Speaking the Truth in Maritime Risk Assessment

Jason R. W. Merrick — Virginia Commonwealth University

J. Rene van Dorp — The George Washington University

Abstract

Several major risk studies have been performed in recent years in the maritime transportation domain. These studies have had significant impact on management practices. The first, the Prince William Sound Risk Assessment, was reviewed by the National Research Council and found to be promising but incomplete, as the uncertainty in its results was not assessed. The difficulty in incorporating this uncertainty is the different techniques that need to be used to model risk in a dynamic and data-scarce application area. In this paper, we combine a Bayesian simulation of the occurrence of situations with accident potential and a Bayesian multivariate regression analysis of the relationship between factors describing these situations and expert judgments of accident risk. These techniques are applied to a risk case study involving an assessment of the effects of proposed ferry service expansions in San Francisco Bay. This paper can be considered an innovative application of Bayesian simulation and Bayesian multivariate regression to assess uncertainty in risk analysis of a maritime transportation system.

Keywords: Uncertainty Analysis; Risk Analysis; Bayesian simulation; Bayesian multivariate regression; Maritime Transportation; Expert judgment.

1. Introduction

Maritime transportation is a critical part of the US economy; excluding Mexico and Canada, 95 percent of foreign trade and 25 percent of domestic trade depends on maritime transportation, cargo worth a total of $1.0 trillion of per year (National Research Council 2000, page 53). However, examples of accidents are easy to recollect; the grounding of the *Exxon Valdez*, the capsize of the *Herald of Free Enterprise* and the *Estonia* passenger ferries are some of the most widely publicized accidents in maritime transportation. The consequences of these accidents ranged from severe environmental damage to large-scale loss of life, but also severe economic problems for the companies involved. The Exxon Valdez disaster cost Exxon $2.2 billion in clean up costs alone. This leads to the immediate questions of how to prevent such accidents in the future and how to mitigate their consequences if they should occur.

Risk management has become a major part of operating decisions for companies in the maritime transportation sector and thus an important research domain (National Research Council, 2000). Early work concentrated on assessing the safety of individual vessels or marine structures, such as nuclear powered vessels (Pravda & Lightner, 1966), vessels transporting liquefied natural gas (Stiehl, 1977) and offshore oil and gas platforms (Paté-Cornell, 1990). More recently, Probabilistic Risk Assessment (Bedford and Cooke, 2002) has been introduced in the assessment of risk in the maritime domain (Roeleven et al., 1995; Kite-Powell, 1996; Slob, 1998; Fowler and Sorgard, 2000; Trbojevic and Carr, 2000; Wang, 2000; Guedes Soares and Teixeira, 2001).

The Prince William Sound (PWS) Risk Assessment (Merrick et al., 2000, 2002), Washington State Ferries (WSF) Risk Assessment (van Dorp et al. (2001) and an exposure assessment for ferries in San Francisco Bay (Merrick et al., 2003) are three examples of successful risk studies in this domain, combining system simulation with probabilistic risk assessment techniques. Their results have been used in major investment decisions and have played a significant role in the management of maritime transportation in the US. Figure 1 shows the risk intervention effectiveness estimates from the WSF Risk Assessment. The actual risk intervention cases modeled are described in van Dorp et al. (2001). The figures shows the total percentage reduction in collision probability for the WSF system for various risk management alternative broken down by the severity of the accidents, classified as Minimum Required Response Times (MRRT): less than one hour, between one and six hours and above six hours.

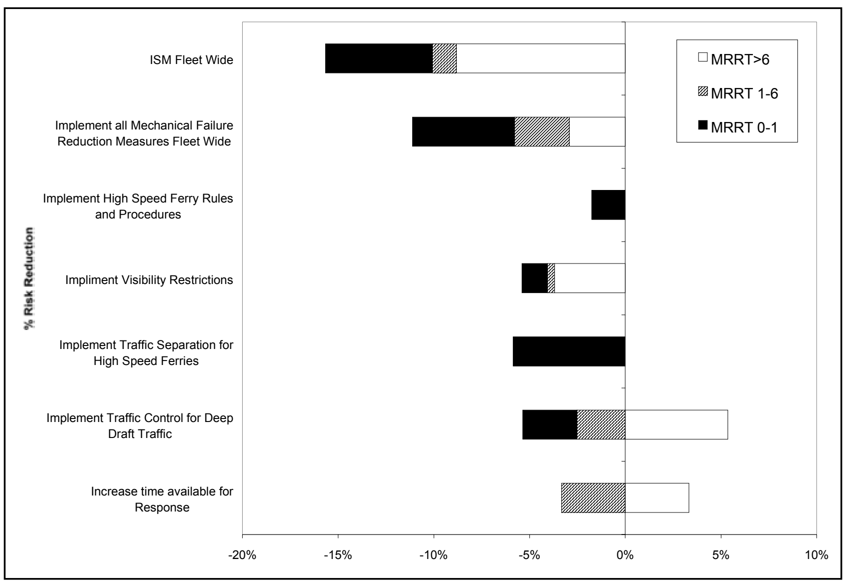

Figure 1. An assessment of risk intervention effectiveness for proposed safety improvements for the Washington State Ferries.

As another example, Figure 2 shows the results from an analysis of proposed ferry service expansions in San Francisco Bay. The estimates show the frequency of interactions between ferries and other vessels for the current ferry system (Base Case) and three alternative expansion scenarios which increase the total number of ferry transits per year.

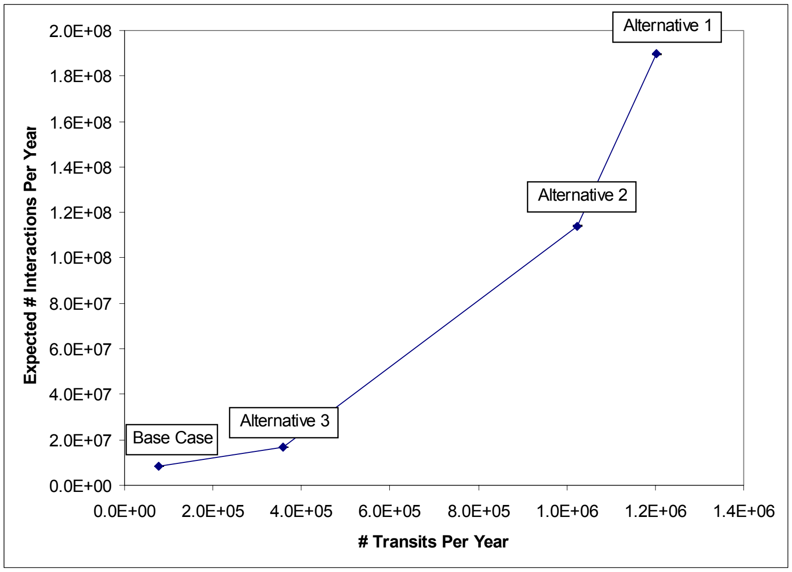

Figure 2. An assessment of alternative expansion scenarios for ferries in San Francisco Bay.

