Analysis of Correlated Expert Judgments from Extended Pairwise Comparisons

Jason R. W. Merrick — Virginia Commonwealth University

J. Rene van Dorp — The George Washington University

Amita Singh — The George Washington University

Abstract

We develop a Bayesian multivariate analysis of expert judgment elicited using an extended form of pair-wise comparisons. The method can be used to estimate the effect of multiple factors on the probability of an event and can be applied in risk analysis and other decision problems. The analysis provides variance predictions of the quantity of interest that incorporate dependencies amongst the various experts. Unlike other combination methods for expert judgment, in this form we may learn about the dependencies between the experts from their responses. The analysis is applied to a data set of expert judgments elicited during the Washington State Ferries Risk Assessment. The effect of the statistical dependence amongst experts is compared to an analysis assuming independence amongst them.

Keywords: Expert judgment; Pairwise Comparisons; Bayesian statistics; Multivariate analysis.

1. Introduction

Pate-Cornell (1996) discusses the necessity of using expert judgment in risk analysis when sufficient data is not available. Many applications of expert judgment involve the estimation of the probabilities of events, often with these probabilities affected by multiple factors. Merrick et al. (2000) propose an expert judgment elicitation method that estimates the effect of multiple factors on the probability of an event. This form of elicitation has been applied in the Prince William Sound (PWS) Risk Assessment (Merrick et al. 2002) and the Washington State Ferries Risk Assessment (WSF) van Dorp et al. 2001) to estimate the probability of human error given organizational factors, such as the experience and training of the crew, and to estimate the probability of an accident given situational factors, such as the proximity and type of nearby vessels and the environmental conditions at the time. While the elicitation method was proposed for use in risk analysis, it can be applied in other decision situations where applicable data is lacking.

The form of the elicitation is pairwise comparison of scenarios in which the event might occur. Multiple factors describe the two scenarios to the expert in a meaningful manner and in each comparison one factor is changed between the two scenarios. The method is akin to that in Bradley and Terry (1952), but the aim is to estimate the effect of the multiple factors rather than developing a scale for a single factor. In the PWS and WSF studies, a classical multiple regression was performed on the elicited expert data by assuming a log-linear model for the accident probability as a function of the factors.

The presence of uncertainty in analyzing risk is well recognized and discussed in the literature (Cooke 1997). However, these uncertainties are often ignored or underreported in studies of controversial or politically sensitive issues (Paté-Cornell 1996). The Bayesian paradigm allows the representation of both aleatory uncertainty (the randomness of the system itself) through the probability model and epistemic uncertainty (the lack of knowledge about the system) through prior distributions on the model parameters (Apostalakis 1978; Cooke 1991; Hofer 1996; Hora 1996; Winkler 1996). It is also well accepted that the judgments of multiple experts can be correlated and that the treatment of these correlations is necessary for proper analysis of such data (Winkler 1981; French 1980 1981; Lindley 1983 1985; Mosleh et al. 1988; Clemen 1987; Jouini and Clemen 2002). Such correlation is often introduced, in the language of Clemen (1987), by overlapping information available to the experts and thus used in determining their responses to the questionnaires.

Szwed et al. (2004) develop a Bayesian analysis of expert judgments elicited using the pairwise comparisons from Merrick et al. (2002). However, the Bayesian analysis in Szwed et al. assumes that the responses of the experts are independent. In this paper, we follow the development of Winkler (1981) by assuming that the errors in the judgments of the experts are drawn from a multivariate normal distribution. There is a fundamental difference, however. In Winkler's work, the experts are assessing the value of one continuous quantity and the covariance matrix is treated either as a hyperparameter (i.e. the components of the matrix need to be assessed by the decision maker) or can be assigned a prior distribution that is not updated in the final analysis. In our case, the experts are assessing the impact of multiple factors and thus give multiple judgments of related quantities. This allows us to treat the covariance matrix as a model parameter and thus learn about the dependencies between the experts.

Our approach can be considered a regression extension of Winkler (1981). We develop a multivariate regression analysis of the responses to the questions that allows for the correlation between experts. While the analysis mirrors the development of Bayesian multivariate regression (Press 1982), it is a special case as each expert is providing judgments on the same quantities, not different quantities as in the case of a full multivariate regression. This requires the complete development of the likelihood, prior and posterior forms. We note, however, that the ability to learn about the dependencies between the experts is gained at a price. While Winkler obtains the decision maker's predictive distribution on the quantity of interest, the multivariate regression set-up obtains the decision maker's multivariate predictive distribution of the judgments the various experts would provide. This requires the aggregation of the predictions to obtain a single distribution. To achieve this aggregation, we follow the minimum variance combination of Newbold and Granger (1974), the formula for which is in fact the same as that obtained by Winkler (1981).

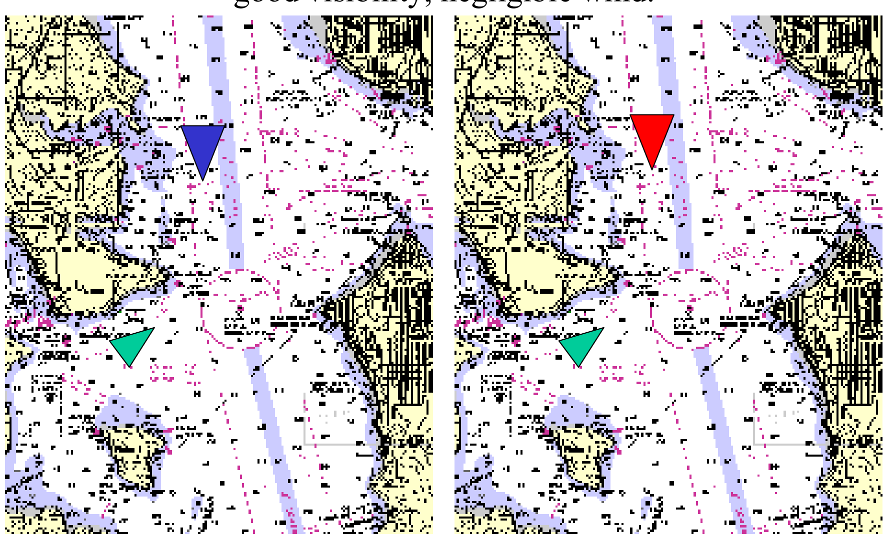
The outline of the paper is as follows. In Section 2, we illustrate the type of question used in our elicitation technique with an example drawn from a maritime risk study. We then illustrate the form of the underlying probability model assumed and show how this leads to regression as a suitable analysis methodology. Our multivariate extension is justified and developed in Section 3. The expert judgment data collected in the WSF Risk Assessment and the process used to collect it is described in Section 4 and this data is used to illustrate the use of our analysis method. We compare the analysis incorporating dependencies herein with one that is analogous to Szwed et al. (2004) assuming independence. Some concluding remarks are drawn in Section 5.

2. The Elicitation Method

2.1 The Questions

The following discussion will be based on the questionnaire used in the WSF Risk Assessment, although the same technique was used in the PWS Risk Assessment and can be used in many risk analysis and general decision problems. As an example, we shall examine the questionnaire for the likelihood of a collision between a ferry and another vessel given that the ferry has suffered a navigational aid (radar) failure. To assess the probability of an accident, experts were asked to compare two situations, as shown in Figure 1. This is essentially a pairwise comparison type of question (Bradley and Terry 1952). However, the questionnaires are used to estimate the effect of several factors, rather than the single factor in standard pairwise comparisons.

Issaquah class ferry on the Bremerton to Seattle route in a
crossing situation within 15 minutes, no other vessels around,
good visibility, negligible wind.



Other vessel is a navy vessel | Other vessel is a product tanker

Figure 1. An example of the type of question used in the expert judgement

The questions ask the expert to consider two situations between which only one factor has changed. The basic situation in Figure 1 is an Issaquah class ferry traveling from Bremerton to Seattle on a clear day with no wind. There is another vessel crossing the bow of the ferry less than 1 mile away. In the situation on the left-hand side, the other vessel is a Navy vessel, while on the right-hand side it is a product tanker. The questions were asked in the format of Figure 2. The responses were given on the scale at the bottom of Figure 2, which is taken from Saaty (1977). We interpret their response as the ratio of the probability of an accident in the two situations pictured. If the expert circled a "1", the two probabilities would be equal. We assume that if the expert circled the "9" on the right (left) then the ratio of the probabilities would be 9 (1/9).

