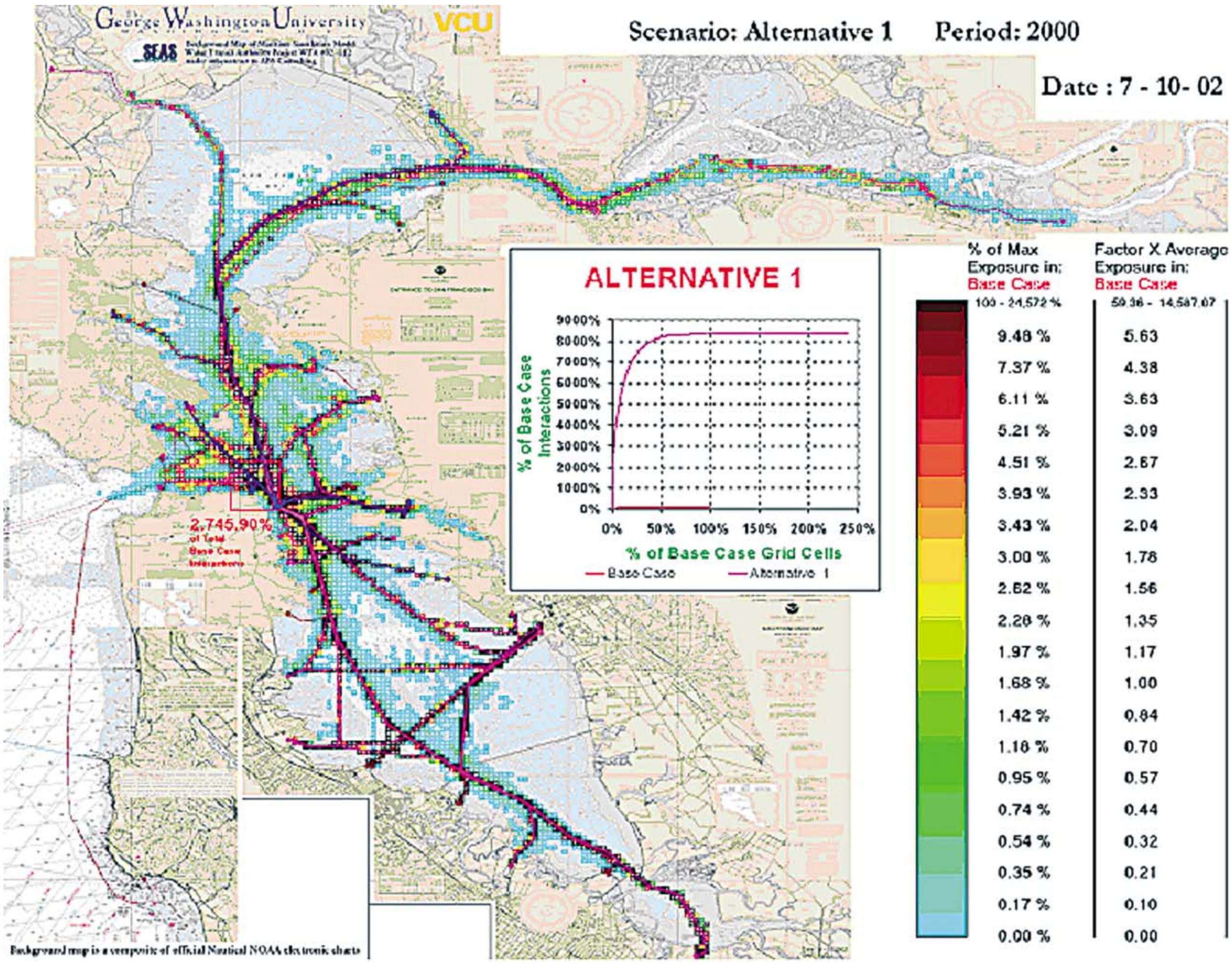


Fig. 11. The full Alternative 1 simulation results.

between Alternative 1 and Alternative 3). Note that the legend has not changed to allow the comparison to the Base Case. Notice that the same red square around the Ferry Building in Alternative 3 (Fig. 10) now contains 3.7 times as many interactions as the whole Base Case and that much of the area within the red square is now colored solid black, indicating that there are more interactions in those grid cells than the maximum for any grid cell in the Base Case. Similar conclusions can be drawn from Fig. 11 showing the geographic interaction profile for Alternative 1 (the most aggressive expansion of future ferry service). Notice that, the same red square around the Ferry Building now contains approximately 27 times as many interactions as the whole Base Case and again much of the area is colored solid black, indicating that there are more interactions in those grid cells than the maximum for any grid cell in the Base Case.

Of particular concern are interactions that occur in restricted visibility. Recall from Table 2 that 1.10 times as many interactions occur in Alternative 3 in restricted visibility than the whole Base Case (regardless of visibility). Moreover, these interactions cover only 91% of the coverage area in the Base Case. Fig. 12 displays the results for Alternative 3 counting only those interactions that occur in restricted visibility. Concentrating on the red square in Fig. 12, it follows that 57.92% of the interactions in the whole Base Case (regardless of visibility) are now occurring in the red square in restricted visibility conditions in Alternative 3. In the Base Case, 6.57% of the total interactions occurred in restricted visibility in the red square. Hence, although Alternative 3 (the least aggressive ferry expansion) resulted in an increase from the Base Case of 3.65 times as many interactions overall, an approximate increase of 8.82 (= 57.92/6.57%) times as many interactions are observed in the red square in Fig. 12 in restricted visibility. These restricted visibility interactions involve both regular and high-speed ferries in an area that is already the most congested in the Base Case. Findings of this nature should be of concern to those planning for future ferry expansions.

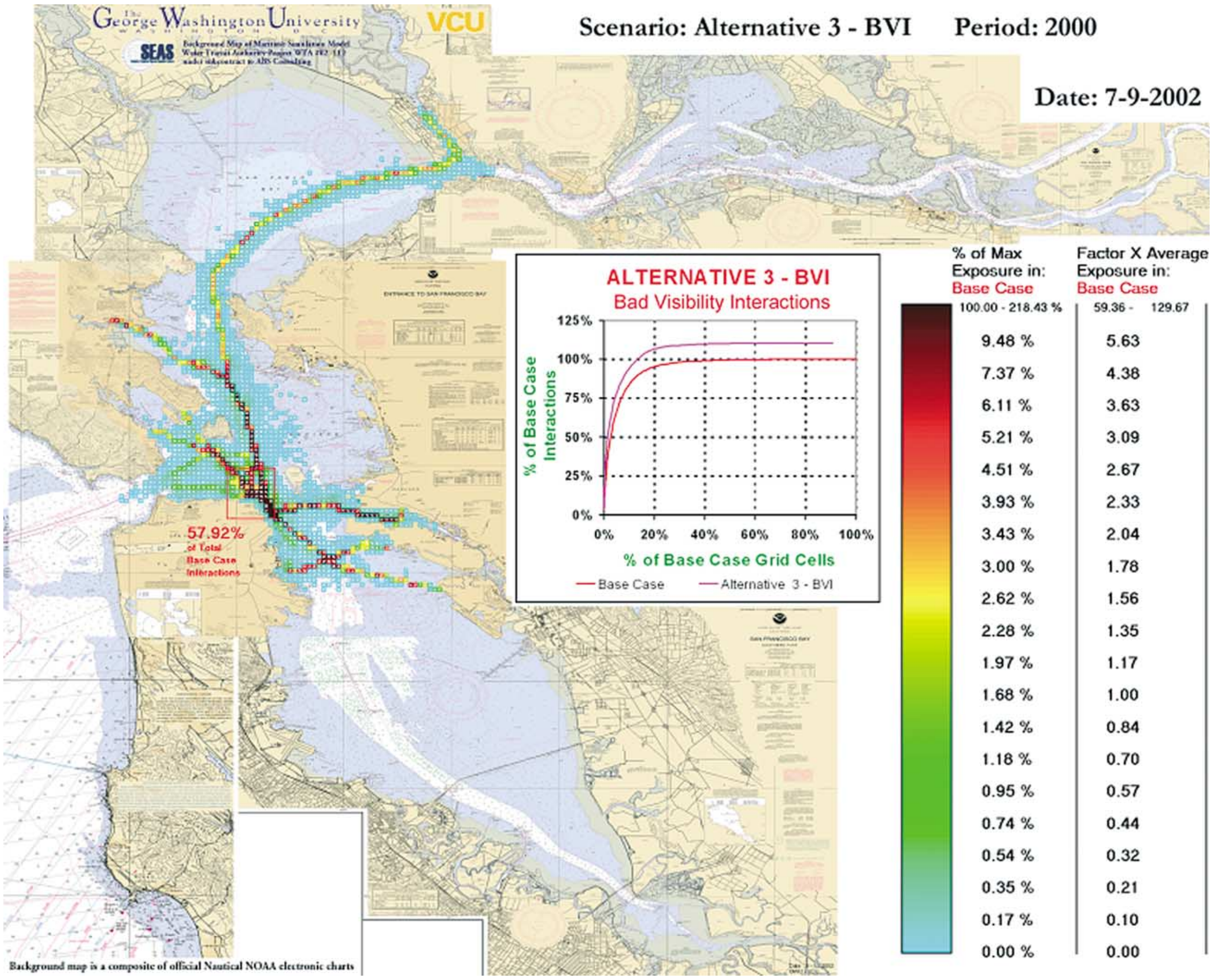


Fig. 12. Alternative 3 results counting only restricted visibility interactions.

7. Conclusions and recommendations

The analysis discussed herein is only one part of the overall assessment of the proposed ferry service expansion by the WTA. Digital movies of the simulation were requested by the WTA allowing the decision-makers to visualize the reality of their proposed ferry service expansions. In addition, other projects are underway or have been completed examining environmental issues, ferry terminal expansions, ridership, intermodal transportation issues, and new technologies (http://www.watertransit.org). Each of these studies will be summarized in the *Implementation and Operations Plan* to be submitted to the California Legislature on December 12th 2002, with review continuing through the summer of 2003.

The vessel interaction analysis presented in this paper provides a foundation for examining the risk inherent in such a major expansion of service and is a first step in a full risk assessment that would satisfy the requirements of the US Coast Guard Captain of the Port. The vessel interaction analysis results can be combined in follow on steps with a conditional accident probability model and an accident damage model for an overall estimate of MTS accident risk [5]. These results, however, do give an initial indication of where high accident risk spikes may occur by illustrating the occurrences of added congestion and their location. In addition, the results seem to indicate that the safety levels currently enjoyed by the SF Bay ferry service cannot be maintained under the planned expansion scenarios without equally aggressive investment in risk intervention. With the broader picture of risk in mind, the project team made the following recommendations to the WTA at the conclusion of the project:

1. Use the results of the simulation analysis in a PRA similar to that of the Washington State Ferry Risk Assessment, where output analyses is presented in terms of expected number of accidents per year.
2. Consider the current SF Bay Ferry Operations and future planned ferry operations as a MTS rather than a collection of individual ferry routes by
   a. Designing a ferry traffic routes system that allows for increased ferry traffic while limiting the increase in expected number of accidents per year.
   b. Designing ferry schedules utilizing this ferry traffic route system that allow for increased ferry traffic while limiting the increase in expected number of accidents per year. A consideration in the development of these future schedules should be the time between arrivals and departures at ferry terminals to allow for sufficient time of loading and unloading passengers.
3. Develop other risk intervention measures that can reduce the number of interactions and the probability of accidents given an interaction.
4. Investigate the effect of proposed risk intervention measures on the accident probability using the full PRA model.
5. Perform an uncertainty analysis of accident risk and risk intervention evaluation to provide estimates of annual accident risk and risk intervention effectiveness in terms of probability intervals rather than point estimates.