One problem with the representations in Figure 1 and Figure 2 is the apparent finality of the results. The decision-maker is led to believe that the results are definitive and are in no way uncertain. In fact, the National Research Council performed a peer review of the PWS Risk Assessment and concluded that the underlying methodology shows "promise" to serve as a systematic approach for making risk management decisions for marine systems (National Research Council 1998). However, to speak the truth in maritime risk assessments, the degree of uncertainty needed to be communicated (Kaplan 1997). *"Risk management … should answer whether evidence is sufficient to prove specific risks and benefits"* (A. Elmer, President, SeaRiver Maritime, Inc. in National Research Council, 2000).

In this article, we discuss an innovative application of Bayesian simulation and Bayesian multivariate regression to perform a complete assessment of risk and uncertainty for dynamic systems. These techniques have not been applied previously in a risk assessment setting. The methodology has applications beyond maritime accident risk, such as port security and aviation safety and security. A summary of the article is as follows. Section 2 discusses uncertainty and how it is best represented in risk analysis. The framework for a full uncertainty analysis of the results of the maritime probabilistic risk assessment models is summarized in Section 3. The results of an uncertainty case study are offered in Section 4, where the robustness of conclusions drawn in a study of ferry expansions in San Francisco Bay are assessed. Conclusions are drawn in Section 5.

2. Uncertainty Analysis

The presence of uncertainty in analyzing risk is well recognized and discussed in the literature. However, these uncertainties are often ignored or under-reported in studies of controversial or politically sensitive issues (Pate-Cornell, 1996). Two types of uncertainty are discussed in the literature, aleatory uncertainty (the randomness of the system itself) and epistemic uncertainty (the lack of knowledge about the system). In a modeling sense, aleatory uncertainty is represented by probability models that give probabilistic risk analysis its name, while epistemic uncertainty is represented by lack of knowledge concerning the parameters of the model (Parry, 1996). In the same manner that addressing aleatory uncertainty is critical through probabilistic risk analysis, addressing epistemic uncertainty is critical to allow meaningful decision-making. Cooke (1997) offers several examples of the conclusions of an analysis changing when uncertainty is correctly modeled.

While epistemic uncertainty can be addressed through frequentist statistical techniques such as bootstrap or likelihood based methods (Frey and Burmaster, 1999), the Bayesian paradigm is widely accepted as a method for dealing with both types of uncertainty (Apostolakis, 1978; Mosleh et al., 1988; Hora, 1996; Hofer, 1996; Cooke, 1991). However, as pointed out by Winkler (1996), there is no foundational Bayesian argument for the separation of these types of uncertainty. Ferson and Ginzburg (1996) use the terminology variability for aleatory uncertainty and ignorance for epistemic. Winkler's argument essentially says that variability is purely ignorance of which event will occur.

The distinction of types of uncertainty, however, does have certain uses in the risk assessment process (Anderson et al., 1999). Specifically, the distinction is useful when explaining model results to decision-makers and the public and when expending resources for data collection. In the communication case, the distinction must be drawn between the statements "we don't know if the event will occur" and "we don't know the probability that the event will occur." In the data collection case, epistemic uncertainty can be reduced by further study and data collection, whereas aleatory uncertainty is irreducible, as it is a property of the system itself (Hora, 1996). Bayesian modeling can allow for the distinction and handle the underlying differences inherently. Monte Carlo simulation (Vose, 2003) can be used to propagate uncertainty through a model (requiring significant computer power), while Bayesian analytical techniques can be used for analyzing data and expert judgments (Cooke, 1991).

3. Modeling Uncertainty in Dynamic Risk Assessment

We will use the examples of ferry risk assessment to demonstrate the application of Bayesian simulation and Bayesian multivariate regression to assessing risk for a dynamic system, such as a maritime transportation system. We shall use the expert judgments from the Washington State Ferries Risk Assessment analyzed using the Bayesian multivariate regression techniques developed in Merrick et al. (2004b) and apply them to the output from the Bayesian simulation of San Francisco Bay ferries described in Merrick et al. (2004a). While both studies considered ferries, these results should not be taken as a definitive analysis for either application as the experts from WSF were not considering San Francisco Bay when responding to the expert surveys. In each of these studies, one type of accident was considered, specifically collisions between a ferry and another vessel. Collisions occur within a situation defined by factors that affect the probability of occurrence. Table 1 shows the factors that were used to describe the situations in the WSF Risk Assessment.

3.1 A Probabilistic Risk Framework

The accident probability model is based on the notion of conditional probability, conditioning on the factors that determine the level of accident potential in a situation. To estimate the probability of a collision, we sum over the possible situations giving

$$P(Collision) = \sum_{j=1}^{k} P(Collision \mid Situation_j) P(Situation_j) \qquad (1)$$

where $Situation_j$ denotes the possible combinations of values of the factors in Table 1 for $j = 1, \ldots, k$ and k is the total number of possible combinations (2,163,200 in Table 1).

Table 1. The risk factors included in the expert judgment questionnaires.

Variable	Description	Notation	Values
X_1	Ferry route and class	FR_FC	26
X_2	Type of 1st interacting vessel	TT_1	13
X_3	Scenario of 1st interacting vessel	TS_1	4
X_4	Proximity of 1st interacting vessel	TP_1	Binary
X_5	Type of 2nd interacting vessel	TT_2	5
X_6	Scenario of 2nd interacting vessel	TS_2	4
X_7	Proximity of 2nd interacting vessel	TP_2	Binary
X_8	Visibility	VIS	Binary
X_9	Wind direction	WD	Binary
X_{10}	Wind speed	WS	Continuous

Thus the accident probability model consists of two parts:

- $P(Situation_j)$: the probability that particular combination of values of the factors occurs in the system
- $P(Collision \mid Situation_j)$: the probability that an accident occurs in the defined situation.