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 2. An example of the question format

2.2 The Probability Model

The model assumed in the PWS and WSF Risk Assessments takes the form of a proportional probabilities model, based on the idea of the proportional hazards model (Cox 1972). Let $X = \left(x_1, \ldots, x_q\right)^T$ denote the q factors describing a situation in which the event of interest could occur. The conditional probability of the event, given the situation defined by X, is assumed to be

$$P(Event \mid X) = p_0 \exp\left(X^T \beta\right), \tag{1}$$

where $\beta = \left(\beta_1, \ldots, \beta_q\right)^T$ is a vector of q parameters and p_0 is a baseline probability parameter. Examining the ratio of the probability of the event in two situations reveals the convenience of this form. Consider two situations defined by the factor vectors X_1 and X_2. The ratio of the probabilities is

$$\frac{P(Event \mid X_1)}{P(Event \mid X_2)} = \frac{p_0 \exp(X_1^T \beta)}{p_0 \exp(X_2^T \beta)} = \exp\left((X_1 - X_2)^T \beta\right), \tag{2}$$

where (X_1 - X_2) denotes the difference vector between the two factor vectors. Thus, for this probability model, the ratio of the probabilities of the event given the two situations depends solely upon the difference between the two situations and the parameter vector β.

2.3 Analyzing the Experts' Responses

Recall the format of the questionnaires demonstrated in Figure 1. Each question asked the experts to assess the relative likelihood (or ratio of probabilities) of the event (a collision) given the two situations. Multiple experts complete each questionnaire, so there are multiple responses to each question. Let the experts be indexed by $j(= 1, \ldots, p)$ and the questions be indexed by $i(= 1, \ldots, N)$, so the experts' responses can be denoted $z_{i,j}$. We now have that $z_{i,j}$ is the j-th expert's estimate of the ratio of probabilities for the i-th question, while the model gives this relative probability as $\exp\left(X_i^T \beta\right)$, where X_i is a vector representing the difference between the two situations in question i $\left(X_{i,1} - X_{i,2}\right)$. This gives the basis for the regression equation used, specifically

$$\ln(z_{i,j}) = X_i^T \beta + u_{i,j} \tag{3}$$

where $u_{i,j}$ is the residual error term representing the variation between the experts' responses around the model.

Assuming that the errors $u_{i,j}$ are independent and normally distributed with zero mean and variance σ^2, this equation is a standard linear regression, where $y_{i,j} = \ln\left(z_{i,j}\right)$ is the dependent variable, X_i is the vector of independent variables, β is a vector of regression parameters and $u_{i,j}$ is the error term. Clemen and Reilly (1999) observe that it is often necessary in expert judgment analysis to use such transformations to arrive at the normal distribution. Kadane et al. (1980) develop a method for assessing prior hyperparameters on a linear regression model; however, their approach is based on direct assessments rather than pairwise comparisons. A conjugate Bayesian analysis of (3) is developed in Szwed et al. (2004) assuming conditional independence of the experts' responses given the model parameters. Pulkkinen (1993 1994a 1994b) was first to introduce, to the best of our knowledge, a Bayesian analysis of pairwise comparisons, but his Bayesian paired comparison inference model does not allow for updating of the dependence amongst experts

3. Analysis for Correlated Experts

3.1 A Multivariate Model

Clemen (1986 1987), Winkler (1981) and Mosleh et al. (1988) discuss the need for the representation of correlation between the experts in the analysis of expert judgment data. Winkler (1981) develops an aggregation technique for experts' assessments of a single, continuous quantity θ using the multivariate normal distribution, although here we follow more the form and notation of Clemen and Winkler (1985). If we denote the experts' point estimates of θ as $\mu = (\mu_1, \ldots \mu_p)$ and let $e_i = \mu_i - \theta$ be their judgment errors around the parameter θ, then Winkler's likelihood is formed by assuming that

$$e = \begin{pmatrix} e_1 \\ \vdots \\ e_p \end{pmatrix} \sim MVNormal(\underline{0}, \mathbf{\Sigma}),$$

where $MVNormal(\underline{0}, \mathbf{\Sigma})$ denotes a multivariate normal distribution with mean vector $\underline{0}$, a vector of p zeros, and covariance matrix $\mathbf{\Sigma}$. Winkler specifies the decision maker's prior distribution on θ as diffuse and updates using the multivariate normal likelihood $L(\theta; \mu_1, \ldots, \mu_p, \mathbf{\Sigma})$. Winkler's initial set-up requires the decision maker to specify the covariance matrix $\mathbf{\Sigma}$ as a hyperparameter of the analysis. Winkler shows that the posterior distribution of θ can then be re-written as

$$\pi(\theta; \mu, \mathbf{\Sigma}) \propto \exp\left(-\left(\theta - \mu^*\right)^2 / 2\sigma^{*2}\right) \tag{4}$$

where

$$\mu^* = \underline{1}^T \mathbf{\Sigma}^{-1} \mu / \underline{1}^T \mathbf{\Sigma}^{-1} \underline{1} \tag{5}$$

$$\sigma^{*2} = 1 / \underline{1}^T \mathbf{\Sigma}^{-1} \underline{1} \tag{6}$$

and $\underline{1}^T = (1,\ldots,1)$ is a vector of p 1's. Winkler's second set-up allows the decision maker to specify a prior distribution on $\mathbf{\Sigma}$, specifically an inverted Wishart distribution. However, the prior is not updated in the posterior analysis as each expert is only supplying one estimate and thus there is not sufficient information.

In our case, the single quantity θ is replaced by the multiple assessments of $\exp(X_i^T \beta)$ ($i = 1,\ldots,N$) that are linked by the common parameter vector β. Thus there are multiple assessments made by multiple experts, which we denoted by $y_{i,j} = \ln(z_{i,j})$. We may mirror Winkler's development by defining $u_{i,j} = y_{i,j} - X_i^T \beta$ and letting

$$u_i^T = \begin{pmatrix} u_{i,1} \\ \vdots \\ u_{i,p} \end{pmatrix} \sim MVNormal(\underline{0}, \mathbf{\Sigma}). \tag{7}$$

We may re-write this model in matrix form to obtain

$$\begin{pmatrix} y_{1,1} & \cdots & y_{1,p} \\ \vdots & \ddots & \vdots \\ y_{N,1} & \cdots & y_{N,p} \end{pmatrix} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,q} \\ \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,q} \end{pmatrix} \begin{pmatrix} \beta_1 & \cdots & \beta_1 \\ \vdots & \ddots & \vdots \\ \beta_q & \cdots & \beta_q \end{pmatrix} + \begin{pmatrix} u_{1,1} & \cdots & u_{1,p} \\ \vdots & \ddots & \vdots \\ u_{N,1} & \cdots & u_{N,p} \end{pmatrix}$$

or

$$\mathbf{Y} = \mathbf{X}\beta\underline{1}^T + \mathbf{U} \tag{8}$$

This equation is similar to a full multivariate regression model

$$\mathbf{Y} = \mathbf{XB} + \mathbf{U}, \tag{9}$$

where **X** is a $(N \times q)$-matrix of differences between the q covariates for the N questions, **B** is a $(q \times p)$-matrix where each column represents the covariate effect parameters for an expert and **U** is a $(N \times p)$-vector of residual errors. The difference between (8) and (9) is that in (8) columns of the regression parameter matrix are restricted to be equal as each expert is providing estimates of the same quantity.

The form in (8) suggests that we follow the analysis of a multivariate regression model, such as that developed in Press (1982). Equation (7) implies that the rows of $\mathbf{U}$ are independent vectors distributed according to a multivariate normal with a zero mean vector and covariance matrix $\mathbf{\Sigma}$. The rows of $\mathbf{U}$ are assumed to be independent as they are responses to the individual questions, but the columns are dependent as they represent the responses of the experts to each question. Analyzing the model in (7) will make our analysis different from Winkler's, as the prior distribution on $\mathbf{\Sigma}$ will be updated by the judgments of the experts.