> The truth is that we are uncertain. The language of uncertainty is probability. Therefore, speaking the truth means to develop analyses results in terms of probability curves rather than in terms of point estimates [33].

Acknowledgements

Special thanks to the San Francisco Bay Area Water Transit Authority for the opportunity to conduct this project. We would also like to extent our gratitude to the following project members for their support: Walt E. Hanson from ABS Consulting provided project management and data collection support; Stacey W. Shonk from the California Maritime Academy provided detailed knowledge about the Maritime Transportation System in the San Francisco Bay, facilitated meetings with local area users and provided data collection support; Philip B. Harms, Jr. from the California Maritime Academy provided help in constructing a large scale nautical map of the San Francisco Bay Area; Lt. Black and Alan M. San from US Coast Guard Vessel Traffic Service San Francisco provided traffic data, recreational vessel information and experts willing to fill out in the restricted visibility questionnaires; URS Corporation provided maps detailing existing and future planned ferry routes; The San Francisco Bar Pilots donated their time and knowledge on vessel movements and visibility; the ferry operators Blue and Gold and Golden Gate Bridge allowed us to ride ferries and providing access to ferry captains for discussions while underway.

Finally, we would like to thank the Editor in Chief and the referee for their helpful comments, which substantially improved the first version of this paper. The research described herein was partially supported by San Francisco Bay Water Transit Authority Project WTA #02-112 and partially supported by NSF grants SES 0213627 and SES 0213700.

References

[1] Harrald J, Mazzuchi T, Merrick J, van Dorp JR, Spahn J. Using system simulation to model the impact of human error in a maritime system. Safety Sci 1998;30(1/2):235–47.

[2] Merrick J, van Dorp JR, Harrald J, Mazzuchi T, Spahn J, Grabowski M. A systems approach to managing oil transportation risk in Prince William Sound. Syst Engng 2000;3(3):128–42.

[3] Merrick J, van Dorp JR, Harrald J, Mazzuchi T, Spahn J, Grabowski M. The Prince William Sound risk assessment. Interfaces 2002;32(6):25–40.

[4] Grabowski M, Merrick J, Harrald J, Mazzuchi T, van Dorp JR. Risk modeling in distributed, large-scale systems. IEEE Syst, Man Cybernet Part A: Syst Humans 2001;30(6):651–60.

[5] van Dorp JR, Merrick J, Harrald J, Mazzuchi T, Grabowski M. A risk management procedure for the Washington State ferries. Risk Anal 2001;21:127–42.

[6] National Research Council. Crew size and maritime safety. Washington, DC: National Academy Press; 1986.

[7] National Research Council. Tanker spills: prevention by design. Washington, DC: National Academy Press; 1991.

[8] National Research Council. Minding the helm: marine navigation and piloting. Washington, DC: National Academy Press; 1994.

[9] National Research Council. Risk management in the marine transportation system. Washington, DC: National Academy Press; 2000.

[10] Pravda MF, Lightner RG. Conceptual study of a supercritical reactor plant for merchant ships. Mar Technol 1966;4:230–8.

[11] Stiehl GL. Prospects for shipping liquefied natural gas. Mar Technol 1977;14(4):351–78.

[12] Paté-Cornell ME. Organizational aspects of engineering system safety: the case of offshore platforms. Science 1990;250(4985):1210–7.

[13] US Coast Guard. Vessel traffic systems: analysis of port needs. Report No. AD-770 710. Washington, DC: US Coast Guard; 1973.

[14] Maio D, Ricci R, Rossetti M, Schwenk J, Liu T. Port needs study. Report No. DOT-CG-N-01-91-1.2. Prepared by John A. Volpe, National Transportation Systems Center. Washington, DC: US Coast Guard; 1991.

[15] Bedford TM, Cooke RM. Probabilistic risk analysis: foundations and method. Cambridge, UK: Cambridge University Press; 2001.

[16] Hara K, Nakamura S. A comprehensive assessment system for the maritime traffic environment. Safety Sci 1995;19(2/3):203–15.

[17] Roeleven D, Kok M, Stipdonk HL, de Vries WA. Inland waterway transport: modeling the probabilities of accidents. Safety Sci 1995;19(2/3):191–202.

[18] Kite-Powell HL, Jin D, Patrikalis NM, Jebsen J, Papakonstantinou V. Formulation of a model for ship transit risk. MIT Sea Grant Technical Report, Cambridge, MA, 96–19; 1996.

[19] Slob W. Determination of risks on inland waterways. J Hazard Mater 1998;61(1–3):363–70.

[20] Fowler TG, Sorgard E. Modeling ship transportation risk. Risk Anal 2000;20(2):225–44.

[21] Trbojevic VM, Carr BJ. Risk based methodology for safety improvements in ports. J Hazard Mater 2000;71(1–3):467–80.

[22] Wang J. A subjective modeling tool applied to formal ship safety assessment. Ocean Engng 2000;27(10):1019–35.

[23] Guedes Soares C, Teixeira AP. Risk assessment in maritime transportation. Reliab Engng Syst Safety 2001;74(3):299–309.

[24] Andrews S, Murphy FH, Wang XP, Welch S. Modeling crude oil lightering in Delaware Bay. Interfaces 1996;26(6):68–78.

[25] Golkar J, Shekhar A, Buddhavarapu S. Panama canal simulation model. Proceedings of the 1998 Winter Simulation Conference; 1998. p. 1229–37.

[26] Ryan NK. The future of maritime facility designs and operations. Proceedings of the 1998 Winter Simulation Conference; 1998. p. 1223–7.

[27] Merrick J, van Dorp JR, Mazzuchi T, Harrald J. Modeling risk in the dynamic environment of maritime transportation. Proceedings of the 2001 Winter Simulation Conference; 2001. p. 1090–8.

[28] Sargent RG. Validation and verification of simulation models. Proceedings of the 1999 Winter Simulation Conference; 1999. p. 39–48.

[29] Sanderson R. Meteorology at sea, Stanford Maritime Limited; 1982.

[30] Evans DL, Grudes SB, Davidson MA. United States Coast Pilot Volume 7, Pacific Coast, California, Oregon, Washington and Hawai. Washington, DC: National Ocean Service, US Department of Commerce; 2001.

[31] Saaty T. The analytic hierarchy process. New York: McGraw-Hill; 1980.

[32] Vargas LG. An overview of the analytic hierarchy process and its applications. Eur J Oper Res 1990;48:2–8.

[33] Kaplan S. The words of risk analysis. Risk Anal 1997;17(4):407–17.

SUB-APPENDIX:

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.

Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 169 (2006) 157–177

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/ejor

Stochastics and Statistics

A Bayesian paired comparison approach for relative accident probability assessment with covariate information

P. Szwed [a,b], J. Rene van Dorp [b,*], J.R.W. Merrick [c], T.A. Mazzuchi [b], A. Singh [b]

[a] *Department of Management, US Coast Guard Academy, 31 Mohegan Avenue, New London, CT 06320-8103, USA*
[b] *Engineering Management and Systems Engineering Department, The George Washington University, 1776 G Street NW, Suite 110, Washington, DC 20052, USA*
[c] *Department of Statistical Sciences and Operations Research, Virginia Commonwealth University, P.O. Box 843083, VCU, Richmond, VA 23284-2014, USA*

Received 23 September 2003; accepted 22 April 2004
Available online 3 August 2004

Abstract

One of the challenges managers face when trying to understand complex, technological systems (in their efforts to mitigate system risks) is the quantification of accident probability, particularly in the case of rare events. Once this risk information has been quantified, managers and decision makers can use it to develop appropriate policies, design projects, and/or allocate resources that will mitigate risk. However, rare event risk information inherently suffers from a sparseness of accident data. Therefore, expert judgment is often elicited to develop frequency data for these high-consequence rare events. When applied appropriately, expert judgment can serve as an important (and, at times, the only) source of risk information. This paper presents a Bayesian methodology for assessing relative accident probabilities and their uncertainty using paired comparison to elicit expert judgments. The approach is illustrated using expert judgment data elicited for a risk study of the largest passenger ferry system in the US.

Keywords: Applied probability; Expert judgment; Risk analysis

1. Introduction

The concepts of risk analysis and management is becoming more and more relevant in our complex technological environment. Numerous papers and books have been written in the last 20 years on this topic

* Corresponding author. Tel.: +1 202 994 6638; fax: +1 202 994 0245.
E-mail addresses: pszwed@cga.uscg.mil (P. Szwed), dorpjr@gwu.edu (J. Rene van Dorp), jrmerric@vcu.edu (J.R.W. Merrick), mazzu@gwu.edu (T.A. Mazzuchi), amitas@gwu.edu (A. Singh).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2004.04.047

(see, e.g., Shrader-Frechette, 1985; Paté-Cornell, 1996; Kumamoto and Henley, 1996; Kaplan, 1997; Koller, 2000; Bedford and Cooke, 2001). Risk analysis, also known as risk assessment, is widely recognized as a systematic, science-based process for quantitatively describing risk (see, e.g., Vose, 1996). Risk, itself, is commonly defined as a quantitative measure combining the likelihood of the occurrence of an undesirable event (accident) and its consequences. Assessment of risk may be separated into the quantitative assessments of accident probabilities and consequences. Kaplan (1997) among others discusses the definition of risk in more detail. Regardless of exactly how these quantitative measures are combined into a single risk measure, separate information about accident probability and consequences are critically important to managers who are charged with risk mitigation because different risk interventions follow from accident probability reduction and consequence reduction.

The quantification of risk models for policy and decision-making often requires the elicitation of expert judgments (see, e.g., Moslesh et al., 1988; Bonano et al., 1989; Morgan and Henrion, 1991; Cooke, 1991). In fact, as long as the fundamental mechanisms that drive a system remain poorly known, the encoding of expert knowledge will be required (see Paté-Cornell, 1996). Nevertheless, as noted by Anderson et al. (1999), expert judgment must be used with care. It is not evidence per se, but an individual's or group's inference based on available evidence. Kahneman et al. (1982) (a Nobel Prize winner in 2002) discuss the numerous biases and heuristics that are introduced when humans process information and attempt to provide judgments.

Winkler (1996) points out that due to the general belief that "several heads are better than one", information is usually elicited from several experts. Numerous techniques exist for the aggregation of multiple experts' responses (see, e.g., Morris, 1974; Winkler, 1981; Genest and Zidek, 1986; Clemen, 1989; Mendel and Sheridan, 1989; Cooke, 1991; DeWispelare et al., 1995). In recent reviews of the techniques, Clemen and Winkler (1990, 1999) note that often the simple aggregation techniques may work just as well as the more complex methods. The Bayesian paradigm, however, seems to supply at the present the most natural and unambiguous approach towards the aggregation problem while addressing uncertainty in the expert judgment at the same time.

While a number of different elicitation methods are available (see, e.g., Cooke (1991) for an excellent overview), the paired comparisons elicitation method seems to be quite popular. The elicitation method to be discussed in this paper belongs to this class. In the next section we reflect on the origins of the paired comparisons elicitation method.

1.1. Paired comparisons elicitation approaches

Origins of this class can be traced back to Thustone'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) via the Bradley and Terry (1952) paired comparisons method. An examination of the latter method is provided by Cooke (1991), among other numerous sources.

