

Speaking the Truth in Maritime Risk Assessment

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 in the industry. 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 assessing this uncertainty is the different techniques that need to be used to model risk in this dynamic and data-scarce application area. In previous articles, we have developed the two pieces of methodology necessary to assess uncertainty in maritime risk assessment, 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. In this paper, we combine the methods to perform a full-scale assessment of risk and uncertainty for two case studies. The first is an assessment of the effects of proposed ferry service expansions in San Francisco Bay. The second is an assessment of risk for the Washington State Ferries, the largest ferry system in the United States.

Keywords: Uncertainty Analysis; Risk Analysis; Maritime Transportation.

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 capsizing 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. Their results have been used in major investment decisions and have played a significant role in the management of maritime transportation in the US.

In a maritime port system, traffic patterns change over time in a complex manner. The dynamic nature of these traffic patterns and, indeed, other situational variables, such as wind, visibility, and ice condition, mean that risk levels change over time. This requires the use of simulation to accurately model the impact of changes that affect the traffic patterns, such as introducing new traffic rules and increases or decreases in the volume of traffic in a given port. Furthermore, accident data that is relevant to a given port is often scarce, necessitating the use of expert judgment to estimate accident probabilities.

Figure 1 shows the results from an analysis of proposed ferry service expansions in San Francisco Bay.

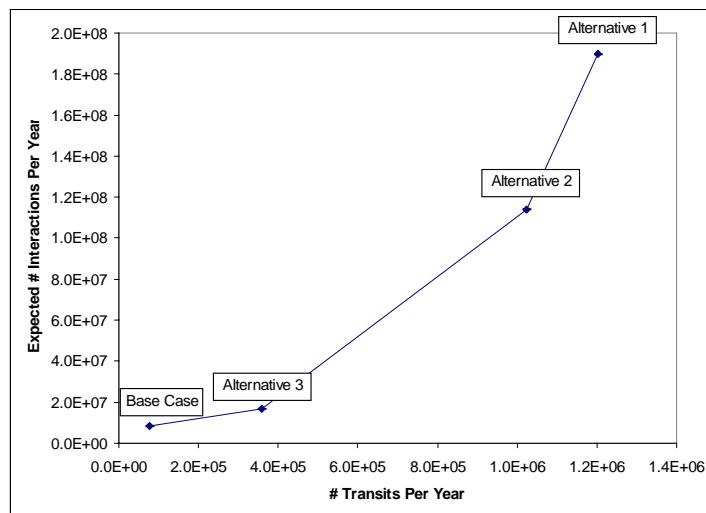


Figure 1. An assessment of alternative expansion scenarios for ferries in San Francisco Bay.

The estimates show the frequency of situations that could lead to a collision 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. As another example, Figure 2 shows the risk intervention effectiveness estimates from the WSF Risk Assessment (van Dorp et al., 2001). The figure shows the total percentage reduction in collision probability for the WSF system for various risk management alternatives.

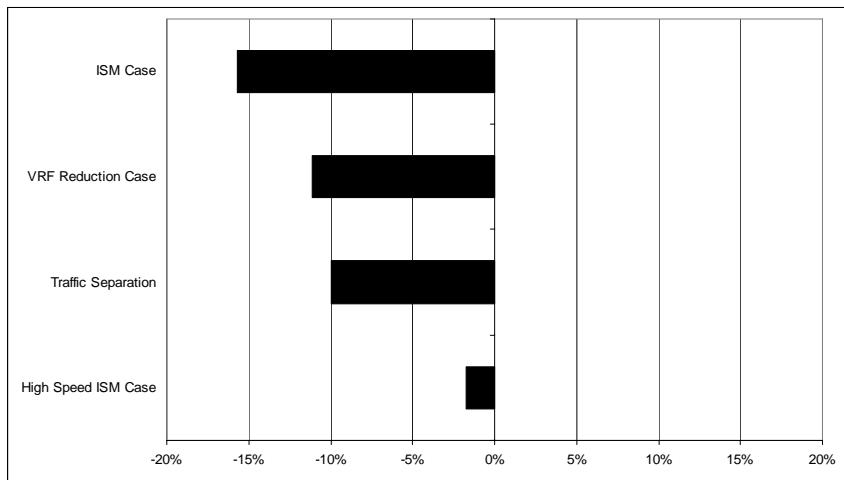


Figure 2. An assessment of risk intervention effectiveness for proposed safety improvements for the Washington State Ferries.

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).

The two pieces of methodology necessary to perform an uncertainty analysis for our maritime risk approach were developed in Merrick et al. (2005a) and Merrick et al. (2005b). In this article, we provide an overall framework for this combination. We use this framework to examine the uncertainty for two of our previous risk studies. We perform a complete uncertainty analysis of each result from the San Francisco Bay study (Figure 1) to demonstrate the complete approach and then examine the uncertainty in the assessment of risk intervention effectiveness for the WSF Risk Assessment (Figure 2) to show the use of this approach in the decision making process. While the framework is generic enough to be applied to other maritime risk work, and even other forms of transportation, some parts of the model are necessarily dependent on the availability of data in each study.