3.2 Posterior Analysis

While the following analysis mirrors the Bayesian analysis in Press (1982), the likelihood and posterior distributions for the column restricted form in (8) requires full development which can be found in Appendices A and B. We summarize these results here for brevity. The likelihood for the parameter restricted multivariate regression model in (8) is

$$p(\mathbf{Y} \mid \beta, \mathbf{\Sigma}, \mathbf{X}) \propto |\Sigma|^{-1/2} \exp\left\{-\tfrac{1}{2} tr\left(\mathbf{V}\Sigma^{-1}\right)\right\} \exp\left\{-\tfrac{1}{2}\left(\beta - \overline{B}\right)^T \left(\sigma_\Sigma^2 \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\right)^{-1} \left(\beta - \overline{B}\right)\right\} \tag{10}$$

where $\mathbf{V} = \left(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\right)^T \left(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\right)$ is the usual sufficient statistic for the unrestricted model, $\widehat{\mathbf{B}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}$ is the least squares estimates of $\mathbf{B}$ under the unrestricted model, $\overline{B} = \frac{1}{p}\widehat{\mathbf{B}}\underline{1}$ is the average of the p experts' least squares estimates under the unrestricted model and $\sigma_{\Sigma}^{2} = \frac{1}{p^{2}} \underline{1}^{T} \Sigma \underline{1}$ is the average variance across all experts. See Appendix A for the derivation of this form.

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

$$(\Sigma) \sim Inv - Wishart(\mathbf{G}, m), \tag{11}$$

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

$$(\beta \mid \mathbf{\Sigma}) \sim MVNormal(\phi, \sigma_{\Sigma}^{2}\mathbf{A}). \tag{12}$$

ϕ, $\mathbf{A}$, $\mathbf{G}$ and m are arbitrary prior hyperparameters determined by the decision maker.

Given the experts' responses to the questionnaires, the posterior distributions obtained in Appendix B are

$$(\mathbf{\Sigma} \mid \mathbf{Y}, \mathbf{X}) \sim Inv - Wishart(\mathbf{G} + \mathbf{V}, m + q), \tag{11}$$

and

$$(\beta \mid \mathbf{Y}, \mathbf{X}, \mathbf{\Sigma}) \sim MVNormal\left(\left(\mathbf{X}^{T}\mathbf{X} + \mathbf{A}\right)^{-1}\left(\mathbf{X}^{T}\mathbf{X}\overline{\mathbf{B}} + \mathbf{A}\phi\right), \sigma_{\Sigma}^{2}\left(\mathbf{X}^{T}\mathbf{X} + \mathbf{A}\right)\right). \tag{12}$$

Thus the analysis is conjugate, making calculation, and therefore application, easier.

A result that will be useful later in the analysis is the marginal posterior distribution of the β_k $(k = 1,...,q)$ that will be univariate student-t distributions with $m + q + 1$ degrees of freedom given by

$$(\beta_k \mid \mathbf{Y}, \mathbf{X}) \sim student - t\left(\left(\mathbf{X}^{T}\mathbf{X} + \mathbf{A}\right)^{-1}\left(\mathbf{X}^{T}\mathbf{X}\overline{\mathbf{B}} + \mathbf{A}\phi\right), \frac{\underline{1}^{T}(\mathbf{G} + \mathbf{V})\underline{1}}{(m+q)p^{2}} \left(\mathbf{X}^{T}\mathbf{X} + \mathbf{A}\right)_{j,j}^{-1}\right)$$

where $\left(\mathbf{X}^T\mathbf{X}+\mathbf{A}\right)^{-1}_{j,j}$ is the j-th diagonal element of $\left(\mathbf{X}^T\mathbf{X}+\mathbf{A}\right)^{-1}$. As $\mathbf{\Sigma}$ follows an inverted Wishart distribution the marginal distributions of the off-diagonal elements are unknown, so samples are taken to observe their form.

3.3 Prediction

Once the Bayesian update has been performed, the next step is to develop the predictive distribution. In this case, we wish to predict how much more likely an accident is in one scenario compared to the other, or the ratio of the probabilities of an accident in the two scenarios. One would imagine that the development of a predictive distribution would mirror the development for multivariate regression, but with posterior distributions drawn from our parameter restricted model form. However, considering (8) reveals that if we denote the difference between the two scenarios by x^*, then the prediction would be of $y^* = \left(y_1^*,\ldots,y_p^*\right)$, a vector of the various experts' judgments of the ratio. We merely wish to estimate one value for the ratio of probabilities, not one for each expert.

Instead, let us predict the weighted average of the p experts' responses. In the following we apply the minimum variance weighted average aggregation of Newbold and Granger (1974). As noted by Clemen and Reilly (1999), the formula obtained by Newbold and Granger is the same as that obtained by Winkler (1981). We wish to predict the single quantity $\bar{y}^* = w^T y^*$, where $w = \left(w_1,\ldots,w_p\right)$ is a vector of weights that sum to one. The model form in (8) implies that

$$\left(\mathbf{y}^* \mid \mathbf{x}^*, \beta, \mathbf{\Sigma}\right) \sim MVNormal\left(\mathbf{x}^{*T}\beta\underline{1}, \mathbf{\Sigma}\right) \tag{13}$$

with each expert having the same mean, but not necessarily the same variance or covariances amongst each other.. Using the results from Newbold and Granger, the minimum variance aggregation $\bar{y}^*$, conditioned on β and $\mathbf{\Sigma}$, will be a normal distribution with mean and variance mirroring (5) and (6), specifically

$$\mu^* = \underline{1}^T \mathbf{\Sigma}^{-1} x^{*^T} \beta \underline{1} / \underline{1}^T \mathbf{\Sigma}^{-1} \underline{1} \tag{14}$$

$$\sigma^{*2} = 1 / \underline{1}^T \mathbf{\Sigma}^{-1} \underline{1} \tag{15}$$

Note that $x^{*^T} \beta$ is a scalar, so can be brought to the front of (14) leaving the rest of the terms to cancel out, implying that $\mu^* = x^{*^T} \beta$. This is to be expected as each expert had the same mean prediction in (13) and the weights sum to one. So the minimum variance weighted prediction, conditioned on β and $\mathbf{\Sigma}$, is

$$\left(\mathbf{y}^* \mid \mathbf{x}^*, \beta, \Sigma\right) \sim MVNormal\left(\mathbf{x}^{*T} \beta, \sigma^{*2}\right) \tag{16}$$

The posterior distribution of σ^{*2} is scalar Wishart with posterior parameters $1/\underline{1}^T (\mathbf{G} + \mathbf{V})^{-1} \underline{1}$ and $m + q$ using (11) and formulae from Dawid (1981). The posterior distribution of β is given in (12). Thus the posterior predictive distribution of $\bar{y}^*$ will then be a student-t distribution with $m + p + q - 1$ degrees of freedom given by

$$\bar{y}^* \sim student - t\left(\mathbf{x}^{*T} \left(\mathbf{X}^T \mathbf{X} + \mathbf{A}\right)^{-1} \left(\mathbf{X}^T \mathbf{X} \overline{\mathbf{B}} + \mathbf{A} \phi\right), (m + q) \underline{1}^T (\mathbf{G} + \mathbf{V})^{-1} \underline{1} \mathbf{x}^{*T} \left(\mathbf{X}^T \mathbf{X} + \mathbf{A}\right) \mathbf{x}^*\right) \tag{14}$$

To complete the prediction recall that $y_{i,j} = \ln\left(z_{i,j}\right)$, the prediction for $z_{i,j}$ is then a log student-t distribution.

4. Example Results

4.1 Elicitation for the WSF Risk Assessment

Expert judgment was used in the WSF Risk Assessment to estimate the effect of risk factors on the probability of a collision given the occurrence of some triggering incident. The risk factors are listed in Table 1 and include the ferry class and route, the type, proximity and angle of interaction of the closest two vessels, the visibility conditions and wind speed and direction. For a discussion of the derivation of the scales used for these risk factors see Szwed et al. (2004).