Another popular paired comparison elicitation technique is called the Analytical Hierarchy Process (AHP) developed by Saaty (1977, 1980). The AHP Process is primarily used for the construction of value functions $V(\underline{X})$ involving multiple contributing factors $\underline{X}=(X_1, X_2, \ldots, X_p)$ (see, e.g. Foreman and Selly, 2002). The construction of a value function in this manner extends the construction of a utility function based on paired comparisons. The theoretical foundation for developing the latter has been provided by the Nobel Laureate G. Debreu (see, e.g., Debreu, 1986). The popularity of the elicitation methods above can perhaps be contributed to the observation that experts are more comfortable making paired comparisons rather than directly assessing a quantity of interest. It should however be mentioned that paired comparisons may lead occasionally to the so-called Simpson paradox-lack of transitivity (see Simpson, 1951).

To the best of our knowledge, Pulkkinen (1993, 1994a,b) was first to introduce a Bayesian paired comparison aggregation method for the elements of a multivariate random vector $\underline{\beta}=(\beta_1,\beta_2,\ldots,\beta_p)$ by multiple experts. Experts are asked to compare the pair of random variables β_i to β_j, $i\neq j$, $i=1,\ldots,p$, and respond in terms of an indicator function $1_{[\beta_i \geqslant \beta_j]}$ (i.e. 1 when the expert judges $\beta_i \geqslant \beta_j$ and 0 otherwise). The paired judgments in Pulkkinen's analysis are assumed to be consistent. Pulkkinen's (1993, 1994a,b) exposition is mainly theoretical and limited to a discussion of mathematical properties of the aggregation method, but mentions that applications of his method in the reliability engineering and system safety domain are self-evident.

We shall report herein on what appears to be a novel paired comparison elicitation method for accident probabilities. We take as an application of this approach an actual case study "The Washington State Ferry (henceforth WSF) Risk Assessment" where paired comparisons were elicited from experts. The next section discusses an overview of the WSF Risk Assessment (see also Van Dorp et al. (2001) for a more detailed description).

1.2. Overview of the WSF risk assessment

The WSF system is the largest ferry system in the United States. In 1997, total ridership for the ferries serving the central Puget Sound region was nearly 23 million, a 4% increase over 1996 ridership, and more

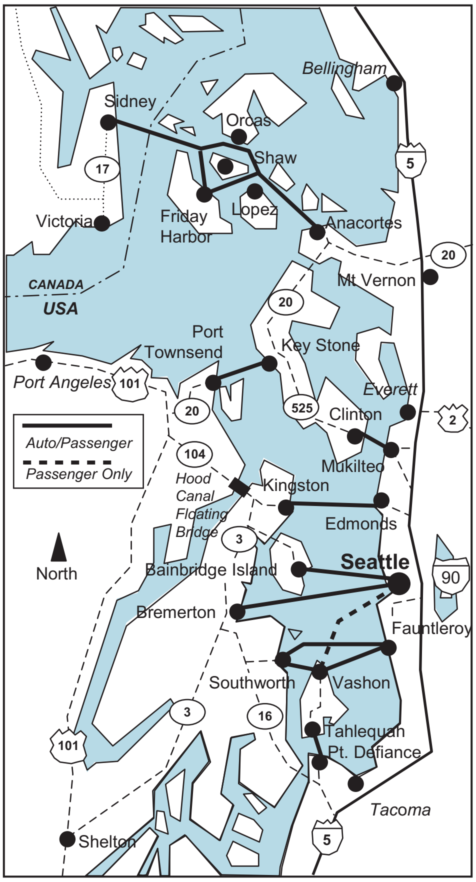


Fig. 1. Washington state ferry system map.

passengers than Amtrak, the US passenger rail carrier, handles in a year. Fig. 1 shows the current ferry routes for the central Puget Sound region. This map illustrates the ferry system's role in linking together the Washington State highway system in the Puget Sound region.

In part due to the introduction of high speed ferries, the State of Washington established an independent Blue Ribbon Panel to assess the adequacy of requirements for passenger and crew safety aboard the Washington State Ferries. On July 9, 1998, the Blue Ribbon Panel engaged a consultant team from The George Washington University and Rensselaer Polytechnic Institute/Le Moyne College to assess the adequacy of passenger and crew safety in the WSF system, to evaluate the level of risk present in the WSF system, and to develop recommendations for prioritized risk reduction measures which, once implemented, can improve the level of safety in the WSF system. The probability of ferry collisions in the WSF system was assessed using a dynamic simulation methodology that extends the scope of available data with expert judgment.

Experts were selected amongst WSF captains and WSF first mates who had extensive experience with all 13 different ferry routes over an extended period of time (more than 5 years). During the WSF risk assessment in 1998 expert responses to paired comparisons were aggregated by taking geometric means of their responses and using them in a classical log linear regression analysis approach to assess relative collision probabilities. The classical analysis conducted during the WSF risk assessment only resulted in point estimates of relative collision probabilities. We shall improve on the previous classical analysis by providing distributional results on these relative collision probabilities by developing a Bayesian inference engine for the paired comparison questionnaires administered during the WSF Risk Assessment. This is in compliance with the almost classical "speaking the truth in risk assessment" argument (see, e.g., Kaplan, 1997, p. 412) originating from the early 1980's when the International Society for Risk Analysis was founded: "*Since the truth is, we always have uncertainty, we say that speaking in probability curves is telling the truth*". The paired comparison elicitation method developed herein is not limited to the maritime domain and may generally be applicable to relative accident probability estimation when limited or no data is available. The research conducted by us is part of a larger project funded by the National Science Foundation to address uncertainty in large scale maritime risk assessments in a coherent manner.

1.3. Bayesian paired comparison approach for relative accident probabilities

Similar to the AHP process, we are interested in the functional relationship between contributing factors $\underline{X}=(X_1, X_2, \ldots, X_p)$ and an accident probability (rather than a value function). Our accident probability behaves much like a value function. That is, not only is the order amongst different sets of contributing factors (or covariates) $\underline{X}$ important, but also the differences in their values. Whereas Pulkkinen's focus (1993, 1994a,b) is on the multivariate distribution of a random vector $\underline{\beta}$, our focus is more applied and based on the distribution of an accident probability $\Pr(Accident|Incident, \underline{X})$ defined by

$$\Pr(Accident|Incident, \underline{X}) = P_0 Exp(\underline{\beta}^{\mathrm{T}}\underline{X}), \tag{1}$$

where $\underline{X}=(X_1, X_2, \ldots, X_p)$ describes a system state during which an incident (e.g. a mechanical failure) occurred. The accident probability model (1) has been proposed in previous maritime risk assessments (see, e.g., Roeleven et al., 1995; Merrick et al., 2000; Van Dorp et al., 2001), resembles the well-known proportional hazards model originally proposed by Cox (1972) and builds on the assumption that accident risk behaves exponentially rather than linearly with changes in covariate values. Our goal is to establish the uncertainty distribution of the accident probability $\Pr(Accident|Incident, \underline{X})$ in entirety rather than a point estimate. Similarly to Pulkkinen (1993, 1994a,b), our aggregation method of the expert judgment paired comparisons will follow the Bayesian paradigm. A questionnaire of paired comparisons is used to elicit the relative contribution of the elements of $\underline{X}$ to the accident probability and update its uncertainty, initially captured by (1) and a prior multivariate distribution of the random vector $\underline{\beta}$.

The Bayesian analysis conducted herein exploits the structure of (1) to result in a conjugate analysis (i.e. the prior and posterior distributions belong to the same family of distributions) involving a multivariate normal prior for the parameter vector $\underline{\beta}$ and a univariate gamma prior on an expert's precision (or, perhaps more appropriately, imprecision). In Section 2, we provide some background surrounding the use of the accident probability model (1) in large maritime risk assessments drawing primarily from the Washington State Ferry (WSF) Risk assessment (see Van Dorp et al., 2001). The likelihood of the expert responses to the paired comparison questionnaire is presented in Section 3. The prior distribution on the parameter vector $\underline{\beta}=(\beta_1,\beta_2,\ldots,\beta_p)$ and the expert judgment's precision is discussed in Section 4. The conjugate analysis deriving the posterior distribution of $\underline{\beta}=(\beta_1,\beta_2,\ldots,\beta_p)$ and the expert judgment's precision is presented in Section 5. In addition, parameter uncertainty in $\underline{\beta}=(\beta_1,\beta_2,\ldots,\beta_p)$ and uncertainty in the expert judgment is propagated through the accident probability model $\Pr(Accident|Incident,\underline{X})$ to arrive at closed form expressions for prior and posterior distributions of relative accident probabilities. A calculation example is presented using expert judgment data elicited during the WSF risk assessment (see, Van Dorp et al., 2001) in Section 6. Some concluding remarks are provided in Section 7.

2. Accident probability model

An accident is not a single event, but can be considered to be the culmination of a series of cascading events (see Garrick, 1984) starting with a triggering incident. In the maritime accident probability model in Merrick et al. (2000) and Van Dorp et al. (2001), triggering incidents have been further categorized as mechanical failures and human errors. Accidents and triggering incidents occur within the context of a system defined by ever changing combinations of contributing factors. Contributing factors may be further classified in organizational factors (OF) and situational factors (SF). In the WSF risk assessment an example of an organizational factor is a specific ferry route and ferry class combination (since operating teams are assigned by ferry class and route), whereas examples of situational factors are the changing weather conditions and traffic patterns while a ferry is underway. Fig. 2 provides an example of an accident probability model, the time sequence of the accident event chain and the influence of contributing factors on this chain. The accident probability model in Fig. 2 is based on the notion of conditional probability. The levels of conditional probability reflected in Fig. 2 are

- $\Pr(OF,SF)$: the probability that a particular set of organizational and situational factors occur in the system,
- $\Pr(Incident|OF)$: the probability that an incident occurs given the organizational factors and
- $\Pr(Accident|Incident,OF,SF)$: the probability that an accident occurs given that a triggering incident has occurred under the organizational and situational factors.

To perform an assessment of the annual accident risk and its uncertainty using this model, each term in the probability model and its uncertainty distribution needs to be estimated and propagated through the law of total probability.

Bayesian simulation techniques may be used to assess the exposure distribution of contributing factors, i.e. the distribution of $\Pr(OF,SF)$ (see, e.g., Merrick et al., 2003). As more data tends to be available at the triggering incident level rather than at the accident level, the distribution of $\Pr(Incident|OF)$ may be assessed utilizing the traditional Bayesian estimation techniques. For example, by updating a Poisson process for the occurrences of mechanical failures with a gamma prior distribution on the rate of occurrences, with mechanical failure data. In this paper we shall concentrate on the assessment of $\Pr(Accident|Incident,OF,SF)$ where the contributing factors (OF,SF) are described by a vector $\underline{X}=(X_1,X_2,\ldots,X_p)$ and only limited accident data is available.

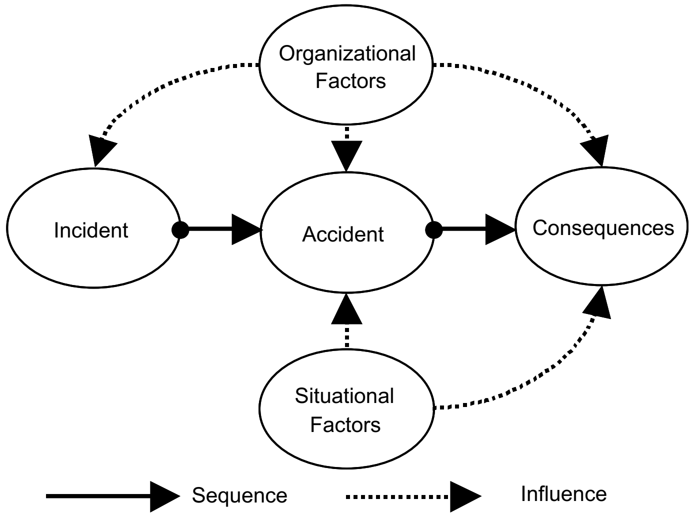


Fig. 2. The accident probability model.