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 and the WSF Risk Assessment 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. Another form of uncertainty concerning the model itself is discussed in the literature, the uncertainty regarding which of a list of alternative models most closely represents the real system (Nilsen and Aven, 2003). However, in this article we take the first step, adding consideration of epistemic uncertainty. The computational complexity of the analysis limited our scope at this point in time.

While epistemic uncertainty can be addressed through frequentist statistical techniques such as bootstrap or likelihood based methods (Frey and Burmester, 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 Maritime Risk Assessment

We will use the example of a ferry system to demonstrate the application of uncertainty analysis for a maritime risk assessment (van Dorp et al. 2001; Merrick et al. 2003). We consider one type of accident, collisions between a ferry and another vessel. Collisions are caused by a triggering incident, specifically propulsion failure, steering failure, navigational aid failure, human error or error by a nearby vessel (which we do not disaggregate). While multiple errors and/or failures could be involved in the collision, the triggering incident is the error or failure that directly causes the collision itself if corrective actions do not succeed. Triggering incidents and collisions occur within a situation defined by factors that affect their 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 in a given time period, we sum over the possible situations giving

$$P(\text{Collision}) = \sum_{j=1}^k \sum_{i=1}^l P(\text{Collision} | \text{Incident}_i, \text{Situation}_j) P(\text{Incident}_i | \text{Situation}_j) P(\text{Situation}_j) \quad (1)$$

where Situation_j denotes the possible combinations of values of the factors for $j = 1, \dots, k$ and k is the total number of possible combinations, and Incident_i denotes the i -th of l possible triggering incidents (5 in our example here). The expected yearly frequency of collisions can then be calculated by multiplying the probability in (1) by the number of time periods in a year.

The accident probability model consists of three parts:

- $P(\text{Situation}_j)$: the probability that particular combination of values of the factors occurs in the system;
- $P(\text{Incident}_i | \text{Situation}_j)$: the probability that a particular triggering incident occurs in the given situation; and
- $P(\text{Collision} | \text{Incident}_i, \text{Situation}_j)$: the probability that an accident occurs in the defined situation once the triggering incident has occurred.

To perform an assessment of the risk of an accident using this model, each term in the probability model needs to be estimated.

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 1 st interacting vessel	TT_1	13
X_3	Scenario of 1 st interacting vessel	TS_1	4
X_4	Proximity of 1 st interacting vessel	TP_1	Binary
X_5	Type of 2 nd interacting vessel	TT_2	5
X_6	Scenario of 2 nd interacting vessel	TS_2	4
X_7	Proximity of 2 nd interacting vessel	TP_2	Binary
X_8	Visibility	VIS	Binary
X_9	Wind direction	WD	Binary
X_{10}	Wind speed	WS	Continuous

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 in each simulated time period where there is the potential that a ferry could be involved in an accident, 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 situations with collision potential). A multi-year simulation is run and for each time period in the simulation the situations that occur are counted. So the probability of situations with particular values of the factors in a given time period, 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, WSF and SF Bay studies meaning that only point estimates of $P(Situation_j)$ were obtained.

The preferred method for estimating $P(Collision | Incident_i, Situation_j)$ 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 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 $P(Collision | Incident_i, Situation_j)$. An example of the form of the questions drawn from the WSF risk assessment project is shown in Figure 3.

Situation 1	Attribute	Situation 2
Issaquah	Ferry Class	-
SEA-BRE(A)	Ferry Route	-
Navy	1st Interacting Vessel	Product Tanker
Crossing	Traffic Scenario 1 st Vessel	-
< 1 mile	Traffic Proximity 1 st Vessel	-
No Vessel	2nd Interacting Vessel	-
No Vessel	Traffic Scenario 2 nd Vessel	-
No Vessel	Traffic Proximity 2 nd 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

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. The responses to the questions are in terms of relative probabilities of the event in the two situations. 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

$$P(\text{Collision} | \text{Incident}_i, \text{Situation}_j) = p_0 \exp\{\text{Situation}_j^T \beta\}, \quad (2)$$

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 $\exp((\text{Situation}_L - \text{Situation}_R)^T \beta)$, where Situation_L and Situation_R are the vectors of factors for the situations on the left and right sides of the question respectively. If we equate the natural logarithm of the experts' responses and the corresponding model terms, the analysis can be performed using linear regression techniques. Thus again, only point estimates were obtained. One should also note that it is possible to calculate probabilities above one using this method (but not below zero), requiring truncation. However, while the form of the model would allow incoherent values, they are extremely unlikely as we are dealing with low probability events in this context.

The determination of $P(\text{Incident}_i | \text{Situation}_j)$ depends on the availability of data. While mechanical failure data is readily available, the availability of human error data is variable from port to port, meaning that there is no standard approach thus far for

estimating $P(\text{Incident}_i | \text{Situation}_j)$ in our studies. We will discuss this further in Section 3.4.

Figure 4 shows a simple influence diagram of the probability model in (1) along with the data used in estimating each conditional probability below each node. 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 $P(\text{Situation}_j)$, uncertainties in the experts' assessments of the conditional probabilities $P(\text{Collision} | \text{Incident}_i, \text{Situation}_j)$ and uncertainties in the estimation of mechanical failure and human error rates from available data and expert judgments to obtain $P(\text{Incident}_i | \text{Situation}_j)$. We must then propagate these uncertainties through the framework expressed by (1).