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

Experts may be classified in three categories (DeWispelare et al. 1995):

- **normative experts** who have the analysis background to quantify the judgments of the substantive experts and combine their judgments.
- **generalists** who have a thorough understanding of the project and play a role in defining the issues addressed and communicating with the experts
- **substantive experts** who have the deep knowledge and experience of a system that allow them to provide information about the functioning of that system and

Certain members of the risk assessment team were normative experts, with knowledge of decision theory, probabilistic reasoning and expert elicitation techniques. Other members were generalists with both maritime experience, knowledge of maritime risk issues and systems engineering techniques. The substantive experts used in the study were the ferry captains that worked relief, filling in for captains on vacation or sick leave across all ferry routes. This ensured that the experts had a thorough knowledge of the entire system, not just a specific route. Each of the experts used had over 10 years of experience with the WSF.

The elicitation team first provided to the substantive experts some background on the project followed by an explanation of the questionnaires and their purpose. Example questions were presented similar to Figure 1, but in the context of driving a car on the highway. This context was also explained in terms of several risk factors. The highway transportation mode was chosen over maritime examples to avoid biasing the experts before beginning the questionnaires and because everyone was familiar with the situations defined. The experts were then given an example question to consider in the driving example and discussion encouraged between the experts to ensure the idea was understood. It was important to remind the experts to look at all the risk factors in the question, rather than just the one that changed between the two situations as there can be interactions between the risk factors.

Each questionnaire consisted of sixty comparisons of the type shown in Figure 2. The questionnaires were designed to collect the maximum amount of information from the sixty questions and to ensure that sufficient information was elicited to ensure the estimation of the main ten risk factors and six pre-defined interactions between risk factors. The questions were asked in random order. The randomization of the questions meant that deliberate attempts to bias the results were difficult. Tests on the responses were performed to ensure that the experts' responses were not affected by fatigue.

4.2 Prior Distributions

The first step in analyzing the expert judgment data is the specification of the prior hyperparameters. Clemen (1986) discusses the concept of aggregation of the decision maker's beliefs with those of the experts. In our applications the decision makers have claimed ignorance of the effect of the factors in Table 1 on the probability of a collision and wished for the experts' beliefs to dominate the predictions. In the Bayesian sense, this means specifying suitably vague priors.

We assumed that ϕ, the vector of the prior means on β, is a vector of zeros, which indicates that a priori all covariates have on average no effect on the probability of an accident. The prior matrix $\mathbf{A}$ is assumed to be an identity matrix to indicate no prior covariance between the parameters in β. The prior matrix $\mathbf{G}$ is assumed to be an identity matrix, indicating no prior knowledge of correlations between the experts, while m is assumed to be 0.380341 calculated by Szwed et al. (2004) to represent a priori that all expert respond to the N questions completely at random..

Figure 3 shows the resulting marginal prior distributions of the β parameters, represented by a circle for their mean and whiskers showing their prior 90% credibility interval. One may see that the prior assumptions are diffuse. Figure 4 shows the prior distribution of $\frac{P(Collision \mid Nav.Fail, X_1)}{P(Collision \mid Nav.Fail, X_2)}$ where X_1 is the scenario on the left and X_2 is the scenario on the right of Figure 1. Note that the distribution has median of one, meaning collisions are equally likely in each situation, and a wide variability. In fact a 90% credibility interval for this ratio of probabilities is $1.88x10^{-35}$ to $5.32x10^{34}$.

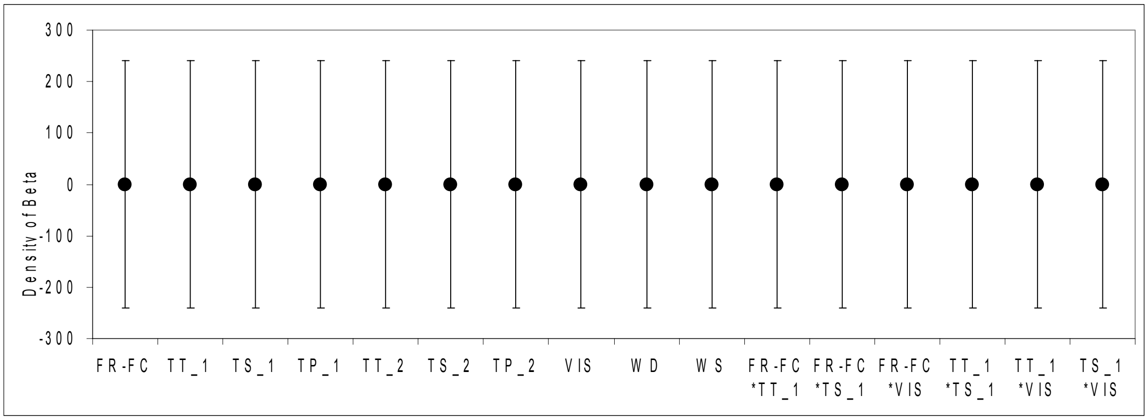


Figure 3. The marginal prior distribution of β.

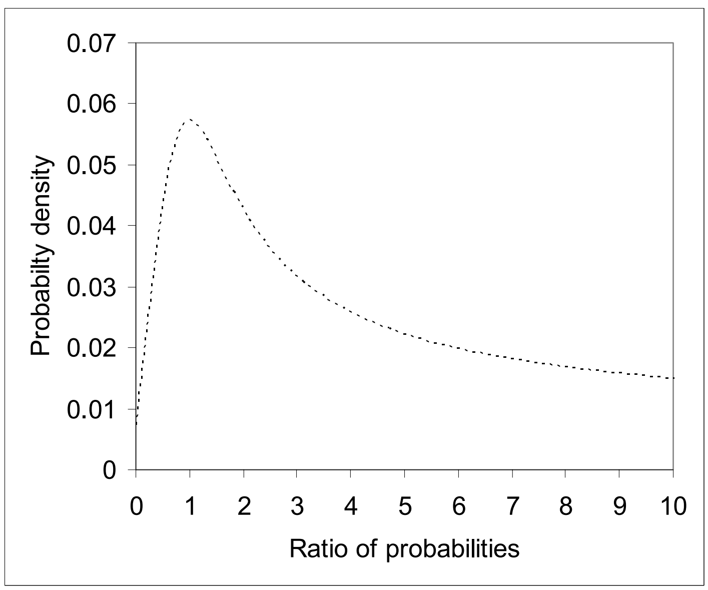


Figure 4. The prior distribution of the ratio of probabilities for the scenarios pictured in Figure 1.

4.2 Posterior Distributions

After updating with the experts' responses using the formulae developed in Section 3.2, the marginal posterior distributions of the β parameters are as shown in Figure 5. To demonstrate the advantage of our model including dependence, we compare the results to the independent experts model in (9) developed in Szwed et al. (2004). Figure 6 shows the marginal posterior distributions of the β parameters obtained using the independent experts model from Szwed et al. (2004), on the same scale as Figure 5. Note that while the mean values are similar, some slight differences can be observed amongst them in Figures 5 and 6. More noticeable, however, is that the posterior variance of every parameter is less in the dependent experts model (Figure 5) than that observed in the independent one by Szwed et al. (Figure 6).

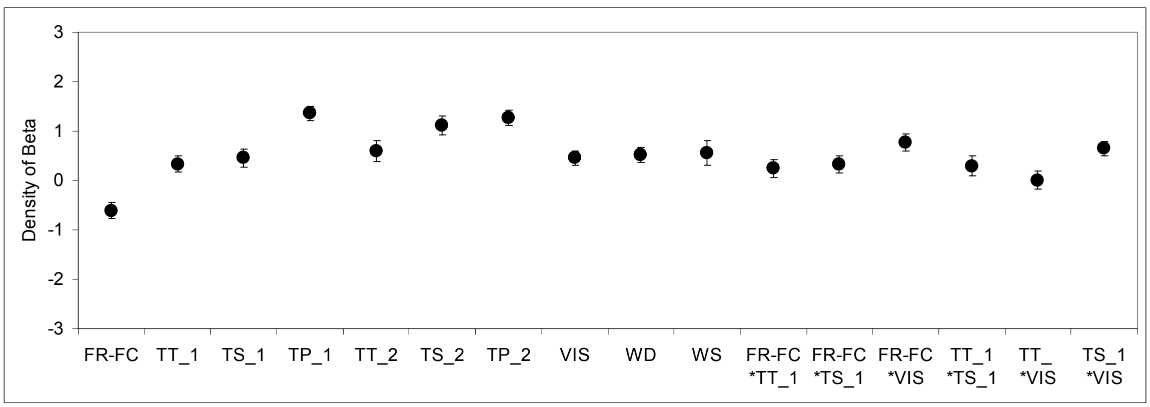


Figure 5. The marginal posterior distribution of β assuming dependent experts.