To perform an assessment of the risk of an accident using this model, both terms in the probability model need to be estimated.

The system simulation is used to count the occurrence of situations with different values of the defining factors. A simulation of the maritime transportation system is created incorporating vessel movements and environmental conditions. A situation is counted for each situation with accident potential, in the case of collisions this occurs when a vessel is considered to be interacting with a ferry (see van Dorp et al. (2001) for a definition of interactions with collision potential). A multi-year simulation is run and for each time period in the simulation the situations that occur are counted. Thus the average yearly frequency of situations with particular values of the factors, denoted $P(Situation_j)$, could be estimated using the simulation. The use of a system simulation also allows for the system wide evaluation of risk reduction and risk migration effects potentially associated with the implementation of particular risk intervention measures (see, e.g., Merrick et al., 2000, Merrick et al., 2002). Classical simulation techniques were used in the PWS and WSF studies, thus only point estimates of $P(Situation_j)$ were obtained.

The next step in the estimation of accident frequency is to estimate the conditional probability of accidents $P(Collision \mid Situation_j)$. The preferred method for estimating these probabilities is through the statistical analysis of accident data. However, expert judgment elicitation is often crucial in performing risk analyses (Cooke, 1991). In both the PWS and WSF Risk Assessments less than three relevant accidents had been recorded. Thus the analysis had to rely, at least in part, on expert judgment. The expert judgment method used to estimate $P(Collision \mid Situation_j)$ was based on pairwise comparisons and the expert responses were analyzed using classical statistical regression techniques. Thus again, only point estimates were obtained.

To address the uncertainties in the PWS/WSF Risk Assessment approach in a comprehensive and coherent manner we need to separately address uncertainty in the simulation estimates of $P(Situation_j)$ and uncertainties in the experts' assessments of the conditional probabilities and $P(Collision \mid Situation_j)$. We must then propagate these uncertainties though the framework expressed by (1).

3.2 Bayesian Simulation of a Maritime Transportation

Bayesian simulation differs from classical simulation analysis in that probability distributions are used to represent the uncertainty about model parameters rather than point estimates and confidence intervals. Such treatment is applied to both random inputs to the model and the outputs from the model. In the language of uncertainty, classical simulation models only aleatory uncertainty, while Bayesian simulation models both the aleatory and epistemic uncertainty. In this section, we discuss the development of a Bayesian simulation of the San Francisco Bay area. A classical simulation was developed by the authors for a study examining the effect of proposed service expansions under consideration by the California legislature (Merrick et al. 2003). As part of our uncertainty modeling, Merrick et al. (2004a) extended the SF Bay simulation model using Bayesian input and output modeling techniques.

In the existing simulation, the ferry transits were based on fixed schedules for the current ferry system and for each of the alternative expansion plans. Visibility and wind conditions were incorporated by tracing large databases of environmental data obtained from National Oceanographic and Atmospheric Administration (NOAA) observation stations in the study area. However, the arrivals for non-ferry traffic was based an historical data. Input uncertainty should be incorporated in the analysis to reflect the limited data available to populate the parameters of the arrival processes in a simulation model (Chick 2001). Thus Bayesian renewal process models of traffic arrivals were created in Merrick et al. (2004a) for all 5,277 arrivals processes.

The presence of input uncertainty means that there will be uncertainty in the outputs as well. This will include aleatory uncertainty as this is a stochastic simulation, but also epistemic uncertainty as the simulation is run for a finite period. In our risk assessment methodology, the data obtained from the simulation in each replication will be the number of vessel interactions occurring in each replication of the simulation, denoted $N_{r,j}$, for the r-th replication ($r = 1,...,s$ for s replications) and the j-th combination of values of the factors ($j = 1,...,k$). However, as we wish to propagate uncertainty throughout the overall model, a probability model will be hypothesized for these output statistics. Chick (1997) notes that this can be thought of as a Bayesian version of metamodeling (Law and Kelton 2001).

As our output data is in the form of a count, the number of vessel interactions for the j-th combination of values of the factors can be naturally modeled using a Poisson distribution with rate μ_j, with a conjugate gamma distributed prior on μ_j with shape α_j and scale γ_j. The posterior distribution of the expected vessel interactions frequency for the j-th combination of values of the factors is given by

$$\left(\mu_j \mid n_{1,j},\ldots,n_{s,j}\right) \sim gamma\left(\alpha_j + \sum_{i=1}^{s} n_{i,j}, \gamma_j + s\right) \tag{2}$$

The predictive distribution of $P(Situation_j)$ is then a Poisson-gamma distribution in the sense of Bernado and Smith (2000). Note that the epistemic uncertainty here can be reduced by running longer simulations, the aleatory uncertainty cannot; this would require additional traffic data.

3.3 Bayesian Multivariate Regression for Expert Judgments

Merrick et al. (2004b) propose a multivariate Bayesian analysis of expert judgments for an extended form of pairwise comparisons (Bradley and Terry, 1952) that accounts for correlations between the experts' responses (Clemen and Reilly, 1999). The aim of the expert elicitation method, as applied to maritime risk, is to estimate the effect of multiple factors on the probability of a collision, denoted $P(Collision \mid Situation_j)$. An example of the form of the questions drawn from the WSF risk assessment project is shown in Figure 3. Note that in each comparison, the situation is completely described in terms of the factors and only one factor is changed between the two situations the expert is asked to compare.

Situation 1	Attribute	Situation 2
Issaquah	Ferry Class	-
SEA-BRE(A)	Ferry Route	-
Navy	1st Interacting Vessel	Product Tanker
Crossing	Traffic Scenario 1st Vessel	-
< 1 mile	Traffic Proximity 1st Vessel	-
No Vessel	2nd Interacting Vessel	-
No Vessel	Traffic Scenario 2nd Vessel	-
No Vessel	Traffic Proximity 2nd Vessel	-
> 0.5 Miles	Visibility	-
Along Ferry	Wind Direction	-
0	Wind Speed	-
	Likelihood of Collision	
	9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9	

Figure 3. An example of the question format

The responses to the questions are in terms of relative probabilities of the event in the two situations. Thus, if the expert circles a “1”, this means they believe that the two probabilities would be equal, or if the expert circles a “9” on the right (left) then they believe the ratio of the probabilities is 9 (1/9) (Saaty, 1977).