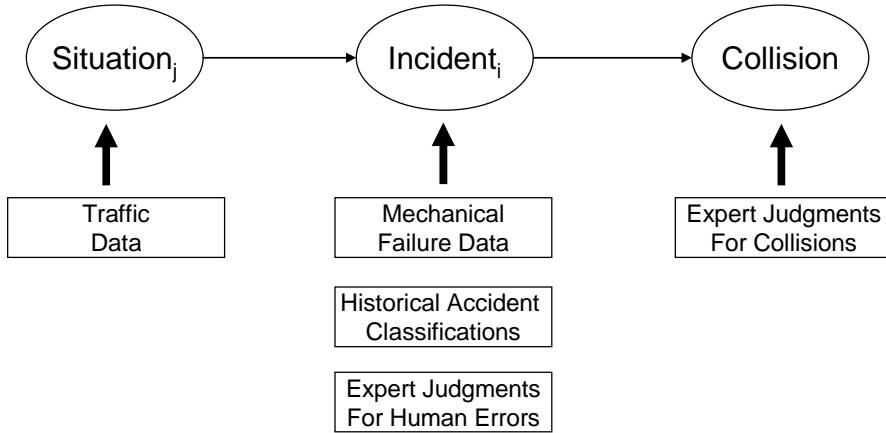


Figure 4. A simplified influence diagram of the accident probability model along with the data used in estimation.

3.2 Modeling Uncertainty in the Simulation

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. We should note, however, that our use of the term Bayesian simulation is akin to Chick (1997), rather than that used in Bier and Andradottir (2000) where the output data is used to update the decision maker's beliefs about both the input and output distributions simultaneously. Our approach follows Chick in updating the input distributions based on traffic data, passing the inputs through the simulation and learning about the outputs alone from the output data.

In Bayesian simulation, 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). Hence Bayesian renewal process models of traffic arrivals were created in Merrick et al. (2005a) for all vessel arrivals processes in the San Francisco Bay simulation. The prior distributions on the parameters for these renewal processes were updated with available traffic data from a Coast Guard Vessel Traffic Center. This input uncertainty is then propagated through the multiplication replications of the simulation using the following algorithm (Chick 1997, 2001):

For $r = 1, \dots, n$ replications:

1. Sample values from the posterior distributions for the renewal process parameters to be used in the r -th replication.,
2. For the r -th replication:
 - a. Sample random variates for the inter-arrival times given the parameter values drawn for each renewal process in Step 1;
 - b. Generate the simulation output that is determined by these random variates.

End loop

So aleatory uncertainty is represented within each replication, while epistemic uncertainty is represented across multiple replications.

The presence of input uncertainty means that there will be uncertainty in the outputs as well. This will include both the aleatory and epistemic uncertainty from the input uncertainty propagated through the simulation. 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 ($r = 1, \dots, s$ for s replications) and the j -th combination of values of the factors ($j = 1, \dots, k$). Following the Bayesian approach, we hypothesize a probability model for the random output data and specify our prior beliefs about the parameters of this output model. Chick (1997) notes that this can be thought of as a Bayesian version of meta-modeling (Law and Kelton 2001).

As our output data is in the form of a count, Merrick et al. (2005a) model the number of situations for the j -th combination of values of the factors as a Poisson distribution with rate μ_j , with a conjugate gamma distributed prior on μ_j with shape α_j and scale γ_j . Here α_j can be considered the total number of such situations observed by the decision maker in the real system and γ_j the number of periods the length of a replication for which the decision maker observed the system. The posterior distribution of the expected frequency of situations for the j -th combination of values of the factors is given by

$$(\mu_j | n_{1,j}, \dots, n_{s,j}) \sim \text{gamma}\left(\alpha_j + \sum_{i=1}^s n_{i,j}, \gamma_j + s\right) \quad (3)$$

The predictive distribution of $P(Situation_j)$ is then a Poisson-gamma distribution in the sense of Bernardo and Smith (2000). Note that the aleatory uncertainty here can be reduced by running longer simulations, the epistemic uncertainty cannot; this would require additional traffic data. For additional details of this analysis we refer the reader to Merrick et al. (2005a)

3.3 Modeling Uncertainty in the Collision Probabilities

Szwed et al. (2004) develop a conjugate Bayesian analysis for the questions depicted in Figure 3 and the model form shown in (2) assuming that the experts' responses are not based on overlapping information (Clemen, 1987). Merrick et al. (2005b) 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 $Situation_j = (x_{j,1}, \dots, x_{j,q})$ 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

$$\mathbf{Y} = \mathbf{X}\beta\mathbf{1}^T + \mathbf{U}, \quad (4)$$

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, $\beta = (\beta_1, \dots, \beta_q)^T$ is the vector of regression parameters from (3) and $\mathbf{1} = (1_1, \dots, 1_q)^T$ 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 Σ . Merrick et al.