For example in the WSF Risk Assessment only two collisions occurred over a period of 11 years (see Van Dorp et al., 2001). As an example, Table 1 provides a description of the contributing factors used in the WSF risk assessment. The heading "discretization" in Table 1 indicates the different number of possible scenarios for a contributing factor. For example, any of the following four traffic scenarios applies to the factor TS_1: meeting, passing, crossing astern and crossing the bow. Note that from the description in Table 1 it follows that a WSF Ferry may be interacting with more than one vessel at the same time.

The calculation model suggested for the accident probability given contributing factors $\underline{X}$ is given by (1), where $\underline{X} \in [0,1]^p$, $\underline{\beta} \in R^p$ and $P_0 \in (0,1)$. The covariates X_i, $i=1,\ldots,p$, are normalized so that $X_i=1$ describes the "worst" case scenario and $X_i=0$ describes the "best" case scenario. For example, for the 10th attribute X_{10} in Table 1, $X_{10}=1$ relates to the maximum wind speed typically observed in the given geographic area and $X_{10}=0$ relates to a wind speed of 0 knots. The calibration constant P_0 equals the accident probability when $\underline{X}=\underline{0}$.

In the previous example (dealing wind speed) the ordering from worst to best as it relates to an accident probability is self-evident, but this may not be the case for, for example, the second covariate in Table 1 indicating *vessel class*. In that case, a scale needs to be constructed ranking interacting vessel types

Table 1
Description of 10 contributing factors to $\Pr(Accident|Incident, \underline{X})$ in WSF risk assessment

	Designation	Description	Discretization
X_1	FR_FC	Ferry route—class combination	26
X_2	TT_1	1st Interacting vessel type	13
X_3	TS_1	Scenario of 1st interaction	4
X_4	TP_1	Proximity of 1st interaction	Binary
X_5	TT_2	2nd Interacting vessel type	5
X_6	TS_2	Scenario of 2nd interaction	4
X_7	TP_2	Proximity of 2nd interaction	Binary
X_8	VIS	Visibility	Binary
X_9	WD	Wind direction	Binary
X_{10}	WS	Wind speed	Continuous

according to a level of concern (from a collision perspective) when WSF captains or first mates encounter them on the water way. In the WSF risk assessment (see Van Dorp et al., 2001) a separate Bradley and Terry (1952) paired comparison procedure was used for that purpose, involving also WSF captains and first mates as experts. The Bradley–Terry procedure assumes that each object i is associated with a true scale value. For example, the value $X_2(i)$ is the scale value associated with the vessel type i, $i = 1,\ldots,13$, of the first interacting vessel (see Table 1). Next, experts are asked to respond whether a traffic interaction with a vessel of type j would be preferred over that of type i, $i,j = 1,\ldots,13, j \neq i$. Fig. 3 presents the resulting scale values $X_2(i)$, $i = 1,\ldots,13$, from the Bradley–Terry analysis for the second covariate in Table 1 involving 13 different vessel types.

It follows from Fig. 3 that when encountering these vessel types, the level of concern is the largest when encountering a Naval Vessel and the smallest when encountering a large WSF Ferry. One may argue that the construction of the scale in Fig. 3 introduces a motivational bias as Washington State Ferries consistently received the lowest rankings. On the other hand, when these results were presented to the Blue Ribbon Panel on Ferry Safety (see Van Dorp et al., 2001) it was noted that WSF Ferries interacting with WSF Ferries is an everyday occurrence involving common actors, rather than the far less frequent Naval Vessel whose captain is unknown to the WSF Ferry operators. In a similar manner, covariate scales had to be constructed for $X_1, X_3,\ldots, X_7$ to allow for the use of (1) and their contribution to $\Pr(Accident|Incident, \underline{X})$. Note that some of the elements in X may be used to describe interaction effects. For example, if X_1 relates to the Ferry Route-Ferry Class combination and X_2 relates to the traffic type of the first interacting vessel, one may introduce an 11th factor X_{11} equal to $X_1 \cdot X_2$ to model that accident probability may increase more (or less) as a result of a combined increase in both X_1 and X_2. In principle more complex interactions can be included.

Having selected the contributing factors for $\Pr(Accident|Incident, \underline{X})$ and having constructed the covariate scales of the elements in $\underline{X}$, a paired comparison questionnaire may be designed, each question comparing two different system states $\underline{X}^1$ and $\underline{X}^2$. Fig. 4 provides an example question appearing in one of the questionnaires used in the WSF risk assessment (see Van Dorp et al., 2001). For ease of comparison $\underline{X}^1$ and $\underline{X}^2$ (situations 1 and 2 in Fig. 4) differ only in one contributing factor. By circling a "1" or the midpoint

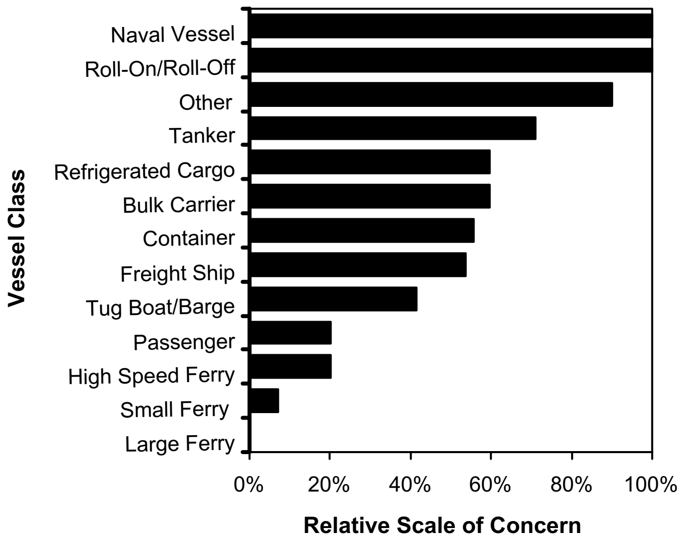


Fig. 3. Constructed covariate scale for interacting vessels.

Question: 32		48
Situation 1	**Attribute**	**Situation 2**
Super	**Ferry Class**	-
SEA-BAI	**Ferry Route**	-
Naval Vessel	**1st Interacting Vessel**	-
Crossing the bow	**Traffic Scenario 1st Vessel**	-
1 to 5 miles	**Traffic Proximity 1st Vessel**	-
Deep Draft	**2nd Interacting Vessel**	-
Crossing the bow	**Traffic Scenario 2nd Vessel**	-
1 to 5 miles	**Traffic Proximity 2nd Vessel**	-
more than 0.5 mile	**Visibility**	less than 0.5 mile
Along Ferry	**Wind Direction**	-
40 knots	**Wind Speed**	-
	9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9	
Situation 1 is worse	**<===================X===================>**	**Situation 2 is worse**

Fig. 4. An example question appearing in one of the questionnaires used in the WSF risk assessment.

of the scale, the expert has indicated that he/she judges the likelihood of a particular accident type to be the same in system state $\underline{X}^1$ as in system state $\underline{X}^2$. If he/she circles, e.g. the number 9 towards Situation 2 (i.e. to the right) we interpret that he/she considers the likelihood of a particular accident type to be 9 times as high in $\underline{X}^2$ as in $\underline{X}^1$ given a particular incident has occurred. In the WSF risk assessment (see Van Dorp et al., 2001) the focus was on collision accidents and incidents were further classified as propulsion, steering and navigation equipment failures, and human error.

If one is interested in paired comparison of accident risk between two different systems states $\underline{X}^1$ and $\underline{X}^2$ given an incident occurred, it is sufficient to estimate the parameter vector $\underline{\beta}$, as the relative accident probability in $\underline{X}^1$ compared to $\underline{X}^2$ (denoted by $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$) follows from (1) yielding

$$P(\underline{X}^1,\underline{X}^2|\underline{\beta}) = Exp\{\underline{\beta}^{\mathrm{T}}(\underline{X}^1-\underline{X}^2)\}. \tag{2}$$

Note that the relative accident probability is not restricted to the support [0,1] but $P(\underline{X}^1,\underline{X}^2|\underline{\beta}) \in [0,\infty]$ and

$$\log\{P(\underline{X}^1,\underline{X}^2|\underline{\beta})\} = \underline{\beta}^{\mathrm{T}}(\underline{X}^1-\underline{X}^2) \in (-\infty,\infty). \tag{3}$$

On the other hand, if one is interested in an absolute accident probability, one is required to estimate P_0 in addition to the parameter vector β. The calibration constant P_0 may be estimated by applying the law of total probability using all probability terms in Fig. 2, the maritime system simulation and average annual accident data, for example the 2 collisions over an 11 year period as was the case in the WSF risk assessment (see Van Dorp et al., 2001). In the following sections, the discussion will be limited to presenting prior and posterior analysis for relative accident probabilities given by (2).

3. The likelihood of a single expert's response

Let Y_j be the response of an expert to a paired comparison question j, comparing two different situations $\underline{X}_j^1$ and $\underline{X}_j^2$ in terms of accidents proneness given an incident has occurred (e.g. a navigation equipment failure), i.e.

$$Y_j = \text{Experts response to ratio } \frac{\Pr(Accident|Incident,\underline{X}_j^1)}{\Pr(Accident|Incident,\underline{X}_j^2)}.$$

Define

$$Z_j = \log Y_j,\ j = 1, \ldots, n,$$

to be experts' log response to question j. The response of the expert to such a question is uncertain and will assumed to be normal distributed such that

$$(Z_j|\mu_j, r) \sim N(\mu_j, r), \tag{4}$$

where $r = 1/\sigma^2$ is the precision that does not depend on the question index j and σ is the standard deviation of the normal distribution in (4), $\sigma > 0$. This is the most common uncertainty model encountered in practice, which seems to be appropriate at least given the support indicated by (3). Utilizing the structure of the accident probability model (1) and (3) we set

$$\mu_j = q_j^{\mathrm{T}} \underline{\beta}, \tag{5}$$

where $\underline{q}_j = (\underline{X}_j^1 - \underline{X}_j^2)$ is a $p \times 1$ vector. The relevance of the paired comparison of situations $\underline{X}_j^1$ and $\underline{X}_j^2$ appears in the distribution (4) of Z_j only via the vector $\underline{q}_j$ (cf. (5)). The likelihood of an expert responding z_j to question j, $f_{Z_j}(z_j)$, follows from (4) as

$$f_{Z_j}(z_j) \propto \sqrt{r} \exp\left\{-\frac{r}{2}(z_j - \mu_j)^2\right\}, \tag{6}$$

where the symbol $\propto$ means "being proportional to".