The form of the underlying probability model is assumed to be

$$P(Collision \mid Situation_j) = p_0 \exp\left\{ Situation_j{}^T \beta \right\}, \tag{3}$$

where p_0 is a baseline probability of a collision and β is a vector of factor effect parameters. Due to this choice of form, the ratio of probabilities will be equal to $\exp\left(\left(Situation_L - Situation_R \right)^T \beta \right)$, where $Situation_L$ and $Situation_R$ are the vectors of factors for the situations on the left and right sides of the question respectively. Thus, if we equate the natural logarithm of the experts’ responses and the corresponding model terms, the analysis can be performed using linear regression techniques.

Szwed et al. (2004) develop a conjugate Bayesian analysis assuming that the experts respond independently. Merrick et al. (2004b) extend this Bayesian analysis to account for the correlations between the responses of the experts by assuming a multivariate normal distribution on the experts’ judgment errors in the manner of Winkler (1981). Suppose we ask p experts to respond to N such questions about q factors. We use the notation $Situation_j = \left(x_{j,1}, \ldots, x_{j,q} \right)$ to denote the differences between the q factors for the j-th question and $y_{j,e}$ for the response to the j-th question by expert e. The multivariate regression model used can be written as

$$\mathbf{Y} = \mathbf{X}\beta \underline{1}^T + \mathbf{U}, \tag{4}$$

where $\mathbf{X}$ is a $(N \times q)$ matrix of differences between the q covariates for N questions, $\mathbf{U}$ is a $(N \times p)$ vector of residual errors, $\beta = (\beta_1, \ldots, \beta_q)^T$ is the vector of regression parameters from (3) and $\underline{1} = (1_1, \ldots, 1_q)^T$ is a vector of p 1's. The multivariate regression model is completed by assuming that the rows of $\mathbf{U}$ are distributed according to a multivariate normal with a zero mean vector and covariance matrix $\mathbf{\Sigma}$.

A natural conjugate analysis is possible by assuming that Σ has an inverse Wishart prior distribution of dimension p with parameter matrix **G** and *m* degrees of freedom and, conditioned on Σ, β has a multivariate normal prior distribution with mean ϕ and variance $\sigma_\Sigma^2 \mathbf{A}$, where $\sigma_\Sigma^2 = \frac{1}{p^2} \underline{1}^T \mathbf{\Sigma} \underline{1}$ is the average variance across all experts. ϕ, $\mathbf{A}$, $\mathbf{G}$ and m are arbitrary prior hyperparameters determined by the decision maker. Given the responses of the experts and applying the minimum variance weighted average of Newbold and Granger (1974), the posterior predictive distribution of $\bar{y}$ is shown in Merrick et al. (2004b) to be a student-t distribution with $m + p + q - 1$ degrees of freedom given by

$$\bar{y}^* \sim student - t\left(\mathbf{x}^{*T}\left(\mathbf{X}^T\mathbf{X} + \mathbf{A}\right)^{-1}\left(\mathbf{X}^T\mathbf{X}\overline{\mathbf{B}} + \mathbf{A}\phi\right), (m+q)\underline{1}^T\left(\mathbf{G} + \mathbf{V}\right)^{-1}\underline{1}\mathbf{x}^{*T}\left(\mathbf{X}^T\mathbf{X} + \mathbf{A}\right)\mathbf{x}^*\right) \quad (5)$$

To complete the prediction recall that the regression was on the natural logarithm of the experts' assessments of the ratios of probabilities, so the actual predictions of the probability is a log student-t distribution.

3.4 Propagating Uncertainties

To perform a full uncertainty analysis of such a maritime risk model, we had to obtain Bayesian predictive distributions for each term in the model, $P(Situation_j)$ and

$P(Collision \mid Situation_j)$, and then propagate the uncertainty expressed in these distributions through the calculations in (1). Given the development in Sections 3.2 and 3.3, the predictive distribution of $P(Collision, Situation_j)$ in (1) cannot be obtained in closed form as the multiplication of a Poisson-gamma distribution with a log student-t distribution does not result in a known form. Monte Carlo simulation is the most commonly used tool for propagating uncertainty through a risk analysis model (Vose 2003). As noted by Winkler (1996) analytical solutions should be used if at all possible. In most cases though, closed form solutions are not possible and the brute force simulation method must be used (Pate-Cornell 1996). To perform Monte Carlo analysis for our model, values for all the parameters of the model are sampled at the beginning of each calculation of (1). These values are then used in the calculation and the value of $P(Collision)$ recorded. Thus samples of the posterior distribution of $P(Collision)$ are obtained and descriptive statistics of the distribution can be estimated.

Such calculations do require significant computational effort. The samples from $P(Situation_j)$ include a sample from the posterior distribution of μ_j in (2) and $N_j \mid \mu_j$, a Poisson distribution, for $j = 1, \ldots, k$. The samples for the conditional probabilities $P(Collision \mid Situation_j)$ are taken from the log student-t distribution in (5). For the simulation of the current SF Bay ferry system, the calculation time for a sample of 1000 values takes approximately 16,000 seconds or about 4½ hours. This does not include the 2 hours that 10 replications of a one-year simulation takes before this analysis can be performed. These times increase significantly as the number of ferry transits increase in the alternatives considered. Thus such uncertainty analysis on a large-scale model like this is highly computationally intensive even with the efficient sampling algorithms we applied.

However, such Monte Carlo analysis is referred to in the language of parallel computing as *embarrassingly parallel*. This means that the computation can be broken down in to independent pieces for calculation on parallel processors or separate computers and then the reconstituted for final analysis. For a 10 processor set-up, we may run one-year of the simulation on each processor instead of 10 years on one processor. As the sufficient statistics for the Bayesian meta-model on the simulation are the sum of the number of occurrences and the number of years simulated, the results from each processor are combined by summing the number of occurrences of each combination of values of the factors. The combined simulation data can then be passed back to the 10 processors for each to sample 100 values from the posterior distribution of $P(Collision)$. The total run time will now be one-tenth as long plus a small time for passing data after the simulation runs and the sampling. While this is not the frontier of work on parallel computing, this simple application can make the difference between running all the analysis needed for a complete decision and not.

4. An Uncertainty Analysis Case Study

In an effort to relieve congestion on freeways, the state of California is proposing to expand ferry operations on San Francisco (SF) Bay by phasing in up to 100 ferries in addition to the 14 currently operating, extending the hours of operation of the ferries, increasing the number of crossings, and employing some high-speed vessels. The state of California has directed the San Francisco Bay Area Water Transit Authority to determine whether the “safe” operation of ferries in San Francisco Bay 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. Of these alternatives, 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.