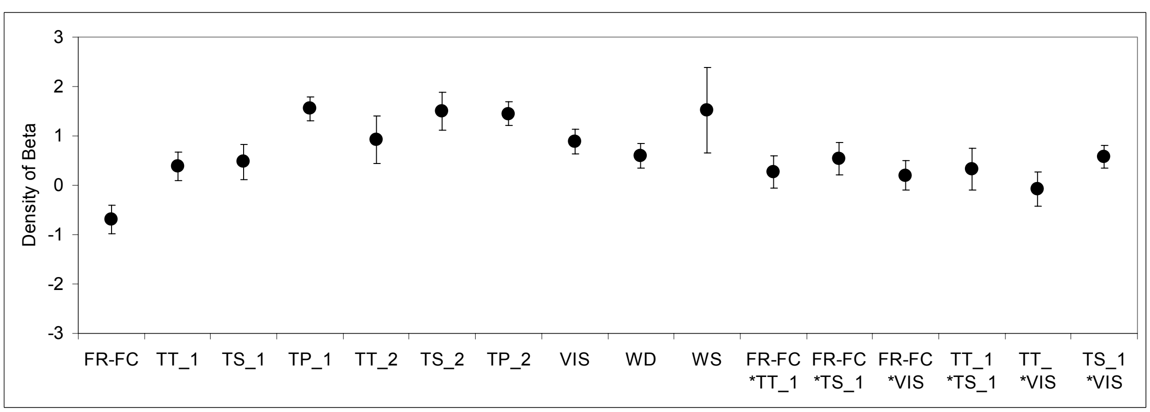


Figure 6. The marginal posterior distribution of β assuming independent experts.

Of particular interest in this analysis, is the covariance matrix $\mathbf{\Sigma}$ representing the dependencies between the experts. Figure 7 shows the posterior distribution of the corresponding correlation matrix with the correlation between the i-th and j-th experts indicated by the notation i,j on the top left of each histogram. Only lower triangular elements are shown to reduce clutter in the figure. A vertical line is drawn at 0, indicating no dependence between the two experts. Thus a histogram showing samples to the right of the line indicates a posterior probability that the two experts have positive dependence or overlapping information, while samples to the left indicates negative dependence or different information.

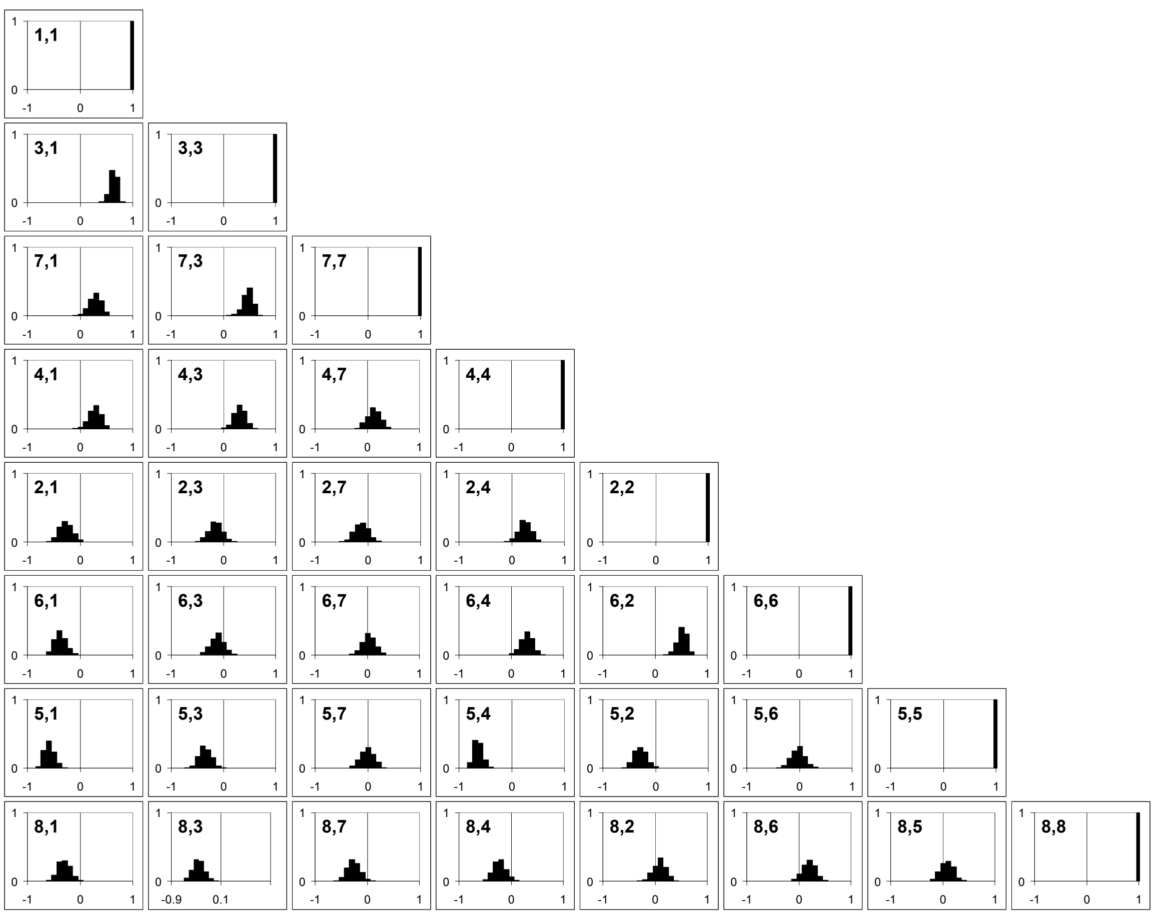


Figure 7. The posterior distribution of the expert correlation matrix.

The experts are re-ordered in Figure 7 to show groupings that appear in the correlations. Note that experts 1, 3 and 7 have a tendency to agree in their responses. Experts 2, 4 and 6 also tend to agree with each other. These groups are not completely disparate, however. Experts 2 and 6 tend to disagree with experts 1, 3 and 7, but expert 4 tends to agree with them. Expert 5 tends to disagree with experts 1, 2, 3, 4 and 6, but does not conclusively agree or disagree with expert 7. Expert 8 disagrees with the group of experts 1, 3 and 7, but does not conclusively agree or disagree with the group 2, 4 and 6 or with 5.

Consider the effect that this correlation between the experts will have on the variance of the experts' responses. The posterior expected standard deviation of the average of the residuals for the p experts, which we denoted σ_{Σ}^2, evaluates to 0.49, whereas the posterior expected standard deviation of the residuals in the independent experts model is 1.05, over twice as high.

4.3 Posterior Predictions

We have seen the difference these correlations make in the variance of the β parameters and to the precision of the residuals, but of more interest is their effect on the predictive distribution. Using the result in (14), Figure 9 shows the posterior distribution of $\frac{P(\text{Collision} \mid \text{Prop. Fail}, X_1)}{P(\text{Collision} \mid \text{Prop. Fail}, X_2)}$ where X_1 is the scenario on the left and X_2 is the scenario on the right of Figure 1. The prior distribution is also shown as a dotted line to show the effect of updating. The posterior 90% credibility interval on the ratio of probabilities is 4.01 to 5.00, with a half-width of 0.50. The reader should note that the question in Figures 1 and 2 is different from that illustrated in Szwed et al. (2004).