(2005b) show that the likelihood for this model can be written as a vector normal distribution with mean vector and covariance matrix given by

$$\mu_*^\beta = \frac{\hat{\mathbf{B}}\Sigma^{-1}\underline{1}}{\underline{1}^T\Sigma^{-1}\underline{1}} \quad (5)$$

and

$$\Sigma_*^\beta = \frac{(X^T X)^{-1}}{\underline{1}^T \Sigma^{-1} \underline{1}}. \quad (6)$$

A natural conjugate analysis is made possible by the following distributional assumptions,

$$(\Sigma) \sim \text{Inv-Wishart}(\mathbf{G}, m), \quad (7)$$

which defines an inverse Wishart distribution of dimension p with parameter matrix \mathbf{G} and m degrees of freedom, and

$$(\beta | \Sigma) \sim \text{MVNormal}\left(\phi, \frac{\mathbf{A}}{\underline{1}^T \Sigma^{-1} \underline{1}}\right). \quad (8)$$

ϕ , \mathbf{A} , \mathbf{G} and m are prior hyperparameters determined by the decision maker. ϕ represents the decision maker's prior mean estimate of β , while the matrix \mathbf{A} defines the prior covariance between the elements of β . The matrix \mathbf{G} represents the decision maker's prior estimate of the covariance between the experts and m represents the degree of certainty in this estimate, with higher values representing more certainty.

Given the experts' responses to the questionnaires, the posterior distributions are

$$(\Sigma | \mathbf{Y}, \mathbf{X}) \sim \text{Inv-Wishart}(\mathbf{G} + \mathbf{V}, m + N) \quad (9)$$

and

$$(\beta | \mathbf{Y}, \mathbf{X}, \Sigma) \sim MVNormal \left(\left(\mathbf{A}^{-1} + \mathbf{X}^T \mathbf{X} \right)^{-1} \left(\mathbf{X}^T \mathbf{X} \frac{\hat{\mathbf{B}} \Sigma^{-1} \underline{1}}{\underline{1}^T \Sigma^{-1} \underline{1}} + \mathbf{A}^{-1} \phi \right), \frac{\left(\mathbf{A}^{-1} + \mathbf{X}^T \mathbf{X} \right)^{-1}}{\underline{1}^T \Sigma^{-1} \underline{1}} \right). \quad (10)$$

The natural logarithm of the ratio to be predicted for two scenarios with difference vector

x^* conditioned on Σ will be a multivariate normal distribution defined by

$$(x^{*T} \beta | \mathbf{Y}, \mathbf{X}, \Sigma) \sim MVNormal \left(x^{*T} \left(\mathbf{A}^{-1} + \mathbf{X}^T \mathbf{X} \right)^{-1} \left(\mathbf{X}^T \mathbf{X} \frac{\hat{\mathbf{B}} \Sigma^{-1} \underline{1}}{\underline{1}^T \Sigma^{-1} \underline{1}} + \mathbf{A}^{-1} \phi \right), x^{*T} \frac{\left(\mathbf{A}^{-1} + \mathbf{X}^T \mathbf{X} \right)^{-1}}{\underline{1}^T \Sigma^{-1} \underline{1}} x^* \right) \quad (11)$$

We may then integrate out Σ using (9) to obtain a student-t distribution. For additional details of this analysis, we refer the reader to Merrick et al. (2005b).

3.4 Modeling Uncertainty in the Triggering Incident Probabilities

There are three pieces necessary to model $P(Incident_i | Situation_j)$ for the different types of triggering incidents. The WSF made failure data available for propulsion failures, steering failures and navigational aid failures. The times between mechanical failures were modeled as exponential distributions with a different parameter λ^f for each type of ferry ($f = 1, \dots, 10$). For the exponential distribution with rate parameter λ^f , the gamma distribution is a natural conjugate prior for λ^f . That is, if λ^f is assumed a priori to be drawn from a gamma distribution with shape parameter a^f and scale parameter b^f , then after updating with the time between failure data (denoted $t_1^f, \dots, t_{m^f}^f$), λ^f will be a

gamma distribution with shape parameter $a^f + \sum_1^{m^f} t_i^f$ and scale parameter $b^f + m^f$.

Hence a simple exponential-gamma model could be used to determine a posterior predictive distribution for the required probabilities given the failure data.

Modeling human error was more difficult due to the lack of human error data (Harrald et al. 1998; Grabowski et al. 2000). The human error rates were differentiated across the various routes and types of ferries. Ferry captains were asked to make pairwise comparisons of the various combinations of routes and types of ferries. These expert judgments were then analyzed using the method discussed in Section 3.3. However, this analysis did not allow estimation the overall number of human errors. In the PWS Risk Assessment, the frequency of human errors was estimated by taking the frequency of mechanical failures and multiplying by four, the commonly assumed 80-20 rule (Harrald et al. 1998). In the WSF Risk Assessment, an analysis of historical ferry accidents revealed that of the 51 triggering incidents leading to accidents in the preceding ten years, 35 were human errors and 16 were mechanical failures. Consequently there were 2.19 times as many triggering incidents that were human errors as opposed to mechanical failures, if we require just a point estimate.

To model the uncertainty in this estimate, we assumed that the triggering incidents were independent with a fixed probability p of being a human error instead of a mechanical failure. This implies that the likelihood for the number of human errors out of a fixed number of triggering incidents is a binomial distribution. We assumed a beta prior distribution on p with parameters a and b . Given the number of triggering incidents, denoted n , and the number of them that were human errors, denoted m , the posterior distribution of p is also a beta distribution with parameters $a+m$ and $b+n-m$. a and b were assumed to be 0.00001 to give a high variance prior, implying that the posterior beta distribution on p had parameters 35.00001 and 16.00001, with mean 0.686. This corresponds to a mean ratio of human errors to mechanical failures as

before, namely 2.19 , but includes the remaining epistemic uncertainty about this estimate.