Merrick et al. (2004b) used the Bayesian simulation of the SF Bay area to analyze the effect of ferry expansions on the number of situations occurring with accident potential. Figure 6 shows a comparison plot of Alternatives 1 and 2.

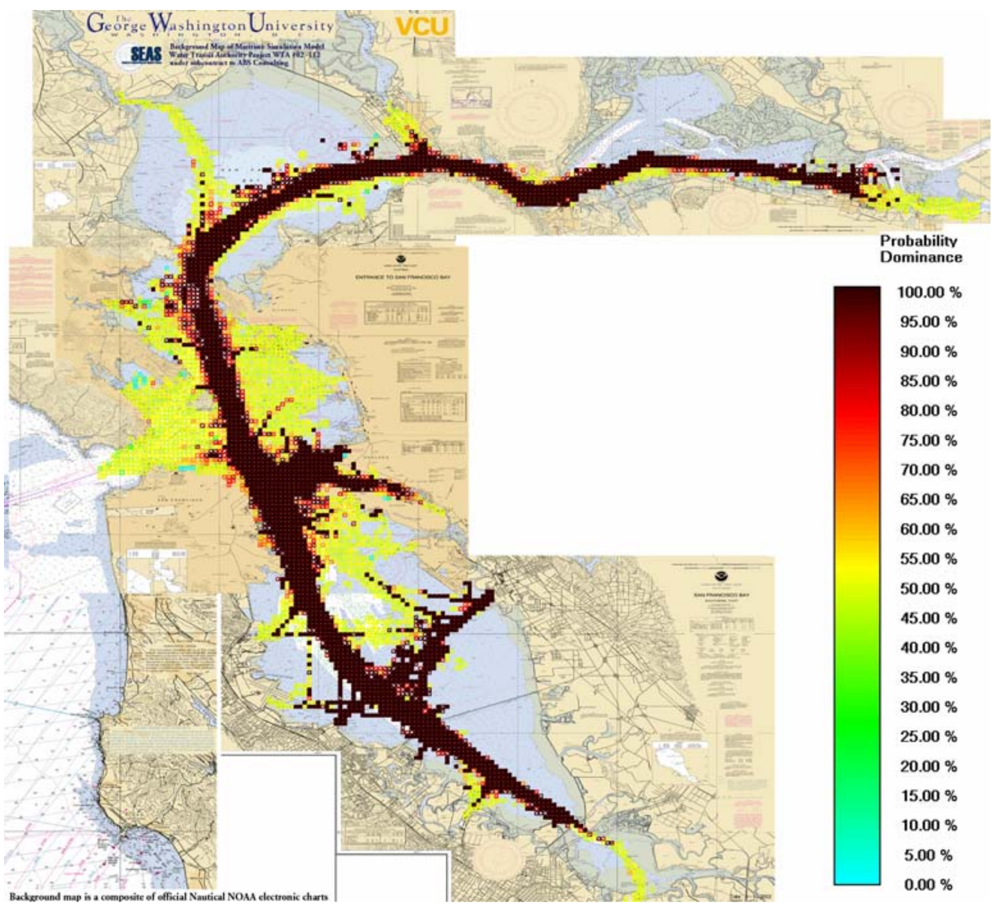

Figure 6. Alternative 2 compared to Alternative 1.

The geographic regions that are colored black show that almost surely more situations will occur in these areas if Alternative 1 is implemented than if Alternative 2 is implemented. The yellow regions show no dominance, with the predictive distributions of the number of situations for the two alternatives being approximately the same. A blue region would indicate the reverse, but does not occur as Alternative 1 will almost surely have as many or more situations occurring in all regions.

Figure 7A shows an aggregate comparison of the alternatives by the total expected yearly number of situations, in this case interactions with other vessels that could lead to a collision.

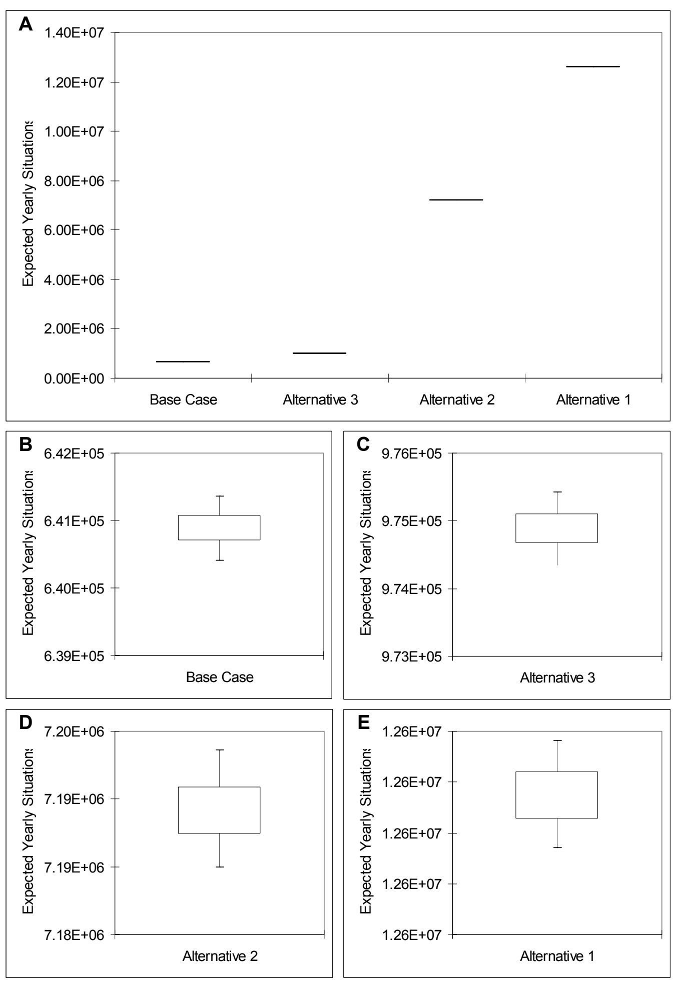

Figure 7. Expected Yearly Situations Comparison.

The lines in Figure 7A are actually box plots of showing the predictive distribution with the interquartile range as the box and the 5th and 95th percentiles of the distribution as the whiskers. However, as the remaining uncertainty in these estimates is low, they do not show up on this comparison scale and are repeated Figures 7B through 7D. It is evident from Figure 7 that there is an increase in the number of situations across the alternatives and this result is not subject to epistemic or aleatory uncertainty.