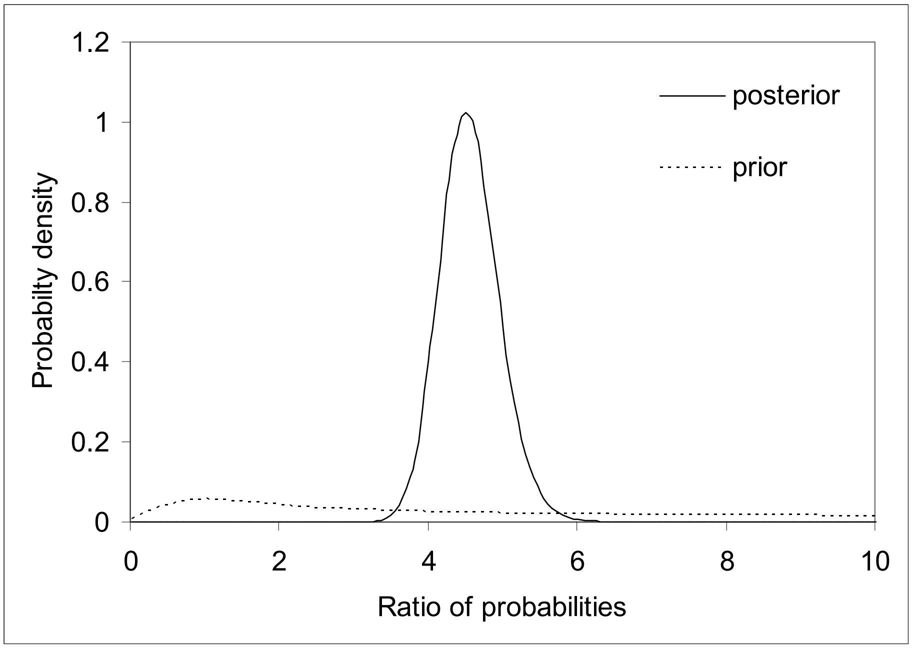


Figure 9. The prior and posterior density of the ratio of probabilities for the scenarios pictured in Figure 1.

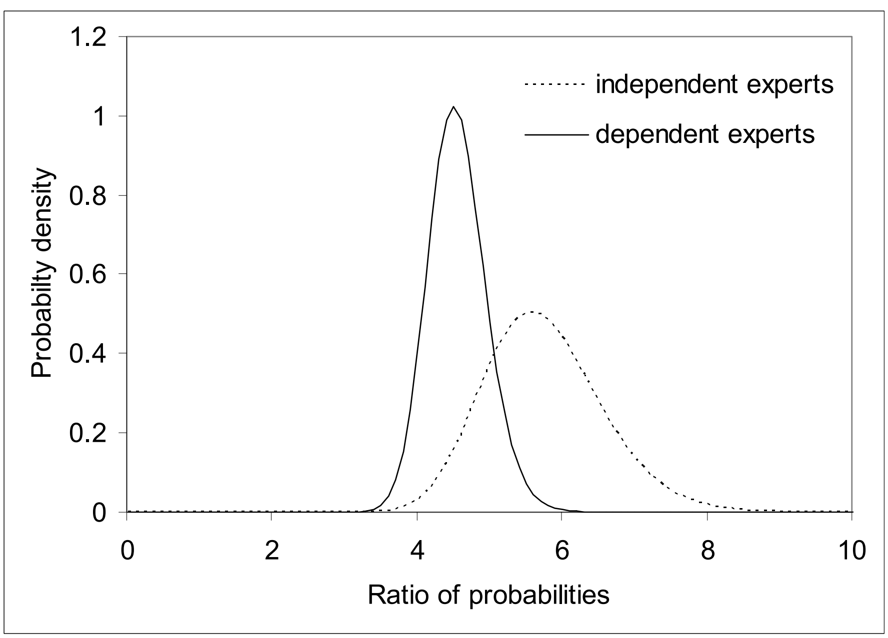


Figure 10. The posterior density of the ratio of probabilities for the scenarios pictured in Figure 1 assuming dependent and independent experts.

Figure 10 compares the posterior distribution of the ratio of probabilities from Figure 9 to the same prediction obtained using the independent experts model of Szwed et al. (2004). The independent expert model gives a higher posterior expected value of the ratio of probabilities, 5.59 as opposed to 4.48 with the dependent experts model, and a larger variance. The posterior 90% credibility interval is now 4.43 to 7.04, with a half-width of 1.3 compared to 0.50 for the dependent expert model.

4.4 The Experts' Posterior Weights

Of interest is the weight that each expert receives in the final predictions. These weights cannot be obtained in closed form and thus are found using sampling approximations. The weights depend only on the posterior distribution of $\mathbf{\Sigma}$ as they are chosen to minimize the variance of $\overline{y}^* = w^T y^*$ which equals $w^T \mathbf{\Sigma} w$ when conditioned on $\mathbf{\Sigma}$. Figure 11 shows 90% credibility intervals of the experts' posterior weights as the whiskers of box plots and the interquartile range as the boxes, each estimated using sampling approximations.

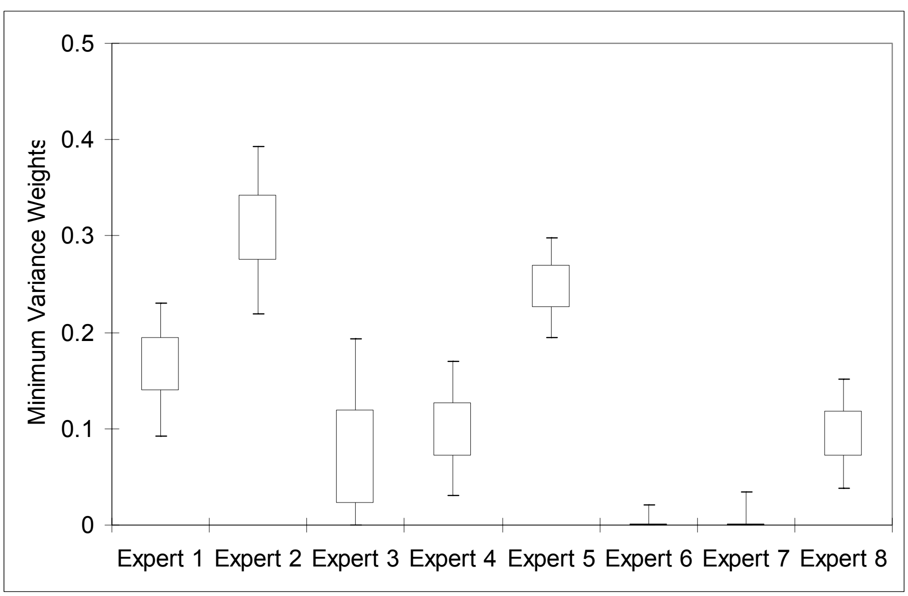


Figure 11. Box plots for the expert weightings using regular weights restrictions.

Winkler and Clemen (1992) discuss the sensitivity of the aggregation formulae from both Winkler (1981) and Newbold and Granger (1974) to high dependencies. They report severe problems including negative weights and weights above one, neither of which is seen in Figure 11 for this expert group's dependence structure.

The results in Figure 11 do raise the question of why certain experts receive a higher weight and some do not. Recall that $\bar{y}^* = w^T y^*$ implying that the variance of the weighted prediction obtained by Newbold and Granger (1974) is given by

$$\mathrm{var}(\bar{y}^* \mid \Sigma) = \sum_i w_i^2 \sigma_{i,i} + \sum_{i \neq j} w_i w_j \sigma_{i,j} ,$$

where $\sigma_{i,j}$ is the element of $\mathbf{\Sigma}$ in the i-th row and j-th column. As it is this expression that is minimized to find the weights of Newbold and Granger, it is apparent that to receive the most weight in the sum, experts should have low individual variance and be negatively correlated or uncorrelated with other experts that receive positive weight. Clemen and Winkler (1985) also conclude that positive (negative) correlations can reduce (increase) the precision of predictions and the value of information from dependent sources, although they use the results from Winkler (1981) in their investigations.

As we have seen in Figure 7, groups of experts with positive dependencies do appear in our expert judgment data set, implying overlapping information. These groups are then uncorrelated or negatively correlated with other individual experts or groups of experts. It does not make sense to include all the members of a group due to their positive covariance. Rather some experts from the group should get chosen to represent the overlapping information from that group. Expert groups or individual experts that contain different information, as represented by negative or zero covariance with other experts, should receive positive weight to include their different information. Recalling the apparent groupings amongst the experts in our analysis discussed in Section 4.2, experts 1 and 3 are chosen to represent the group of experts 1, 3 and 7. Experts 2 and 4 are chosen to represent the group of 2, 4 and 6. Experts 5 and 8 both receive some weight as they represent different information from the other two groups. This is an intuitively appealing feature of this expert weighting technique.