Suppose the expert answers n paired comparison questions defined by the vectors $q_j = (\underline{X}_j^1 - \underline{X}_j^2)$, $j = 1, \ldots, n$, define Q to be the $p \times n$ questionnaire matrix

$$Q = [\underline{q}_1, \ldots, \underline{q}_n] \tag{7}$$

and let the answers of the expert be summarized in the $n \times 1$ response vector

$$\mathscr{Z} = (z_1, \ldots, z_n). \tag{8}$$

Assuming conditional independence between an individual expert's responses to different questions given the precision r and parameter vector $\underline{\beta}$, the likelihood $\mathscr{L}(\mathscr{Z}|\underline{\beta}, r, Q)$ of an expert responding $\mathscr{Z}$ to questionnaire Q, may be derived from (6) as being proportional to

$$r^{\frac{n}{2}} \exp\left\{-\frac{r}{2}\left(\sum_{j=1}^{n} z_j^2 - 2\sum_{j=1}^{n} \mu_j z_j + \sum_{j=1}^{n} \mu_j^2\right)\right\}. \tag{9}$$

The conditional independence assumption implies that the sole source for *dependence* amongst an individual expert's responses to the different questions are the unknown precision r and the unknown parameter vector $\underline{\beta}$ (which seems to be reasonable). In addition, in a Bayesian analysis the standard conditional independence assumption given the unknown parameters is quite natural and is often not explicitly mentioned (see, e.g. Pulkkinen, 1994a). Substituting $\mu_j = q_j^{\mathrm{T}} \underline{\beta}$ (cf. (5)) in (9), yields

$$\begin{aligned} \mathscr{L}(\mathscr{Z}|\underline{\beta}, r, Q) &\propto r^{\frac{n}{2}} \exp\left\{-\frac{r}{2}\left(\sum_{j=1}^{n} z_j^2 - 2\left[\sum_{j=1}^{n} \underline{q}_j z_j\right]^{\mathrm{T}} \underline{\beta} + \underline{\beta}^{\mathrm{T}}\left[\sum_{j=1}^{n} \underline{q}_j \underline{q}_j^{\mathrm{T}}\right]\underline{\beta}\right)\right\} \\ &\propto r^{\frac{n}{2}} \exp\left\{-\frac{r}{2}(c - 2\underline{b}^{\mathrm{T}}\underline{\beta} + \underline{\beta}^{\mathrm{T}} A \underline{\beta})\right\}, \end{aligned} \tag{10}$$

where

$$A = \sum_{j=1}^{n} \underline{q}_j \underline{q}_j^{\mathrm{T}}; \quad \underline{b} = \sum_{j=1}^{n} \underline{q}_j z_j; \quad c = \sum_{j=1}^{n} z_j^2. \tag{11}$$

The matrix A will be referred to as the design matrix of the questionnaire Q. Note that, $A^T = A$. Hence, A is symmetric. Furthermore, for $\underline{x} \neq \underline{0}$ it follows that

$$\underline{x}^{\mathrm{T}} A \underline{x} = \underline{x}^{\mathrm{T}} \left[\sum_{j=1}^{n} \underline{q}_j \underline{q}_j^{\mathrm{T}} \right] \underline{x} = \sum_{j=1}^{n} \underline{x}^{\mathrm{T}} \underline{q}_j \underline{q}_j^{\mathrm{T}} \underline{x} = \sum_{j=1}^{n} (\underline{x}^{\mathrm{T}} \underline{q}_j)^2 > 0 \tag{12}$$

as long as the columns $\underline{q}_j$ of Q span R^p. If the latter condition holds for the questionnaire matrix Q, it follows from (12) that A is positive definite and symmetric and therefore invertible.

4. Prior distribution

To allow for a conjugate Bayesian analysis a multivariate normal/gamma prior is proposed for the joint distribution of $(\underline{\beta}, r)$ similar to the one described in West and Harrison (1989). Conjugate Bayesian analysis is motivated mainly by the desire to simplify calculations of the posterior probability. Nevertheless it proved to be a reliable approach yielding invariably meaningful results.

A Gamma$(\frac{\alpha}{2}, \frac{v}{2})$ will be defined on the precision r and is given by the pdf

$$\prod(r|\alpha, v) = \frac{\frac{v^{\frac{\alpha}{2}}}{2}}{\Gamma\left(\frac{\alpha}{2}\right)} r^{\frac{\alpha}{2}-1} \exp\left(-\frac{r}{2} v\right). \tag{13}$$

The distribution of $(\underline{\beta}|r)$ is assumed to be multivariate normal (MVN) with a prior $p \times 1$ dimensional mean vector $\underline{m}$ and $p \times p$ precision matrix $r\Delta$, i.e.

$$\prod(\underline{\beta}|r) \propto r^{\frac{p}{2}} \exp\left\{-\frac{r}{2}(\underline{\beta} - \underline{m})^{\mathrm{T}} \Delta (\beta - m)\right\}. \tag{14}$$

Hence, from the structure of the MVN it follows that $(r\Delta)^{-1}$ is the variance covariance matrix of $(\underline{\beta}|r)$. The joint prior distribution on $(\underline{\beta}, r)$ follows from (13) and (14) to be

$$\prod(\underline{\beta}, r) \propto r^{\frac{\alpha}{2}-1} \exp\left(-\frac{r}{2} v\right) \times r^{\frac{p}{2}} \exp\left\{-\frac{r}{2}(\underline{\beta} - \underline{m})^{\mathrm{T}} \Delta (\underline{\beta} - \underline{m})\right\}. \tag{15}$$

The marginal distribution of $\underline{\beta}$ may be derived from (15), yielding

$$\prod(\underline{\beta}) \propto \left[1 + \frac{1}{v}(\underline{\beta} - \underline{m})^{\mathrm{T}} \Delta (\underline{\beta} - \underline{m})\right]^{-\frac{\alpha+p}{2}} \tag{16}$$

and is recognized as a p-dimensional multivariate t-distribution with α degrees of freedom, location vector $\underline{m}$ and precision matrix

$$\frac{\alpha}{v} \Delta. \tag{17}$$

Note that, α/v in (17) is the mean value of the precision $r \sim$ Gamma$(\frac{\alpha}{2}, \frac{v}{2})$ and hence the marginal distribution of β integrates the precision given by (13) and that of $(\underline{\beta}|r)$ (cf. (14)). The marginal distribution of β_i, $i = 1, \ldots, p$, follows from (16) as a univariate t-distribution with α degrees of freedom, location parameter m_i and precision parameter $\frac{\alpha}{v}\delta_{ii}$, given by

$$\prod(\beta_i) \propto \left[1 + \frac{\delta_{ii}}{v}(\beta_i - m_i)^2\right]^{-\frac{\alpha+1}{2}}, \tag{18}$$

where δ_{ii} is the ith diagonal element of the precision matrix Δ. From (16) and (3) follows that the log-relative probability $log\{P(\underline{X}^1,\underline{X}^2|\beta)\}$ has a prior t-distribution with mean

$$\underline{m}^{\mathrm{T}}(\underline{X}^1-\underline{X}^2) \tag{19}$$

and precision

$$\frac{\alpha}{v}(\underline{X}^1-\underline{X}^2)^{\mathrm{T}}\Delta(\underline{X}^1-\underline{X}^2). \tag{20}$$

The prior distribution of the relative probability $P(\underline{X}^1,\underline{X}^2|\beta)$ (cf. (2)) thus follows a log-t distribution (see, e.g., McDonald and Butler, 1987) with parameters specified via (19) and (20).

4.1. Prior parameter specification

A prior chi-squared distribution with α degrees of freedom (equivalent to a gamma distribution Gamma$(\frac{\alpha}{2},\frac{v}{2})$ with $v=1$) will be selected for the prior distribution of precision r requiring only specification of the prior parameter α. From (13) it follows that $E[r|\alpha,v=1]=\alpha$. The prior parameter α will be set equal to the reciprocal of the variance of an expert responding to the n paired comparison questions completely at random and depends on the scale that is used in the paired comparison questions to collect the expert responses. In the example of Fig. 4, responses range from $\frac{1}{9},\frac{1}{8},\ldots\frac{1}{2},1,2,\ldots,9$ totaling 17 possible responses per question. With different responses being equally like and mutually independent for an expert responding at random and noting that $\log^2(x^{-1})=\log^2(x)$ it follows that a priori

$$\alpha=E[r|\alpha,v=1]=\frac{1}{\frac{2}{17}\sum_{k=2}^{9}\{\log(k)\}^2}\approx 0.380341. \tag{21}$$

Consistency within an individual expert's response can be observed when the posterior variance decreases as compared to an expert responding at random. The conjugacy of the posterior analysis will allow for straightforward sequential updating using the responses of the k individual experts. Agreement amongst the experts can be identified by further reduction (increase) in the posterior variance (precision) using sequential updating.

During the WSF risk assessment in 1998 geometric means amongst the expert responses were used in a classical log-linear regression analysis approach to assess relative accident probabilities given by (2). Using a best subset regression approach 6 interactions indicated Table 2 were selected and will also be used herein to allow for a comparison in Section 6 between the classical and Bayesian point estimates. Hence, the vector $\underline{\beta}$ to be utilized in our example in Section 6 will be a 1×16 vector.

For the distribution of $(\underline{\beta}|r)$ we may select a priori a location vector

$$\underline{m}=(0,\ldots,0)^{\mathrm{T}} \tag{22}$$

Table 2
Interaction variables associated with the contributing factors in Table 1

	Designation	Description
X_{11}	FR_FC·TT_1	Interaction
X_{12}	FR_FC·TS_1	Interaction
X_{13}	FR_FC·VIS	Interaction
X_{14}	TT_1·TS_1	Interaction
X_{15}	TT_1·VIS	Interaction
X_{16}	TS_1·VIS	Interaction

and the unit precision matrix

$$\Delta = \begin{pmatrix} 1 & & \emptyset \\ & \ddots & \\ \emptyset & & 1 \end{pmatrix}, \tag{23}$$

as long as the resulting marginal distributions of β_i (cf. (18)) are flat, or (perhaps more importantly) as long as the resulting prior distribution on the relative accident probabilities (2) are non-informative. The motivation for a non-informative prior is to "let the evidence speak" (i.e. the expert judgment) (see, e.g., Kaplan, 1997, p. 414). Expression (22) specifies that a priori none of the attributes contribute to accident risk and expression (23) indicates a priori independence between the elements of the parameter vector $\underline{\beta}$.

Fig. 5 below depicts the prior distribution on $(\underline{\beta},r)$ utilizing (21), (22) and (23). Fig. 5A depicts a graph of the prior density function of the precision r. Fig. 5B displays the 90% credibility intervals of $\beta_i, i=1,\ldots,16$ and Fig. 5C provides a graph of prior distribution of the relative probability $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$ associated with the

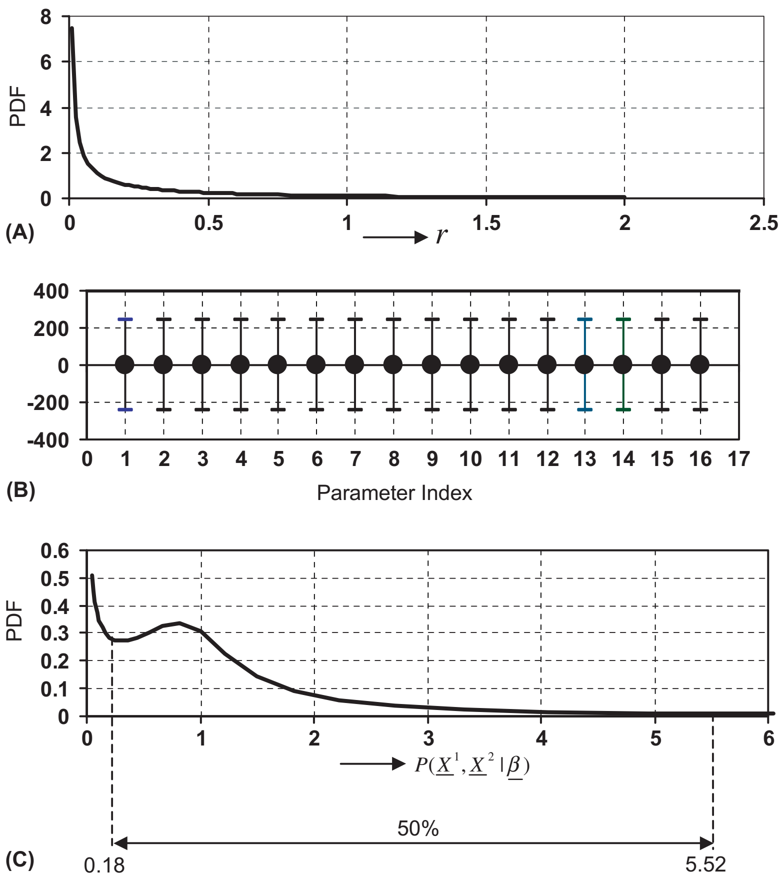


Fig. 5. Prior distribution on $(\underline{\beta},r)$ and $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$ (cf. (2)) for the two scenarios in Fig. 4. (A) Prior marginal distribution on r; (B) Prior 90% credibility intervals for the parameters β_i, $i=1,\ldots,16$; (C) Prior distribution of relative probability $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$ associated with Fig. 4.