However, the results of Merrick et al. (2004a) count each such situation equally. Merrick et al. (2004b) analyzed the expert judgments from the WSF Risk Assessment considering both dependencies between the experts' responses and the remaining uncertainty in the estimates. Figure 8 shows the marginal posterior distributions of the β parameters for the factors listed in Table 1 and six interaction terms. The prior distributions used in Merrick et al. (2004b) were vague. For the model form in (3), a value of zero for these parameters implies that the corresponding factor does not affect the collision probability. A positive (negative) value indicates that an increase in the factor would increase (decrease) the collision probability.

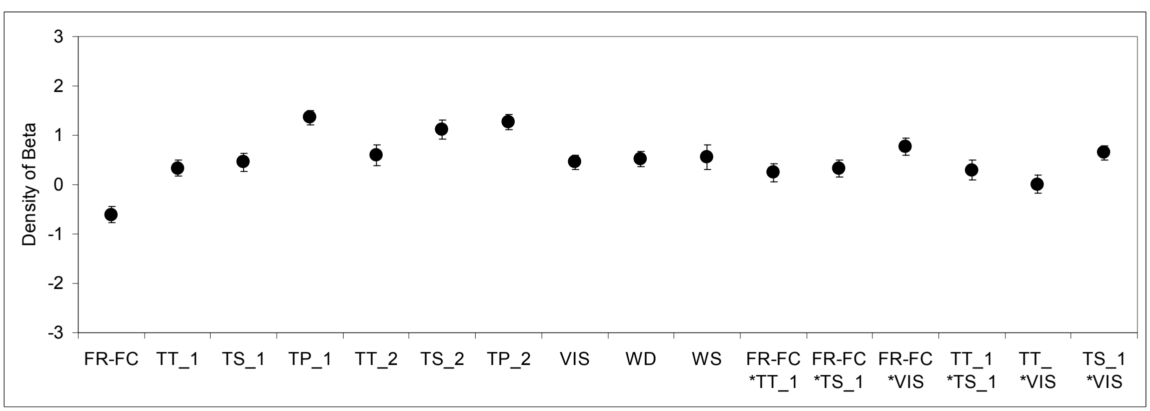

Figure 8. The marginal posterior distribution of the factor effect parameters.

As the factors describing the situations effect the probability of a collision given that a situation occurs, $P(Collision \mid Situation_j)$, the analysis from Merrick et al. (2004a) is useful, but not definitive. Instead we must examine the collision probability itself, $P(Collision)$. Figure 9A shows a similar pattern of increase for the expected yearly number of accidents as seen for the expected yearly situations. However, with the introduction of estimated accident probabilities based on expert judgments, there is significantly more uncertainty evident in these results and this uncertainty cannot be removed by simply running more simulations. The largest uncertainty remains about Alternatives 2 and 1. However, there are almost certainly a higher expected number of accidents in Alternative 1 than Alternative 2. There is not such certainty when comparing the Base Case to Alternative 3.

Whereas there was an almost certain ranking in terms of the expected yearly number of situations, this is not true for the expected yearly number of accidents. As the comparison is not clear on a scale that includes Alternatives 2 and 1, Figures 9B and 9C show the box plots for the Base Case and Alternative 3 respectively; the 90% credibility intervals for the two alternatives are (5.28×10^{-5}, 6.52×10^{-5}) for the Base Case and (6.19×10^{-5}, 7.77×10^{-5}). Thus these distributions do indeed overlap and the best we can say is that Alternative 3 stochastically dominates the Base Case in the sense that their cumulative distribution functions to do cross. This result seems questionable given the results in Figure 7 and to explain why this occurs we must consider the accident probabilities calculated for the occurring situations.

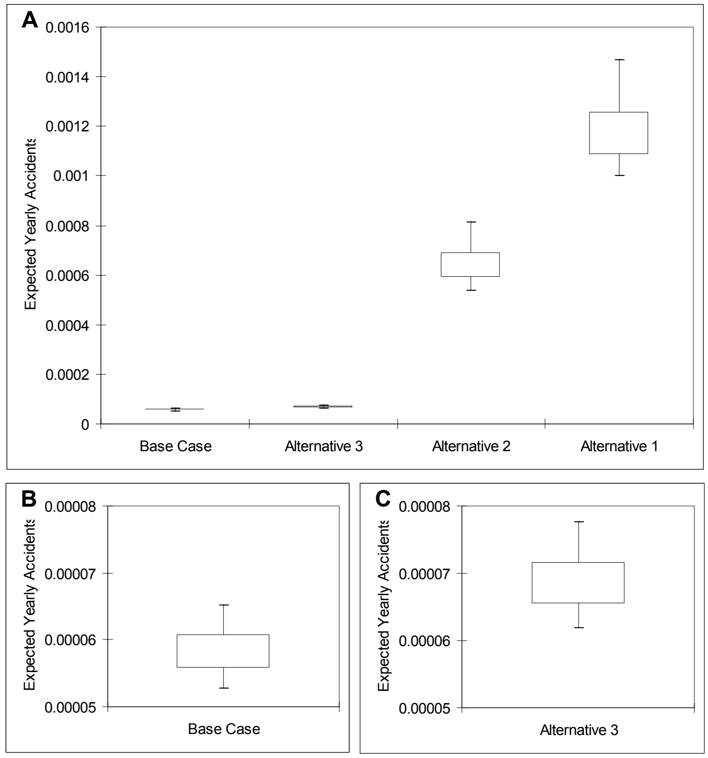

Figure 9. Expected Yearly Accidents Comparison.

It is evident from Table 1 that there will be many possible situations that can be counted in the simulation and from Figure 8 that these situations can have significantly different accident probabilities when they occur. Thus to compare the accident probabilities in the situations occurring in the different alternatives, we take the average accident probability across all situations that occurred in the simulation of each alternative. For each alternative, this involves taking the number of times that a given situation defined by the factors in Table 1 occurs and multiplying by the accident probability given that the situation occurs. We then add these results up for all possible situations and divide by the total number of situations that occurred.