3.5 Propagating Uncertainty through the Risk Framework

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, $P(Situation_j)$,

$P(Incident_i | Situation_j)$ and $P(Collision | Incident_i, Situation_j)$, and then propagate the

uncertainty expressed in these distributions through the calculations in (1). Given the

development in Sections 3.2, 3.3 and 3.4, the predictive distribution of $P(Collision)$ in

(1) cannot be obtained in closed form as the multiplication of the various distributions

does not result in a known distribution. 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). Monte

Carlo simulation is the most commonly used tool for propagating uncertainty through a

risk analysis model (Vose 2003). 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.

In this manner, samples of the posterior distribution of $P(Collision)$ are obtained and

descriptive statistics of the distribution can be estimated. The exact algorithm can be

written as follows:

For r = 1, ..., n replications:

1. Sample from the posterior distributions of the parameters used in calculating $P(Incident_i | Situation_j)$ for each type of incident
 - a. For each type of mechanical failure, sample from λ^f , the rate of each mechanical failure for each type of ferry ($f = 1, \dots, 10$).
 - b. For each type of ferry and route, sample from the posterior distribution of the parameters for the relative probability of human errors discussed in Section 3.4.
 - c. Sample from the probability p of an incident being a human error instead of a mechanical failure discussed in Section 3.4.
2. Sample from the posterior distribution of the parameters used in calculating $P(Collision | Incident_i, Situation_j)$ for each type of incident. For $i = 1, \dots, l$, sample from the posterior distribution of β^i from equation (10) where the responses are drawn from the questionnaire for incident type i .
3. Calculate the overall $P(Collision)$ for each potential situation.

For $j = 1, \dots, k$:

- a. Sample from the posterior distribution of μ_j in (3).
- b. Sample from $N_j | \mu_j$, a Poisson distribution.
- c. Use the samples from Steps 1.b and 1.c to calculate the probability of human errors for the ferry class and route in situation j .
- d. Use the samples from Step 2 to calculate $P(Collision | Incident_i, Situation_j)$ for each incident i in situation j using equation (2).
- e. Use the samples from Steps 3.a to 3.d and 1.a to calculate $P(Collision)$ using equation (1).

Loop

End loop

Such calculations do require significant computational effort. 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. To spread the workload, parallel simulations were implemented on a 10 processor system, decreasing the actual run time by almost a factor of 10.

4. Case Study 1: Expanding San Francisco Bay's Ferry Service

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 asked the authors 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.

A classical simulation was developed by the authors for the original study (Merrick et al. 2003). As part of our uncertainty modeling, Merrick et al. (2005a) extended the SF Bay simulation model using Bayesian input and output modeling techniques discussed in Section 3.2. 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. Non-ferry traffic was modeled using the Bayesian methodology in Section 3.2 using historical traffic data.

We start by examining the number of situations that could occur under each alternative, a result from Merrick et al. (2005a). Figure 5A shows an aggregate comparison of the alternatives by the total expected yearly number of situations, in this case when a ferry is close enough to other vessels that the situation could lead to a collision.

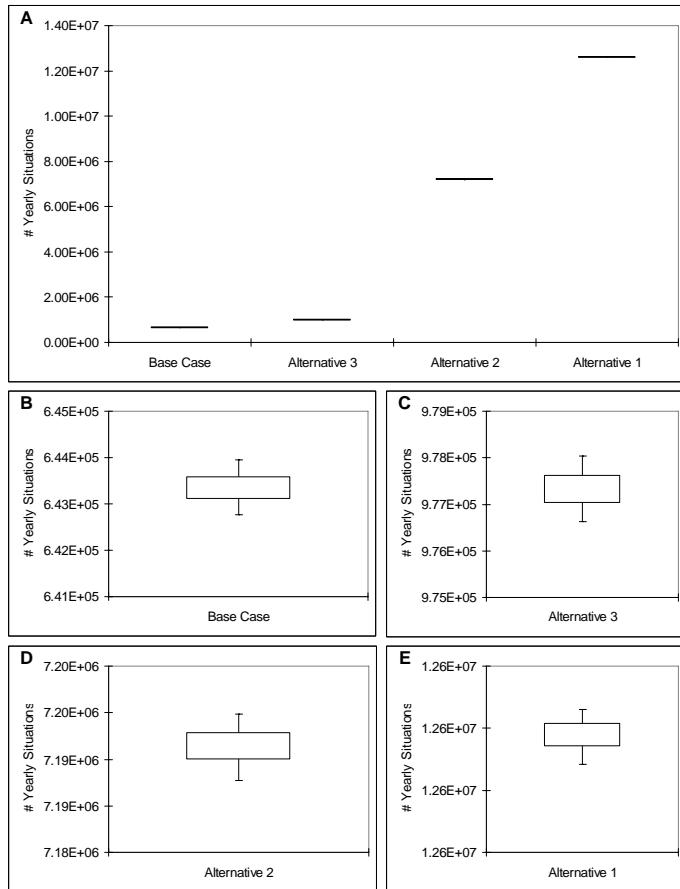


Figure 5. Expected Yearly Situations Comparison.