4.5 Expert Correlation: Good or Bad?

The comparison of the dependent and independent experts models validates the incorporation of expert dependency in the modeling and aggregation of the experts' opinions, but it does not demonstrate the impact that the correlation between the group of experts is having compared to an expert group with the same judgments about the effect of the factors but answering independently. The generalized variance of $\mathbf{\Sigma}$ is defined as the determinant of $\mathbf{\Sigma}$, denoted $|\mathbf{\Sigma}|$, and is used as a measure of the overall variance of correlated variables. Applications include multivariate quality control (Mastrangelo et al. 1996) and general linear models (Finn 1974; Tatsuoka 1971).

To obtain a measure of the effect of the dependencies between the experts, we consider the covariance matrix $\mathbf{\Sigma}$ compared to the same matrix $\mathbf{\Sigma}$ with the off-diagonal elements set to zero, denoted by $\mathbf{\Sigma}_0$. This modification leaves the experts with the same variance for their residuals around the aggregated model $x^T\beta$, but with no correlation. To compare the generalized variance for $\mathbf{\Sigma}$ and $\mathbf{\Sigma}_0$, we take the ratio $|\mathbf{\Sigma}_0|/|\mathbf{\Sigma}|$. This yields a statistic that takes values above (below) one if independent experts would increase (decrease) the generalized variance compared to the correlated experts. Recall, however, that we in fact have a distribution on $\mathbf{\Sigma}$. Thus we take samples from the posterior Wishart distribution of $\mathbf{\Sigma}$ and calculate the generalized variance ratio for each sample matrix. The distribution is difficult to read on this scale as we are trying to assess whether the values are above or below one. Thus we take the natural logarithm, to obtain the log generalized variance ratio. This statistic will take values above (below) zero if independent experts would increase (decrease) the generalized variance compared to the correlated experts.

Figure 8 shows the posterior distribution of the log generalized variance ratio. In Figure 8, the log generalized variance ratio is almost surely positive, indicating that independent experts would increase the generalized variance compared to the correlated experts. Thus the correlation structure of the experts in Figure 7 has in fact reduced the generalized variance compared to independent experts.

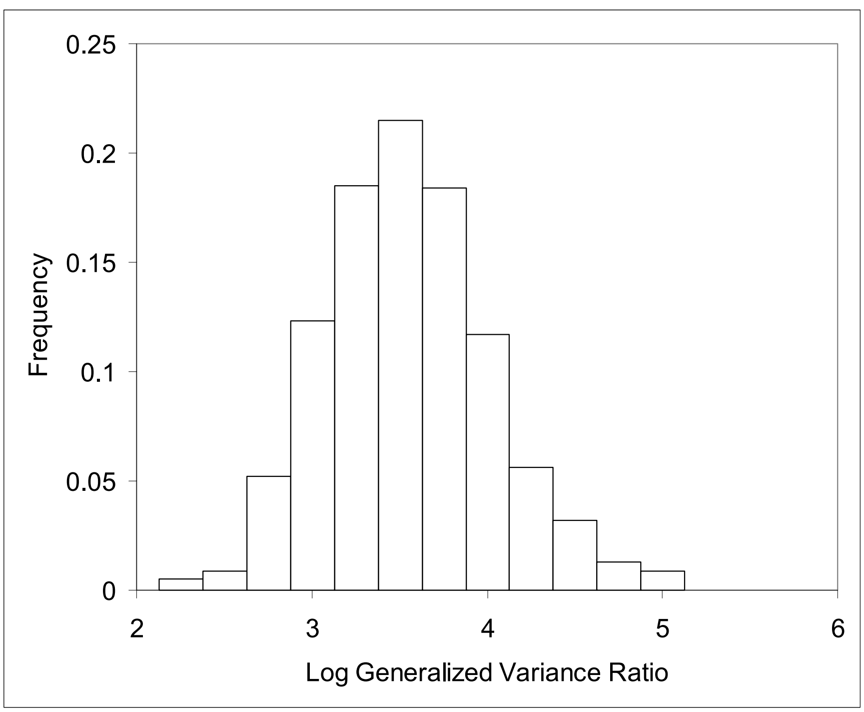


Figure 8. The log generalized variance ratio.

5. Conclusions

We have developed an analysis of an extended form of pairwise comparisons introduced in Merrick et al. (2000) from a Bayesian analysis that assumes that experts' responses are independent (Szwed et al. 2004) to one that allows for correlations between experts. The analysis was set up using the approach of Winkler (1981), but took the form of a special case of Bayesian multivariate regression and this required full development of likelihood and posterior distributions in Appendices A and B. The method was applied to expert judgment data elicited during the WSF Risk Assessment. The empirical results show that there were correlations between the experts in this data and that allowing for these correlations decreases the posterior variance in the predictions made using the model compared to those obtained in Szwed et al. (2004). This reduction in uncertainty could be critical in determining whether to apply proposed risk interventions when such risk interventions are evaluated using the output of this expert judgment methodology.

In a complete decision analysis, we must take care not only in the estimation of our mean predictions, but also their variances. Modeling dependencies amongst experts is a key part of getting the variance predictions right. Thus the development herein for such a widely applicable expert judgment elicitation technique is an important contribution to the literature on aggregation of expert opinion.

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.

Appendix A - The Likelihood

In this appendix, we follow the development of the likelihood for the full multivariate regression model in (9). At the appropriate point, we make an appropriate modification to find the likelihood for the restricted multivariate regression model in (8). Consider the likelihood for the full multivariate model in (9)

$$\mathbf{Y} = \mathbf{XB} + \mathbf{U}.$$

We assume that the rows of $\mathbf{U}$ are independent vectors distributed according to a multivariate normal with a zero mean vector and covariance matrix $\mathbf{\Sigma}$, which gives the likelihood (Press 1982) as

$$p(\mathbf{Y} \mid \mathbf{B}, \mathbf{\Sigma}, \mathbf{X}) \propto |\mathbf{\Sigma}|^{-\frac{N}{2}} \exp\left\{-\tfrac{1}{2} tr\left((\mathbf{Y}-\mathbf{XB})^T (\mathbf{Y}-\mathbf{XB})\mathbf{\Sigma}^{-1}\right)\right\} \tag{A.1}$$

Considering the main part of the term in the exponent, we may introduce terms in each bracket a term that sums to zero giving

$$\begin{aligned}
&(\mathbf{Y}-\mathbf{XB})^T (\mathbf{Y}-\mathbf{XB}) \\
&= \left(\mathbf{Y}-\mathbf{X\hat{B}}+\mathbf{X\hat{B}}-\mathbf{XB}\right)^T \left(\mathbf{Y}-\mathbf{X\hat{B}}+\mathbf{X\hat{B}}-\mathbf{XB}\right) \\
&= \left(\left(\mathbf{Y}-\mathbf{X\hat{B}}\right)+\mathbf{X}\left(\mathbf{\hat{B}}-\mathbf{B}\right)\right)^T \left(\left(\mathbf{Y}-\mathbf{X\hat{B}}\right)+\mathbf{X}\left(\mathbf{\hat{B}}-\mathbf{B}\right)\right) \\
&= \left(\mathbf{Y}-\mathbf{X\hat{B}}\right)^T \left(\mathbf{Y}-\mathbf{X\hat{B}}\right)+\left(\mathbf{\hat{B}}-\mathbf{B}\right)^T \mathbf{X}^T\mathbf{X}\left(\mathbf{\hat{B}}-\mathbf{B}\right) \\
&+\left(\mathbf{Y}-\mathbf{X\hat{B}}\right)^T \mathbf{X}\left(\mathbf{\hat{B}}-\mathbf{B}\right)+\left(\mathbf{\hat{B}}-\mathbf{B}\right)^T \mathbf{X}^T\left(\mathbf{Y}-\mathbf{X\hat{B}}\right)
\end{aligned} \tag{A.2}$$