paired comparison in Fig. 4. The probability density in Fig. 5C is one of a log-t distribution (see, e.g., McDonald and Butler, 1987) with prior parameters (cf. (19) and (20))

$$\underline{m}^{\mathrm{T}}(\underline{X}^1-\underline{X}^2)=0,\ \ \alpha=0.380341,\ \ v=1,\ \ \delta_{ii}=(\underline{X}^1-\underline{X}^2)^{\mathrm{T}}\varDelta(\underline{X}^1-\underline{X}^2)=4.$$

The prior median of $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$ equals 1 (indicating indifference in collision likelihood between system states $\underline{X}^1$ and $\underline{X}^2$). A 50% credibility interval of $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$ in Fig. 5A equals [0.181, 5.515]. A 75% credibility interval of $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$ equals $[2.012\cdot 10^{-5}, 4.971\cdot 10^4]$ (which is quite wide) and hence our prior specification utilizing (21)–(23) may be viewed as sufficiently non-informative.

Previous credibility intervals above and those in Fig. 5B were evaluated utilizing

$$A(u|\alpha,v,\delta_{ii})=\frac{1}{B(\frac{1}{2},\frac{\alpha}{2})}\sqrt{\frac{\delta_{ii}}{v}}\int_{m_i-u}^{m_i+u}\left[1+\frac{\delta_{ii}}{v}(\beta_i-m_i)^2\right]^{-\frac{\alpha+1}{2}}\mathrm{d}\beta_i,$$

where $A(u|\alpha,v,\delta_{ii})$ is the probability mass in a credibility interval $[m_i-u, m_i+u]$ around the location parameter m_i of a t-distribution with precision $\frac{\alpha}{v}\delta_{ii}$. The latter quantity $A(u|\alpha,v,\delta_{ii})$ is related to the well known incomplete beta function

$$B(x|a,b)=\frac{1}{B(a,b)}\int_0^x u^{a-1}(1-u)^{b-1}\,\mathrm{d}u, \tag{24}$$

where $a,b>0$, $x\in[0,1]$, and $B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ via the relationship

$$A(u|\alpha,v,\delta_{ii})=1-B\left(\frac{v}{v+\delta_{ii}u^2}\middle|\frac{\alpha}{2},\frac{1}{2}\right)$$

(see, e.g., Press et al., 1989). Numerical routines for evaluating the incomplete beta function (24) are widely provided in standard PC software such as Microsoft Excel. It should also be noted that due to the value of α (cf. (21)), the moments of β_i, at least a priori, do not exist. However, since the t-distribution is symmetric around m_i, a natural point estimate for β_i is provided by its median value m_i indicated in Fig. 5B, $i=1,\ldots,16$.

5. Posterior analysis

Applying Bayes theorem utilizing the likelihood (10), the prior distribution (15) and data specified via (7) and (8), it follows that the posterior distribution $\prod(\underline{\beta},r|\mathscr{X},Q)$ is proportional to

$$r^{\frac{n}{2}}\exp\left\{-\frac{r}{2}(c-2\underline{b}^{\mathrm{T}}\underline{\beta}+\underline{\beta}^T A\underline{\beta})\right\}\times r^{\frac{\alpha}{2}-1}\exp\left(-\frac{r}{2}\right)\times r^{\frac{p}{2}}\exp\left\{-\frac{r}{2}(\underline{\beta}-\underline{m})^{\mathrm{T}}\varDelta(\underline{\beta}-\underline{m})\right\},$$

where c, $\underline{b}$, and A are given by (11). Combining like terms we obtain

$$\prod(\underline{\beta},r|\mathscr{X},Q)\propto r^{\frac{\alpha+n}{2}-1}\exp\left\{-\frac{r}{2}(1+c+\underline{m}^{\mathrm{T}}\varDelta\underline{m})\right\}\times r^{\frac{p}{2}}$$
$$\times\exp\left\{-\frac{r}{2}\left(-2[\underline{b}+\varDelta\underline{m}]^{\mathrm{T}}\underline{\beta}+\underline{\beta}^{\mathrm{T}}[A+\varDelta]\underline{\beta}\right)\right\}. \tag{25}$$

Defining $\varDelta^u$ to be

$$\varDelta^u=A+\varDelta, \tag{26}$$

it follows from the symmetry and positive definiteness of A (cf. (12)) and $\varDelta$, that $\varDelta^u$ is symmetric and positive definite, and hence invertible. Implicitly defining $\underline{m}^u$ satisfying

$$[\underline{b}+\varDelta\underline{m}]^{\mathrm{T}}\underline{\beta}=[\varDelta^u\underline{m}^u]^{\mathrm{T}}\underline{\beta} \tag{27}$$

for all $\underline{\beta}$, it follows that

$$\underline{b} + \sum \underline{m} = \Delta^u \underline{m}^u \Longleftrightarrow \underline{m}^u = (\Delta^u)^{-1}(\underline{b} + \Delta \underline{m}). \tag{28}$$

Utilizing (27) and (28) we derive from (25) that

$$\prod(\underline{\beta}, r|\mathscr{Z}, Q) \propto r^{\frac{\alpha+n}{2}-1} \exp\left\{-\frac{r}{2}(1 + c + \underline{m}^{\mathrm{T}}\Delta\underline{m} - [\underline{m}^u]^{\mathrm{T}}\Delta^u\underline{m}^u)\right\} \times r^{\frac{p}{2}} \times \exp\left\{-\frac{r}{2}[\underline{\beta} - \underline{m}^u]^{\mathrm{T}}\Delta^u[\underline{\beta} - \underline{m}^u]\right\}. \tag{29}$$

From (29) it follows, utilizing (11), that $(\underline{\beta}|r, \mathscr{Z}, Q) \sim MVN(\underline{m}^u, r\Delta^u)$ where

$$\begin{cases} \Delta^u = \sum_{j=1}^{n} q_j q_j^{\mathrm{T}} + \Delta \\ m^u = (\Delta^u)^{-1}\left(\sum_{j=1}^{n} \underline{q}_j z_j + \Delta\underline{m}\right) \end{cases} \tag{30}$$

and $(r|\mathscr{Z}, Q) \sim \mathrm{Gamma}(\frac{\alpha^u}{2}, \frac{v^u}{2})$ with

$$\begin{cases} \alpha^u = \alpha + n \\ v^u = v + \sum_{j=1}^{n} z_j^2 + \underline{m}^{\mathrm{T}}\Delta\underline{m} - [\underline{m}^u]^{\mathrm{T}}\Delta^u\underline{m}^u \end{cases} \tag{31}$$

and $\underline{m}^u$ and Δ^u are given by (30). From (30), (31), (13) and (14) we deduce that the Bayesian updating procedure above is in fact a conjugate Bayesian analysis. In the next section we shall illustrate the inference procedure using the responses of eight experts to a paired comparison questionnaire containing 60 questions similar to the one in Fig. 4 and administered during the WSF risk assessment in 1998.

6. Example with data elicited during WSF risk assessment

An individual questionnaire was administered to experts for each of the following possible incidents on the Washington State Ferry: propulsion failure, steering failure, navigation equipment failure, human error, as well as an individual questionnaire given an incident (either human error or mechanical failure) which occurred on the nearby vessel. As an illustrative example, we shall demonstrate our Bayesian conjugate analysis utilizing the responses of the 8 experts to the questionnaire involving the navigation equipment failure to derive the posterior distribution of the relative accident probability given by (2) associated with Fig. 4. Combination of the responses of these 8 experts follows naturally by exploiting the conjugacy of the analysis in Section 3–5 through sequential updating.

During the WSF risk assessment in 1998 expert responses were aggregated by taking geometric means of their responses and using them in a classical log linear regression analysis approach to assess relative accident probabilities given by (2). Classical point estimates for the parameters β_i, $i=1,\ldots,16$, associated with the contribution factors (the so-called main effects) in Table 1 and interaction effects in Table 2 will be compared to their Bayesian counterparts following our Bayesian aggregation method.

6.1. The elements A, $\underline{b}$ and $\underline{c}$ of the likelihood given by (11)

Experts were instructed to assume that a navigation equipment failure had occurred on the Washington State Ferry and were next asked to assess how much more likely a collision is to occur in Situation 1 (good visibility in Fig. 4) as compared to Situation 2 (bad visibility in Fig. 4) taking into account the value of all the contributing factors. The additional factors in Fig. 4 (besides visibility) are used to assess interaction effects but also play a role in terms of designing a meaningful question. For example, a question that simply asks an expert to assess the likelihood of collision given a navigation equipment failure in bad visibility compared to good visibility is not meaningful since the expert would have to know for example whether another vessel nearby is crossing or passing and its proximity. Table 3 provides the answers of the eight experts to the question in Fig. 4. Note that Expert 8 responded (presumably inconsistently) that Situation 2 (with bad visibility) has a lower accident probability than Situation 1 (with good visibility). An expert aggregation method combines the responses in Table 3 into a single one.

The questionnaire consisted of sixty questions similar to the one displayed in Fig. 4. The questions were randomized in order and were distributed evenly over the 10 contributing factors in Table 1 (i.e. 6 questions per changing contributing factor). The 16×16 design matrix A of the questionnaire (cf. (11)) is of the following form:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{32}$$

where A_{11} is a 10×10 diagonal matrix with diagonal elements

$$(4.56, 4.33, 2.89, 6, 1.5, 2.44, 6, 6, 6, 0.375) \tag{33}$$

and associated with the contributing factors $X_1, \ldots, X_{10}$. (The matrix A_{11} in (32) is a diagonal matrix since the paired comparison scenarios $\underline{X}^1$ and $\underline{X}^2$ only differed in one covariate (see Fig. 4)). The matrix A_{22} in (32) is a symmetric 6×6 matrix with elements

$$\begin{bmatrix} 3.45 & 0.33 & 0 & 1.44 & 0.76 & 0 \\ 0.33 & 3.45 & 0.44 & 0.33 & 0 & 1 \\ 0 & 0.44 & 4.11 & 0 & 1 & 2.39 \\ 1.44 & 0.33 & 0 & 1.89 & 0.36 & 0.08 \\ 0.76 & 0 & 1 & 0.36 & 3.02 & 2 \\ 0 & 1 & 2.39 & 0.08 & 2 & 6.67 \end{bmatrix} \tag{34}$$

and associated with the interaction effects $X_{11}, \ldots, X_{16}$. Finally, the matrix $A_{21} = A_{12}^{\mathrm{T}}$ is a sparse 10×6 matrix

$$\begin{bmatrix} 1 & 2.82 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2.26 & 0 & 2.12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1.13 & 0 & 0 & 0 & 0 & 0 & 0 & 3.06 & 0 & 0 \\ 0 & 2.13 & 0.52 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.02 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1.56 & 0 & 0 & 0 & 0 & 5.33 & 0 & 0 \end{bmatrix} \tag{35}$$

with only positive elements associated with the contributing factors X_1, X_2, X_3 and X_8 that are included in the interaction effects $X_{11}, \ldots, X_{16}$. The questionnaire was designed in a manner such that the resulting matrix A is positive definite (and thus invertible), but equally important, involved meaningful paired comparisons consistent with realistic scenarios on the Puget Sound. The latter required maritime knowledge about the WSF Ferry system acquired by the team conducting the WSF Risk Assessment.