The lines in Figure 5A 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 5B through 5E. It is evident from Figure 5 that there is an increase in the number of situations across the alternatives and that the amount of uncertainty is small relative to the size of the differences between the alternatives.

However, the results of Merrick et al. (2005a) count each such situation equally. Merrick et al. (2005b) analyzed the expert judgments from the WSF Risk Assessment considering both dependencies between the experts' responses and the remaining uncertainty in the estimates. Figure 6 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. (2005b) were vague. For the model form in (2), 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.

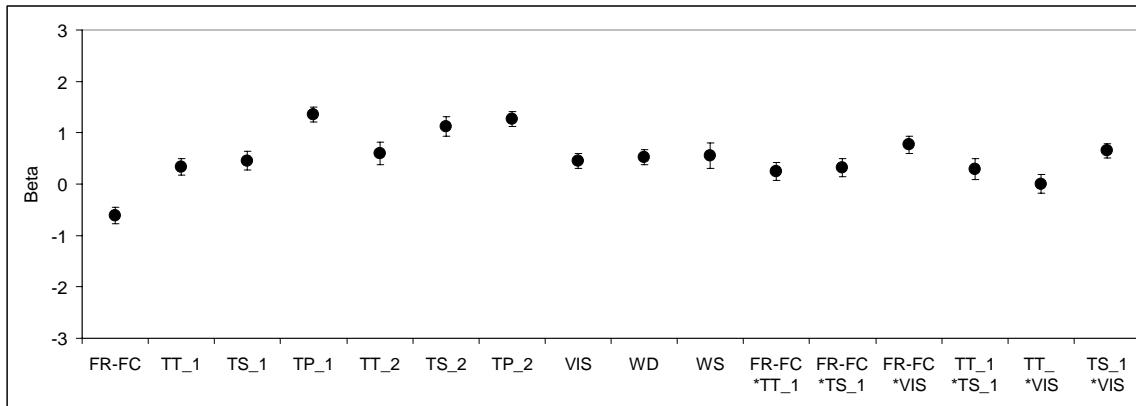


Figure 6. The marginal posterior distribution of the factor effect parameters.

As the factors describing the situations effect the probability of a collision given that a situation occurs, $P(\text{Collision} | \text{Situation}_j)$, the analysis from Merrick et al. (2005a) is useful, but not definitive. Instead we must examine the collision probability itself, $P(\text{Collision})$. As no expert judgments were elicited specifically for the SF Bay study due to a limited scope, we apply the probability assessments in Figure 6 from the WSF Risk Assessment. This part of the analysis should be considered illustrative, not definitive.

Figure 7A shows a similar pattern of increase for the expected yearly number of collisions as seen for the expected yearly situations. However, with the introduction of estimated collision 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.

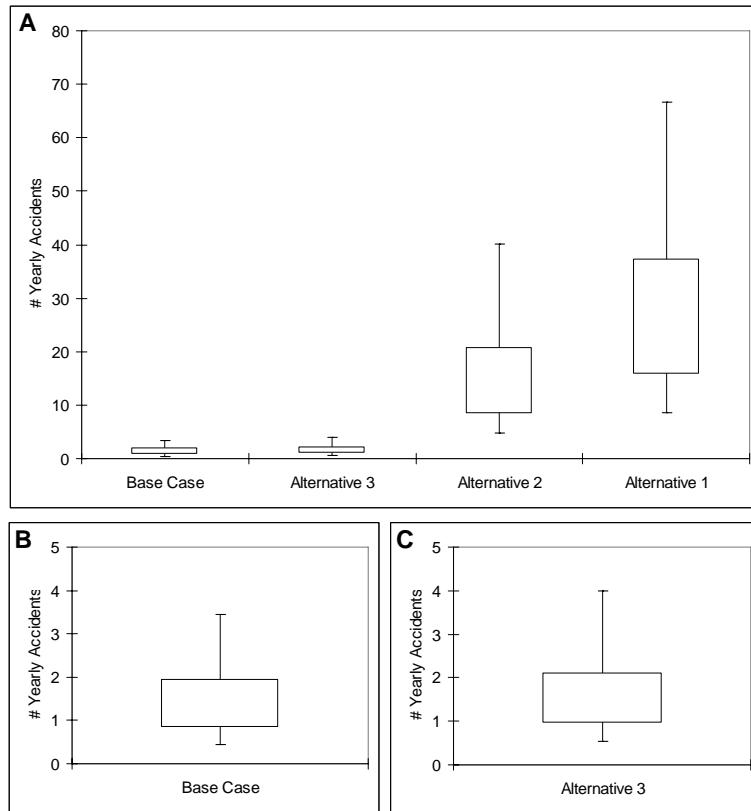


Figure 7. Expected Yearly Collisions Comparison.

The largest uncertainty remains about Alternatives 2 and 1. However, there are almost certainly a higher expected number of collisions 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 collisions. As the comparison is not clear on a scale that includes Alternatives 2 and 1, Figures 7B and 7C show the box plots for the Base Case and Alternative 3 respectively; the 90% credibility intervals for the two alternatives are (0.45,3.44) for the Base Case and (0.54,3.99) for Alternative 3. 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 do not cross. This result seems questionable given the results in Figure 5 and to explain why this occurs we must consider the collision probabilities calculated for the occurring situations.