Figure 10 shows the results of these calculations and the remaining uncertainty about the results for each alternative. Note that the result in Figure 9 can now be explained. Whereas the expected yearly number of interactions increases from the Base

Case to Alternative 3, the average probability of an accident actually decreases, thus causing the distributions of the multiple of these two quantities, the expected yearly number of accidents, to overlap. The average accident probabilities for Alternatives 2 and 1 are about the same as the Base Case, but there is more uncertainty.

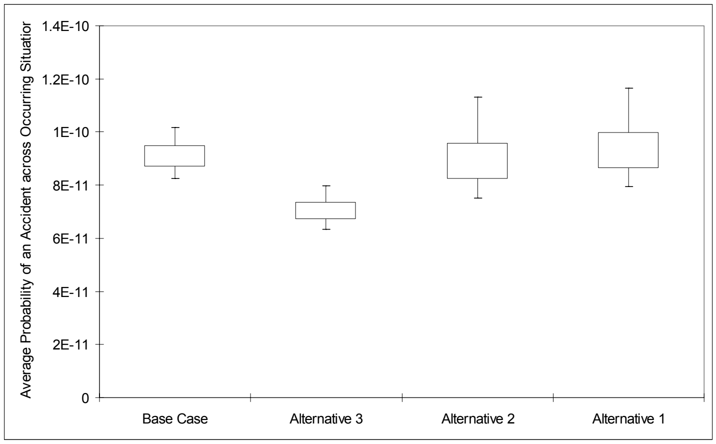

Figure 10. Average Probability of an Accident across Occurring Situations.

While Figure 10 does explain how the strange result in Figure 9 occurs, it does not explain why. What changes occur in Alternative 3 that reduce the average accident probability compared to the Base Case? In Alternative 3, only one additional route is added to the Base Case schedule. However, as discussed in Merrick et al. (2004b), there was a problem with the proposed schedules for the alternatives; they consisted of a start time, end time and time between ferries. For example, the Sausalito and Tiburon ferry schedule starts at 7 am and runs every 30 minutes until 10 pm during the week. At the weekend they run every 60 minutes. This is significantly more than in the Base Case, but this means that there are definite patterns to the transits that are not reflective of a more mature schedule where all vessels don’t start on the hour and run every 15, 30 or 60 minutes. With artificial schedules the ferries do not interact as much because of the timing of the transits. The question then is what effect this has that can make the average accident probabilities reduce from the Base Case to Alternative 3?

The analysis can be decomposed by any of the factors in Table 1. By this we mean that we can examine the expected yearly number of situations or accidents where the first interacting vessel was a navy vessel compared to a product tanker or where there was good visibility versus bad visibility. We can also calculate the average accident probabilities in situations of different types. To explain the decrease in average accident probability from the Base Case to Alternative 3, we broke down the analysis by each of the factors in Table 1 and discovered that, while other smaller effects were contributing to the result, the main effect was a change in the proximities of not the closest interacting vessel, but the second closest interacting vessel. This factor is included in the analysis as multiple nearby vessels are more confusing and thus lead to a higher chance of human error and thus collision.

Figure 11B shows the average accident probabilities for the Base Case for a second vessel within 1 mile of the ferry, the second vessel over 1 mile away (but still closer than 15 minutes away) and for no interacting vessel. It is evident that having a second vessel is considerably more risky than not and that the second vessel being closer in is more risky than further away. This is a logical result. While the exact values in Figure 11B are for the Base Case, they do not change significantly in the other alternatives, so these are omitted. Figure 11A shows the percentage of situations in each alternative that occur in these three classifications for the second interacting vessel. Note that in Alternative 3 there are proportionally fewer situations that have a second interacting vessel within 1 mile, the riskier situation, and proportionally more situations where the second interacting vessel is over 1 mile away, the less risky situation. This is a result of the artificial timings of the schedule tested for Alternative 3. As the average accident probability is averaged over all situations occurring in the simulated alternative, this will mean that the average accident probability will be lower for Alternative 3 than the Base Case.

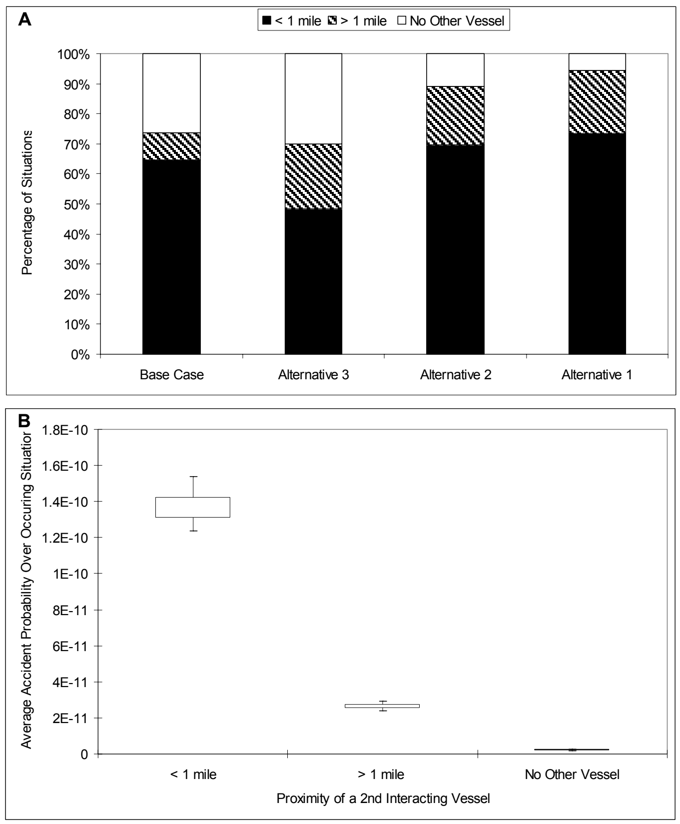

Figure 11. Explaining the Average Probabilities in Terms of the Proximities of 2nd Interacting Vessels.

What would these results mean in terms of the decision to build out the San Francisco Bay ferries if they could be considered more than an academic demonstration of the techniques? Firstly, while Alternative 3 does significantly increase the number of ferries and thus the expected yearly number of situations from the Base Case, there is a decrease in the risk of the situations that occur and thus the comparison in terms of expected yearly accidents is not conclusive. However, as this result appears to be caused by the artificial nature of the schedule tested, the actual schedule to be implemented should be tested in this manner before any decisive conclusions could be reached. We note that such caution would not be engendered by an analysis without uncertainty as the mean values would have implied a definitive ranking and led to the conclusions that Alternative 3 was less safe. Alternatives 2 and 1 do almost certainly increase the expected yearly number of accidents as ferries are added to the schedule. Merrick et al. (2003) concludes that with such a result, measures to reduce accident probability and control the occurrence of interactions should be considered before implementing such a major build out of the San Francisco Bay ferry system.