Table 3
Expert response to the paired comparison in Fig. 4

Expert index	1	2	3	4	5	6	7	8
Response	5	5	3	9	7	9	3	0.5

Fig. 6 below summarizes the vector $\underline{b}$ cf. (11) for each of the eight expert responses to 60 questions in terms of $\sum_{j=1}^{60} q_{ij} z_j$ for each of the contributing factors X_i, $i=1,\ldots,10$, in Table 1 and interaction effects X_i, $i=11,\ldots,16$, in Table 2. Hence, Fig. 6 consists of 16 histograms each one plotting the ith element of the vector $\underline{b}$ cf. (11) for all eight experts. From Fig. 6 we may (visually) assess the consistency in the expert judgment with respect to the ordering of the covariate scale of the elements X_i, $i=1,\ldots,16$. A positive (negative) value indicates agreement with the ordering of that particular scale. For example, the histogram in Fig. 6 associated with the contributing factor TP1 (Traffic Proximity of first interacting vessel) shows that all experts responded (not surprisingly) that vessels further away pose less (immediate) collision risk. The histogram in Fig. 6 associated with the contributing factor VIS provides a similar result to that in Table 3, i.e. that Expert 8 inconsistently rated lower visibility with lower collision risk throughout the questionnaire. The largest discrepancy with the ordering of a covariate scale amongst the 8 experts is observed in the first histogram and is associated with the variable FR-FC (Ferry Route-Ferry Class combination).

The elements $c = \sum_{j=1}^{60} z_j^2$ (cf. (11)) for each individual expert are provided in Table 4. Note that on aggregate particularly both Expert 3 and Expert 8 assessed lower collision likelihoods in their paired comparisons questions.

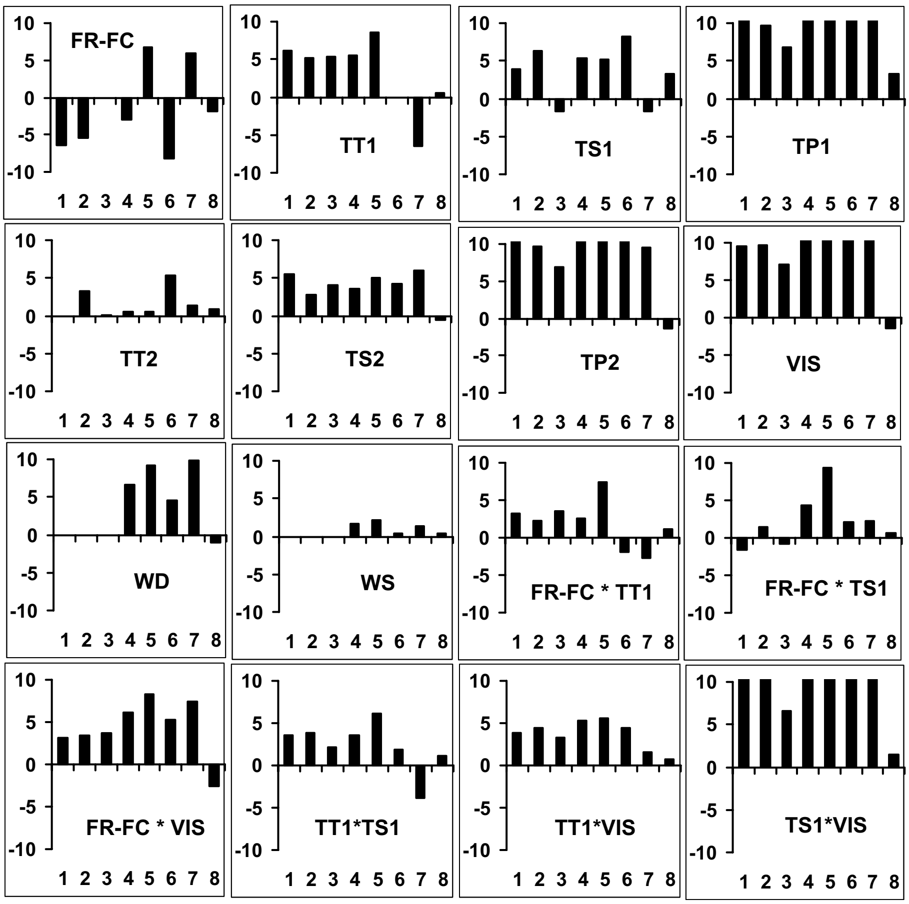


Fig. 6. Summary of individual expert response for 8 WSF experts in terms of ith element of the vector $\underline{b}$ (cf. (11)) for each of the contributing factors X_i, $i=1,\ldots,10$ in Table 1 and interaction effects X_i, $i=11,\ldots,16$ in Table 2.

Table 4
Values for c (cf. (11)) for the eight individual experts

Expert index	1	2	3	4	5	6	7	8
Scalar c	149.07	95.28	55.74	147.93	185.71	177.30	147.12	44.94

6.2. Posterior analysis

Utilizing the aggregate individual expert responses (vectors $\underline{b}$) in Fig. 6, the matrix A specified by (32)–(35), the scalars c in Table 4, we update the prior distribution of $(\underline{\beta},r)$ depicted in Fig. 5 in a Bayesian manner using sequential updating. The resulting posterior distribution on $(\underline{\beta},r)$ is displayed in Fig. 7. Fig. 7A contains a plot of a $\Gamma(\frac{\alpha^u}{2},\frac{v^u}{2})$ density with parameters

$$\alpha^u = 480.38, \quad v^u = 530.95. \tag{36}$$

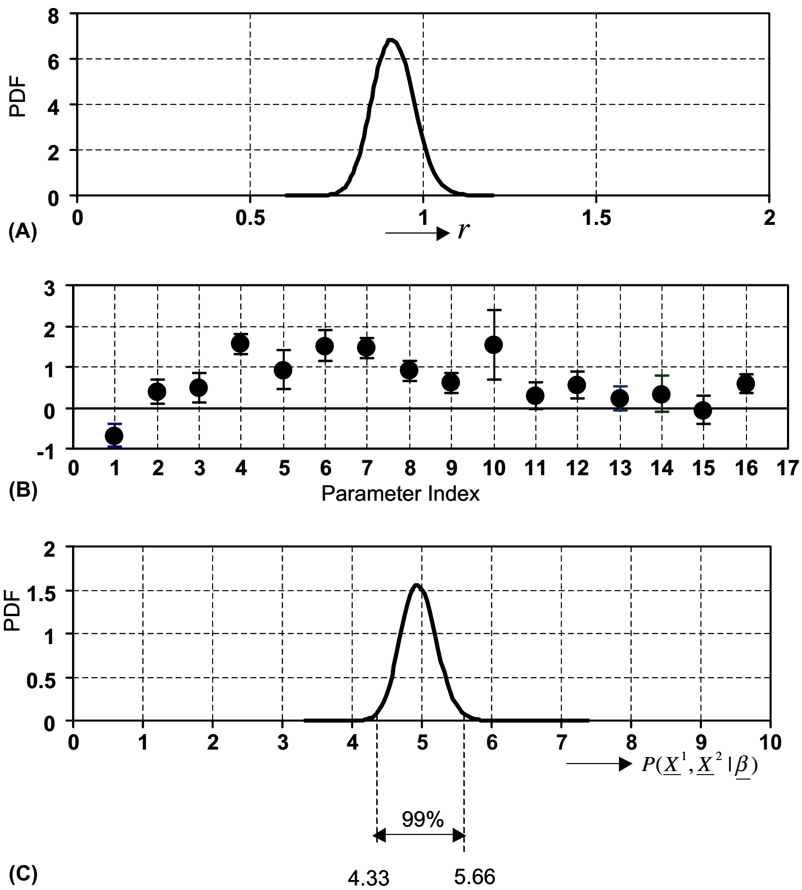


Fig. 7. Posterior distribution on $(\underline{\beta},r)$ and $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$ (cf. (2)) for the two scenarios in Fig. 4. (A) Posterior marginal distribution on r; (B) Posterior 90% credibility intervals for the parameters β_i, $i=1,\ldots,16$; (C) Posterior distribution of relative probability $P(\underline{X}^1,\underline{X}^2|\underline{\beta})$ associated with Fig. 4.

Fig. 7B displays 90% credibility intervals of the posterior distributions of β_i, $i=1,\ldots,16$ and the location parameters m_i^u. The posterior distribution of the parameter vector $\underline{\beta}$ is a multivariate t distribution with location vector $\underline{m}^u$ and precision matrix $\frac{\alpha^u}{\nu^u}\Delta^u$, where α^u, ν^u are given by (36) and

$$\Delta^u = \Delta + 8A$$

(cf. (26)) where the unit matrix Δ is given by (23) and the matrix A by (32)–(35). It can be concluded from Fig. 7B that traffic proximity of the first and second interacting vessel (X_4 and X_7, respectively), traffic scenario of the second interacting vessel X_6 and wind speed X_{10} are the largest contributing factors to accident risk. In addition, the manner in which the first interacting vessel approaches the ferry route–ferry class combination (X_{12}), i.e. crossing, passing or overtaking, and in what visibility conditions (X_{16}) are the largest interacting factors. The posterior location vector m^u is displayed in Fig. 8 together with their classical counterpart estimated via a log-linear regression method utilizing the geometric means of the expert responses. A remarkable agreement should be noted between the Bayesian and classical point estimates provided in Fig. 8, except for a discrepancy associated with the contributing factor WS (Wind Speed). From Fig. 7B, however, it follows that the classical point estimate associated with WS in Fig. 8 is well within the 90% credibility bounds of β_{10}. Finally, Fig. 7C displays the posterior distribution of the relative probability $P(\underline{X}^1, \underline{X}^2|\underline{\beta})$ associated with Fig. 4. We now have for the 50% posterior credibility interval of $P(\underline{X}^1, \underline{X}^2|\underline{\beta})$ the interval [4.78, 5.13]. (Compare this interval with the 50% prior one of [0.18, 5.52] in Fig. 5C.) In addition, the 99% posterior credibility interval [4.33, 5.66] of $P(\underline{X}^1, \underline{X}^2|\underline{\beta})$ is indicated in Fig. 7C (which is remarkably narrow compared to the prior 75% credibility interval of $[2.012\times10^{-5}, 4.971\times10^{4}]$) containing its median point estimate 4.94. Hence, Situation 2 inFig. 4 is approximately 5 times more likely to result in a collision than Situation 1 given that a navigation equipment failure occurred on the ferry.