It is evident from Table 1 that there will be many possible situations that can be counted in the simulation and from Figure 6 that these situations can have significantly different collision probabilities when they occur. To compare the collision probabilities in the situations occurring in the different alternatives, we take the average collision 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 collision 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 8 shows the results of these calculations and the remaining uncertainty about the results for each alternative. Note that the result in Figure 7 can now be explained. Whereas the expected yearly number of situations increases from the Base Case to Alternative 3, the average probability of a collision actually decreases, causing the distributions of the multiple of these two quantities, the expected yearly number of collisions, to overlap. The average collision probabilities for Alternatives 2 and 1 are about the same as the Base Case.

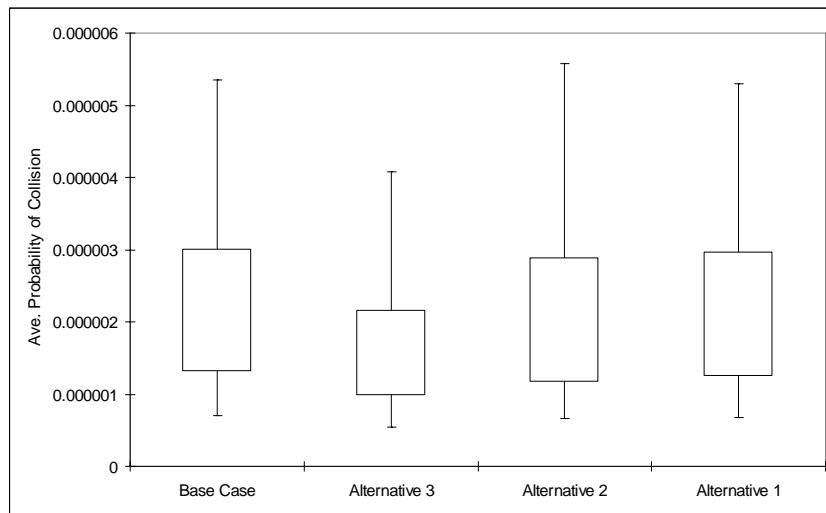


Figure 8. Average Probability of an Accident across Occurring Situations.

Further analysis showed that the reduction in collision probability in Figure 8 is due to the timing of the ferry schedules for Alternative 3. These schedules reduced the proportion of the time that there were two vessels in close proximity to a ferry, creating a less complex and therefore less risky situation.

What do the results in Figure 7 mean in terms of the decision to build out the San Francisco Bay ferries? While we should be careful in overstating these conclusions as the accident probabilities were based on data and expert judgment drawn from the WSF Risk Assessment, the implications are, however, interesting. Firstly, while Alternative 3 does

significantly increase the number of ferries, and consequently the expected yearly number of situations from the Base Case, there is a decrease in the risk of the situations that occur and the comparison in terms of expected yearly collisions is not conclusive. However, as this result appears to be caused by the 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 point estimates 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 collisions 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 situations should be considered before implementing such a major build out of the San Francisco Bay ferry system.

5. Case Study 2: The WSF Risk Assessment

The Washington State Ferries is the largest ferry system in the United States, operating 27 vessels at the time of the WSF Risk Assessment, including 4 passenger only ferries, to twenty terminals on ten routes. In 1998, total ridership for the ferries serving the central Puget Sound region was approximately 26.2 million persons, more passengers than Amtrak handles in a year. In 1998, the Washington State Transportation Commission, at the request of the State Legislature, established an independent Blue Ribbon Panel to assess the adequacy of provisions for passenger and crew safety aboard the Washington State Ferries, following a series of articles in the local newspapers about the adequacy of lifeboats aboard the ferries engendered by the release of the movie Titanic. As a result, the Blue Ribbon Panel engaged a consultant team including the authors to develop

recommendations for prioritized risk reduction measures which, once implemented, can improve the level of safety in the Washington State Ferry system. Figure 2 shows a summary of the results from the WSF Risk Assessment. Additional discussion is given in van Dorp et al. (2001) and Grabowski et al. (2000).

To assess the uncertainty in the risk reduction estimates given in Figure 2, we overlaid the simulation output meta-model discussed in Section 3.2 on a simulation of the WSF system. The predictive distributions obtained for $P(\text{Situation}_j)$ were then combined with the estimates of $P(\text{Collision} \mid \text{Incident}_i, \text{Situation}_j)$ obtained in Section 3.3 and $P(\text{Incident}_i \mid \text{Situation}_j)$ obtained in Section 3.4, both using data and expert judgments from the WSF Risk Assessment. Hence all data used came from the WSF Risk Assessment.

The proposed risk reductions were taken from the WSF Risk Assessment and were not modified for our analysis here. The cases modeled were

1. **Base Case:** representing ferry operations during the base case year for the risk assessment, namely 1998.
2. **ISM Case:** representing implementation of the International Safety Management (ISM) code throughout the WSF fleet. Specifically human errors were reduced 30% (estimated effect, see van Dorp et al. 2001) and mechanical failures were reduced 3.7% (estimated effect, see van Dorp et al. 2001).
3. **High-Speed ISM Case:** representing implementation of the International Safety Management (ISM) code on high-speed ferries only. Human errors and mechanical failure reductions from the ISM Case applied only to high-speed ferries.