As a final remark, alternative comparison maps for the expected yearly number of collisions, like that in Figure 6 for the expected yearly number of situations, would be desirable. They would show where the accidents were most likely to occur and help design effective risk reduction measures. However, as the location would become an extra factor in the analysis with many possible values, this analysis is currently not computationally feasible, even with parallel implementation of the calculations.

5. Conclusions

We have developed an overarching Bayesian framework for addressing uncertainty when simulation of situations that have accident potential is combined with expert judgment to assess risk and uncertainty in a dynamic system, applying this framework to maritime transportation. The combination of the Bayesian simulation and a Bayesian multivariate regression of expert judgments is an original contribution to the field of uncertainty assessment in risk analysis. In the case study, the results in Merrick et al. (2003) were shown to be robust to the aleatory and epistemic uncertainty inherent in assessing risk in such a dynamic and data-scarce system, though surprising results did occur.

The broader impact of this work is primarily drawn from its applicability to areas other than maritime accident risk. **Port security** risk (intentional as opposed to accidental events) has now been recognized as an integral part of homeland security. Subsequent uncertainty assessment of security risk and propagation in security intervention effectiveness needs to be accounted for, since lack of data will be of even greater concern than for accident risk. Furthermore, despite our focus on maritime risk, the framework and methodologies developed will be applicable to other transportation modes as well, such as **aviation safety** including the ever-increasing problem of runway incursions at our national airports (FAA 2003).

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant Nos. SES 0213627 and SES 0213700. Any opinions, findings, 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.

References

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

Apostolakis, G. 1978. Probability and risk assessment: the subjectivist viewpoint and some suggestions. *Nuclear Safety* **19** 305-315.

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

Bernado, J., A. Smith. 1994. *Bayesian Theory*. Chichester, UK: Wiley.

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

Chick, S. 1997. Bayesian analysis for simulation input and output. *Proceedings of the Winter Simulation Conference*. S. Adradottir, K. Healy, D. Withers and B. Nelson eds., 253-260.

Chick, S. 2001. Input distribution selection for simulation experiments: accounting for input uncertainty. *Operations Research* **49**(5) 744-758.

Clemen, R. T., T. Reilly. 1999. Correlations and copulas for decision and risk analysis. *Management Science* **45**(2) 208-224.

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

Cooke, R. 1997. Uncertainty modeling: examples and issues. *Safety Science* **26**(1) 49-60.

Federal Aviation Administration. 2003. Operational Errors and Runway Incursions: Progress Made, but the Number of Incidents is Still High and Presents Serious Safety Risks. *Report Number: AV-2003-040*, Washington DC.

Ferson, S., L. Ginzburg. 1996. Different methods are needed to propagate ignorance and variability. *Reliability Engineering and System Safety* **54** 133-144.

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

Frey, H., D. Burmaster. 1999. Methods for characterizing variability and uncertainty: comparison of bootstrap simulation and likelihood-based approaches. *Risk Analysis* **19**(1) 109-130.

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

Hofer, E. 1996. When to separate uncertainties and when not to separate. *Reliability Engineering and System Safety* **54** 113-118.

Hora, S. 1996. Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management. *Reliability Engineering and System Safety* **54** 217-223.

Kaplan, S. 1997. The words of risk analysis. *Risk Analysis* **17**(4) 407-418.

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, Cambridge, MA, 96-19.

Law, A., W. Kelton 2001. *Simulation Modeling and Analysis*. 3rd edition. New York: McGraw Hill.

Merrick, J. R. W., 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 Engineering **3**(3) 128-142.

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

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

Merrick, J. R. W., J. R. van Dorp, V. Dinesh. 2004a. Assessing uncertainty in simulation based maritime risk assessment. Submitted to *Risk Analysis*.

Merrick, J. R. W., J. R. van Dorp, A. Singh. 2004a. Analysis of correlated expert judgments from extended pairwise comparisons. Submitted to *Decision Analysis*.

Moslesh, A., V. Bier, G. Apostolakis. 1988. Critique of the current practice for the use of Expert opinions in probabilistic risk assessment. *Reliability Engineering and System Safety* **20** 63-85.

National Research Council. 1998. *Review of the Prince William Sound, Alaska, Risk Assessment Study.* Committee on Risk Assessment and Management of Marine Systems, National Academy Press, Washington D.C.

National Research Council. 2000. Risk Management in the Marine Transportation System. *Conference Proceedings 22, National Academy Press, Washington, DC.*

Parry, G. 1996. The characterization of uncertainty in Probabilistic Risk Assessments of complex systems. *Reliability Engineering and System Safety* **54** 119-126.

Pate-Cornell, M. E. 1990. Organizational Aspects of Engineering System Safety: The Case of Offshore Platforms. *Science* **250**(4985) 1210-1217.

Paté-Cornell, M. E. 1996. Uncertainties in risk analysis: six levels of treatment. *Reliability Engineering and 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 Technology* **4** 230-238.

Press, S. J. 1982. *Applied Multivariate Analysis Using Bayesian and Frequentist Methods and Inference*. 2nd Edition. Robert E. Krieger Publishing Company, Malabar, Florida.

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

Saaty, T. 1977. A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology* **15**(3) 234-281.

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

Stiehl, G. L. 1977. Prospects for Shipping Liquefied Natural Gas. *Marine Technology* **14**(4) 351-378.

Szwed, P., J. Rene van Dorp, J. R. W. Merrick, T. A. Mazzuchi, A. Singh. A Bayesian paired comparison approach for relative accident probability assessment with covariate information. Accepted by *European Journal of Operations Research*.

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

van Dorp, J. R., J. R. W. Merrick, J. Harrald, T. Mazzuchi, M. Grabowski. 2001. A Risk Management Procedure for the Washington State Ferries. *Risk Analysis* **21**(1) 127-142.

Vose, D. 2003. *Risk Analysis: a Quantitative Guide*. 3rd edition. New York: Wiley.

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

Winkler, R. L. 1981. Combining probability distributions from dependent information sources. *Management Science* **27** 479-488.

Winkler, R. 1996. Uncertainty in probabilistic risk assessment. *Reliability Engineering and System Safety* **54** 127-132.