Fig. 9 below provides a posterior analysis of point estimates α^u/ν^u of the precision r, where α^u and ν^u are given by (31). Fig. 9A depicts $E[r|Experti]$ obtained by updating the prior precision with the individual

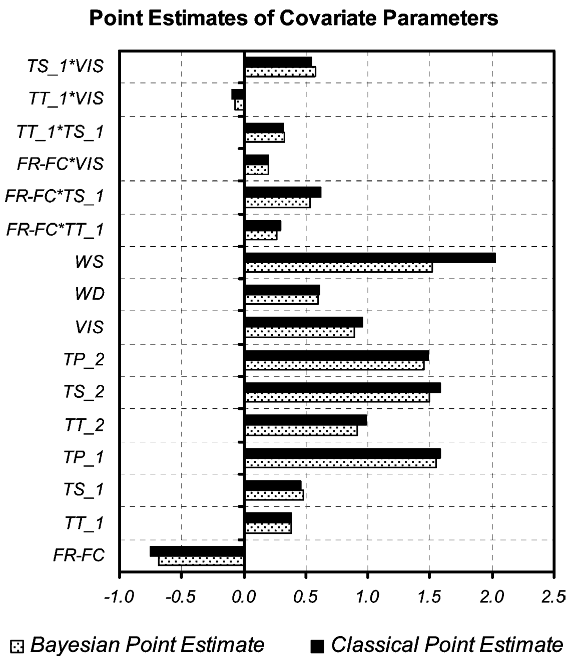


Fig. 8. Comparison of Bayesian and classical point estimates of the parameters β_i, $i=1,\ldots,16$.

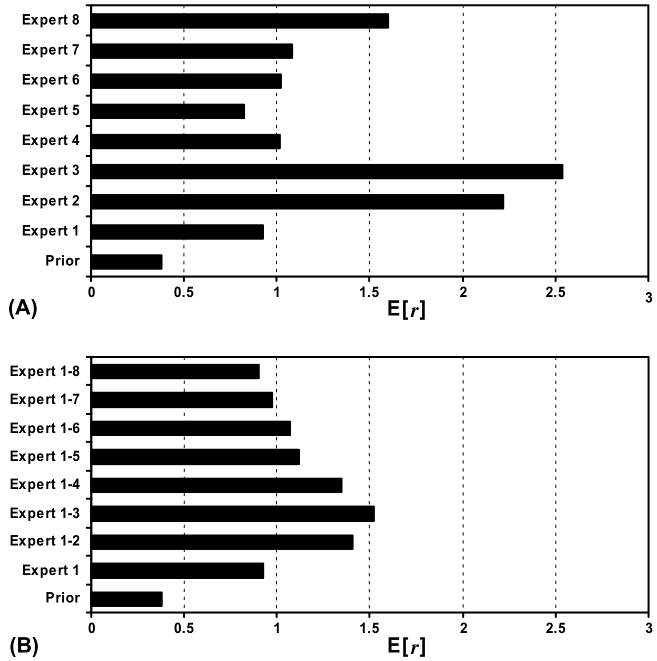


Fig. 9. Prior and posterior points estimates of the precision r (cf. (4) and (13)): (A) Individual posterior estimates for Experts i, $i = 1,\ldots,8$; (B) Sequential posterior estimates involving Experts 1 through i, $i = 1,\ldots,8$.

responses of Expert i, $i = 1,\ldots,8$. Fig. 9B displays $E[r|Expert1-i]$ derived using sequential updating involving Expert 1 through Expert i, $i = 1,\ldots,8$. From Fig. 9A it may be concluded that each expert responded consistently in the sense that posterior precision increased when compared to the precision of an expert responding at random (the prior precision in Fig. 9A). In addition, from Fig. 9B we conclude that at first agreement is present amongst Experts 1–3 due to a continued increase in posterior precision utilizing sequential updating. From Expert 4 onward and including Expert 8, however, a continued disagreement is observed in Fig. 9B due to a continued decline in posterior precision. Note the increasing pattern in Fig. 9A from Expert 5 on compared to the continued decreasing pattern in Fig. 9B from Expert 4 and up. The latter indicates that consistency of an individual expert response does not necessarily result in an increase in agreement amongst a group of experts.

7. Concluding remarks

A Bayesian aggregation method has been developed using responses from multiple experts to a paired comparison questionnaire to assess the distribution of relative accident probabilities. The classical analysis conducted during the WSF risk assessment only resulted in point estimates of relative accident probabilities, not full posterior distributional results as indicated in Fig. 7C. In addition, utilizing posterior distributional results for the parameter vector $\underline{\beta}$ credibility statements can be made for any arbitrary paired comparison. For example setting Situation 1 in (2) to the best possible scenario ($\underline{X}^1 = \underline{0}$) and Situation 2 to the worst possible scenario ($\underline{X}^2 = \underline{1}$) a 99% credibility interval of $P(\underline{X}^1, \underline{X}^2|\underline{\beta})$ equals [31142, 36749]. Therefore, informally, collision risk in the worst possible scenario differs at least by 4 orders of magnitude to that of the best possible scenario while taking uncertainty of the expert judgments into account.

Worst case scenario's however may have a very low incidence of occurrence, which is why all conditional probabilities in Fig. 2 and their uncertainties need to be estimated to assess the distribution of collision risk on for example a per year basis. This paper only provided distributional results for the relative probability given by (2). Merrick et al. (2003) assesses the distribution of $\Pr(OF, SF)$ using Bayesian Simulation techniques. A subsequent paper will integrate the approach herein with that of Merrick et al. (2003) to assess collision risk and its uncertainty in a Bayesian (and therefore coherent) manner.

Acknowledgements

This material is based in part upon work supported by the National Science Foundation under Grant Nos. SES 0213627 and SES 0213700. Any findings, opinions, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Special thanks to our colleague Sam Kotz for his helpful comments that improved both the content and presentation of this paper. We are also indebted to the referees and the editor of *EJOR*, whose comments improved the presentation of the first version.

References

Anderson, E., Hattis, D., Matalas, N., Bier, V., Kaplan, S., Burmaster, D., Conrad, S., Ferson, S., 1999. Foundations. Risk Analysis 19 (1), 47–68.

Bedford, T., Cooke, R., 2001. Probabilistic Risk Analysis: Foundations and Methods. Cambridge University Press, Cambridge, UK.

Bonano, E., Hora, S., Keeney, R., von Winterfeldt, D., 1989. Elicitation and use of expert judgment in performance assessment for high-level radioactive waste repositories. SAND89-1821, NUREG/CR-5411, Sandia National Laboratories, Albuquerque, NM.

Bradley, R., 1953. Some statistical methods in taste testing and quality evaluation. Biometrics 9 (1), 22–38.

Bradley, R.A., Terry, M., 1952. Rank analysis of incomplete block designs I. The method of paired comparisons. Biometrika 39, 324–345.

Clemen, R.T., 1989. Combining forecasts: a review and annotated bibliography. International Journal of Forecasting 5 (4), 559–583.

Clemen, R.T., Winkler, R., 1990. Unanimity and compromise among probability forecasters. Management Science 36 (7), 767–779.

Clemen, R.T., Winkler, R., 1999. Combining probability distributions from experts in risk analysis. Risk Analysis 19 (2), 187–203.

Cooke, R.M., 1991. Experts in Uncertainty: Opinion and Subjective Probability in Science. Oxford University Press, Oxford, UK.

Cox, D.R., 1972. Regression models and life tables (with discussion). Journal of the Royal Statistical Society, Series B 34 (2), 187–220.

Debreu, G., 1986. Theory of Value: An Axiomatic Analysis of Economic Equilibrium. Yale University Press, New Haven, CT.

DeWispelare, A., Herren, L., Clemen, R., 1995. The use of probability elicitation in the high-level nuclear waste recognition program. International Journal of Forecasting 11 (1), 5–24.

Foreman, E.H., Selly, M.A., 2002. Decision by Objectives: How to Convince Others That You Are Right, first ed. World Scientific Pub Co, River-Edge, NJ.

Garrick, B.J., 1984. Recent case studies and advancements in probabilistic risk assessment. Risk Analysis 4 (4), 267–279.

Genest, C., Zidek, J., 1986. Combining probability distributions: A critique and annotated bibliography. Statistical Science 1 (1), 114–148.

Kahneman, D., Slovic, P., Tversky, A., 1982. Judgments Under Uncertainty. Cambridge University Press, Cambridge, UK.

Kaplan, S., 1997. The words of risk analysis. Risk Analysis 17 (4), 407–417.

Koller, G.R., 2000. Risk Modeling for Determining Value and Decision Making. Chapman & Hall, New York.

Kumamoto, H., Henley, E., 1996. Probabilistic Risk Assessment and Management for Engineers and Scientists. IEEE Press, Piscataway, NJ.

McDonald, J.B., Butler, R.J., 1987. Some generalized mixture distributions with an application to unemployment duration. Review of Economics and Statistics 69 (2), 232–240.

Mendel, M., Sheridan, T., 1989. Filtering Information from Human Experts. IEEE Transactions on Systems, Man & Cybernetics 36 (1), 6–16.

Merrick, J.R.W., van Dorp, J.R., Harrald, J.R., Mazzuchi, T.A., Grabowski, M., Spahn, J.E., 2000. A systems approach to managing oil transportation risk in Prince William Sound. Systems Engineering 3 (3), 128–142.

Merrick, J.R.W., Dinesh, V., Singh, A., van Dorp, J.R., Mazzuchi, T., 2003. Propagation of uncertainty in a simulation-based maritime risk assessment model utilizing Bayesian simulation techniques. In: Chick, S., Sanchez, P.J, Ferrin, D., Morrice, D.J. (Eds.), Proceedings of the 2003 Winter Simulation Conference, pp. 449–455.

Morgan, M., Henrion, M., 1991. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge, UK.

Morris, P., 1974. Decision analysis expert use. Management Science 20 (9), 1233–1241.

Moslesh, A., Bier, V., Apostolakis, G., 1988. A critique of current practice for the use of expert opinions in probabilistic risk assessment. Reliability Engineering and System Safety 20 (1), 63–85.

Paté-Cornell, M., 1996. Uncertainties in risk analysis: six levels of treatment. Reliability Engineering and System Safety 54 (2–3), 95–111.

Press, W.H., Flannery, B.P., Teukolsky, S.A., Vettering, W.T., 1989. Numerical Recipes in Pascal. Cambridge University Press, Cambridge.

Pulkkinen, U., 1993. Methods for combination of expert judgments. Reliability Engineering and System Safety 40 (2), 111–118.

Pulkkinen, U., 1994a. Bayesian analysis of consistent paired comparisons. Reliability Engineering and System Safety 43 (1), 1–16.

Pulkkinen, U., 1994b. Gaussian paired comparison models. Reliability Engineering and System Safety 44 (2), 207–217.

Roeleven, D., Kok, M., Stipdonk, H.L., de Vries, W.A., 1995. Inland waterway transport: modeling the probabilities of an accident. Safety Science 19 (2–3), 191–202.

Saaty, T., 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology, New York 15, 234–281.

Saaty, T., 1980. The Analytic Hierarchy Process. McGraw-Hill.

Shrader-Frechette, K., 1985. Risk Analysis and Scientific Method. Reidel, Dordrecht, The Netherlands.

Simpson, E.H., 1951. The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society, Series B 13 (2), 238–241.

Thurstone, L.L., 1927a. A law of comparative judgment. Psychology Review 34, 273–286.

Thurstone, L.L., 1927b. Psychophysical analysis. American Journal of Psychology 38, 368–389.

Van Dorp, J., Merrick, J., Harrald, J., Mazzuchi, T., Grabowski, M., 2001. A Risk Management Procedure for the Washington State Ferries. Risk Analysis 21 (1), 127–142.

Vose, D., 1996. Risk Analysis: A Quantitative Guide to Monte Carlo Simulation Modeling. Wiley, New York.

West, M., Harrison, J., 1989. Bayesian Forecasting and Dynamic Models. Springer-Verlag, New York.

Winkler, R., 1981. Combining probability distributions from dependent information sources. Management Science 27 (5), 479–488.

Winkler, R., 1996. Uncertainty in probabilistic risk assessment. Reliability Engineering and System Safety 54 (2–3), 127–132.