4. **Vessel Reliability Failure (VRF) Reduction Case:** representing improved maintenance practices and system redundancy. Specifically mechanical failures were reduced 50%.

5. **Traffic Separation Case:** increase required route separations for high-speed ferries to decrease interactions within one mile by 50%.

To accurately represent the uncertainty in the percentage risk reduction from the Base Case to each proposed risk reduction case, we must be careful in our Monte Carlo propagation of the various sources of uncertainty from Section 3.5. To estimate the uncertainty in the percentage risk reduction, we must perform Steps 1 and 2 of our propagation algorithm in Section 3.5, then perform Step 3 to estimate $P(\text{Collision})$ for each case using the same sampled values of the parameters of the model, and then calculate the percentage risk reduction from the base case to each proposed risk reduction case for this Monte Carlo iteration. This calculation is akin to common random numbers (Law and Kelton 2001) and means that the percentage risk reductions calculated will only be affected by the relevant uncertainty.

The right of Figure 9 shows the posterior distributions of the percentage risk reductions obtained for each proposed risk reduction case using this method; the left of Figure 9 shows equivalent results without uncertainty repeated from Figure 2, but in the same format for comparison. Figure 9 shows a low uncertainty about the two ISM cases and the VRF Reduction case. These results are only affected by the uncertainty in the proportion of triggering incidents that are human errors versus mechanical failures. The uncertainty about the reduction from the Traffic Separation case is larger as it includes uncertainty from the simulation output as well as uncertainty from the expert judgments

about the relative likelihood of collisions when the traffic separation is more or less than one mile. Overall, however, the case is still clear for ISM across the whole fleet, instead of just the high speed ferries, and the efficacy of both VRF reduction and traffic separation are demonstrated; the conclusions drawn in the original study are confirmed.

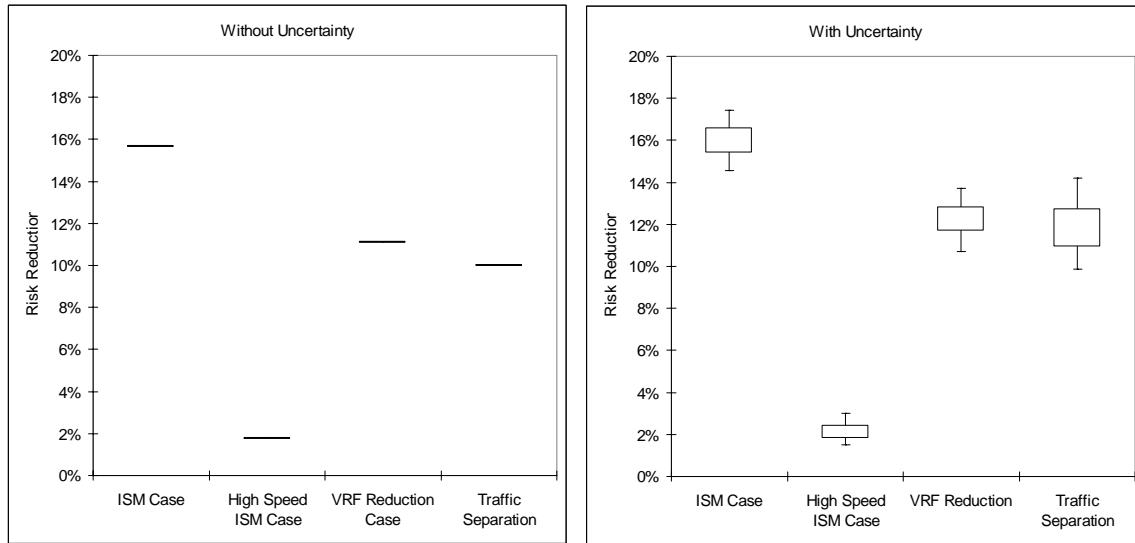


Figure 9. Percentage Risk Reductions from Proposed Interventions in the WSF

Ferries with and without uncertainty.

6. 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. In the case study, the results in Merrick et al. (2003) and van Dorp et al. (2001) 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, the Washington State Ferries, nor the San Francisco Bay Water Transit Authority.

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

Bradley, R., M. Terry. 1952. Rank analysis of incomplete block designs. *Biometrika* **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. 1987. Combining overlapping information. *Management Science* **33**(3) 373-380.

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.
- Grabowski, M., J. R. W. Merrick, J. Harrald, T. Mazzuchi and J. R. van Dorp. 2000. Risk Modeling in Distributed, Large Scale Systems. *IEEE Systems, Man & Cybernetics, Series A* **30**(6) 651-660.
- 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. 2005a. Assessing uncertainty in simulation based maritime risk assessment. *Risk Analysis* **25**(3) 731-743.
- Merrick, J. R. W., J. R. van Dorp, A. Singh. 2005b. Analysis of correlated expert judgments from extended pairwise comparisons. *Decision Analysis* **2**(1) 17-29.
- Mosleh, 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.*
- Nilsen, T. and T. Aven. 2003. Models and model uncertainty in the context of risk analysis. *Reliability Engineering and System Safety* **79** 309-317.
- 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.