If we let $\mathbf{\hat{B}} = \left(\mathbf{X^T X}\right)^{-1} \mathbf{X^T Y}$, then

$$\begin{aligned}
\mathbf{X}^T\left(\mathbf{Y}-\mathbf{X\hat{B}}\right) &= \mathbf{X}^T\left(\mathbf{Y}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}\right) \\
&= \mathbf{X}^T\left(\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right)\mathbf{Y} \\
&= \left(\mathbf{X}^T-\mathbf{X}^T\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\right)\mathbf{Y} \\
&= \left(\mathbf{X}^T-\mathbf{X}^T\right)\mathbf{Y} = \mathbf{0Y} = \mathbf{0}.
\end{aligned}$$

This makes the last two terms in (A.2) zero and our expression becomes

$$\begin{aligned}&(\mathbf{Y}-\mathbf{XB})^T(\mathbf{Y}-\mathbf{XB})\\&=(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})^T(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})+(\hat{\mathbf{B}}-\mathbf{B})^T\mathbf{X}^T\mathbf{X}(\hat{\mathbf{B}}-\mathbf{B})\end{aligned}$$

Substituting back in to the likelihood in (A.1) yields

$$|\mathbf{\Sigma}|^{-\frac{N}{2}}\exp\left\{-\tfrac{1}{2}tr\left((\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})^T(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})\mathbf{\Sigma}^{-1}+(\hat{\mathbf{B}}-\mathbf{B})^T\mathbf{X}^T\mathbf{X}(\hat{\mathbf{B}}-\mathbf{B})\mathbf{\Sigma}^{-1}\right)\right\},$$

which can be written as

$$|\mathbf{\Sigma}|^{-\frac{N}{2}}\exp\left\{-\tfrac{1}{2}tr\left((\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})^T(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})\mathbf{\Sigma}^{-1}\right)\right\}\exp\left\{-\tfrac{1}{2}tr\left((\hat{\mathbf{B}}-\mathbf{B})^T\mathbf{X}^T\mathbf{X}(\hat{\mathbf{B}}-\mathbf{B})\mathbf{\Sigma}^{-1}\right)\right\}$$

(A.3)

The first exponential term in (A.3) involves only $\mathbf{\Sigma}$ as a parameter, while the second involves both $\mathbf{B}$ and $\mathbf{\Sigma}$. Letting $\mathbf{V}=(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})^{\mathbf{T}}(\mathbf{Y}-\mathbf{X}\hat{\mathbf{B}})$ and multiplying each bracket in the quadratic form in the exponent by -1 leaves

$$|\mathbf{\Sigma}|^{-\frac{N}{2}}\exp\left\{-\tfrac{1}{2}tr\left(\mathbf{V}\mathbf{\Sigma}^{-1}\right)\right\}\exp\left\{-\tfrac{1}{2}tr\left((\mathbf{B}-\hat{\mathbf{B}})^T\mathbf{X}^T\mathbf{X}(\mathbf{B}-\hat{\mathbf{B}})\mathbf{\Sigma}^{-1}\right)\right\}$$

Substituting this result in to the likelihood in (A.3) gives

$$p(\mathbf{Y}\,|\,\mathbf{X},\mathbf{B},\mathbf{\Sigma})\propto|\mathbf{\Sigma}|^{-\frac{N}{2}}\exp\left\{-\tfrac{1}{2}tr\left(\mathbf{V}\mathbf{\Sigma}^{-1}\right)\right\}\exp\left\{-\tfrac{1}{2}tr\left((\mathbf{B}-\hat{\mathbf{B}})^T\mathbf{X}^T\mathbf{X}(\mathbf{B}-\hat{\mathbf{B}})\mathbf{\Sigma}^{-1}\right)\right\}$$

(A.4)

The first term in (A.4) can be considered the likelihood of $\mathbf{\Sigma}$, while the second term is the likelihood of $\mathbf{B}$ given $\mathbf{\Sigma}$. In fact, the likelihood of $\mathbf{\Sigma}$ is already in the form of the inverted Wishart distribution. However, while the likelihood of $\mathbf{B}$ given $\mathbf{\Sigma}$ resembles a multivariate normal distribution, we must convert $\mathbf{B}$ in to a vector. Press introduces a stacked vector form of $\mathbf{B}$, which we will denote β^*. The exact form of β^* is

$\left(\beta_{1,1} \cdots \beta_{q,1} \quad \beta_{1,2} \cdots \beta_{q,2} \cdots \beta_{q,1} \cdots \beta_{q,p}\right)^T$ or the vector of parameters for each expert stacked on top of each other.

Press shows that the likelihood can be re-written in terms of β^* as

$$P(\mathbf{Y} \mid \mathbf{X}, \beta, \mathbf{\Sigma}) \propto |\mathbf{\Sigma}|^{-\frac{N}{2}} \exp\left\{-\tfrac{1}{2} tr\left(\mathbf{V\Sigma}^{-1}\right)\right\} \exp\left\{-\tfrac{1}{2}\left(\left(\beta^* - \hat{\beta}^*\right)^T \left(\mathbf{\Sigma} \otimes \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\right)^{-1} \left(\beta^* - \hat{\beta}^*\right)\right)\right\}, \tag{A.5}$$

where $\hat{\beta}^*$ is the stacked vector form of $\hat{\mathbf{B}}$. The second term is now the likelihood of β^* given $\mathbf{\Sigma}$ and is in the form of a multivariate normal with mean $\hat{\beta}^*$ and covariance matrix $\mathbf{\Sigma} \otimes \left(\mathbf{X}^T\mathbf{X}\right)^{-1}$.

To this point, we have been considering the full multivariate regression model (9). Let us now consider the restricted model in (8). The only difference is that the matrix $\mathbf{B}$ is replaced by the expression $\beta \underline{1}^T$, with each expert having the same regression parameters in β. If we stack the matrix defined by $\beta \underline{1}^T$ in the same manner as in (A.5), we obtain a vector $\left(\beta_1 \cdots \beta_q \quad \beta_1 \cdots \beta_q \cdots \beta_1 \cdots \beta_q\right)^T$.

Consider the $(pq \times q)$ matrix $\mathbf{E}$, defined by $\underline{1} \otimes \mathbf{I}$, where $\mathbf{I}$ is a $(q \times q)$ identity matrix. This matrix is a series of p identity matrices stacked on top of each other. If we take $\frac{1}{p}\mathbf{E}^T\beta^*$ under the restricted model, we obtain β, as the matrix $\mathbf{E}$ sums the parameters for each of the experts. Recall that the likelihood of β^* given $\mathbf{\Sigma}$ in (A.5) is in the form of a multivariate normal with mean $\hat{\beta}^*$ and covariance matrix $\mathbf{\Sigma} \otimes \left(\mathbf{X}^T\mathbf{X}\right)^{-1}$. Dawid (1981) states that if z is a multivariate normal of dimension a with mean μ and covariance matrix $\mathbf{\Sigma}$, then if $\mathbf{A}$ is a $(b \times a)$ matrix, then $\mathbf{A}z$ will also be a multivariate normal of dimension b with mean $\mathbf{A}\mu$ and covariance matrix $\mathbf{A\Sigma A}^T$. Thus $\frac{1}{p}\mathbf{E}^T\beta^*$ will be a multivariate normal distribution of dimension q with mean $\frac{1}{p}\mathbf{E}^T\hat{\beta}^*$ and variance $\frac{1}{p^2}\mathbf{E}^T\mathbf{\Sigma}\otimes\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{E}$. The mean evaluates to $\overline{B}=\frac{1}{p}\hat{\mathbf{B}}\underline{1}=\left(\frac{1}{p}\sum_{i=1}^{p}\hat{\beta}_{1,i} \quad \cdots \quad \frac{1}{p}\sum_{i=1}^{p}\hat{\beta}_{q,i}\right)^T$, where $\hat{\mathbf{B}}=\left(\mathbf{X}^\mathbf{T}\mathbf{X}\right)^{\mathbf{-1}}\mathbf{X}^\mathbf{T}\mathbf{Y}$, and the variance evaluates to $\sigma_\Sigma^2\left(\mathbf{X}^T\mathbf{X}\right)^{-1}$, where σ_Σ^2 is the given by $\frac{1}{p^2}\underline{1}^T\mathbf{\Sigma}\underline{1}$.

This arrangement is interesting as $\overline{B}$ is the average across the individual experts of the standard least squares estimates of the parameters of the full multivariate regression model, while σ_Σ^2 is the average across the individual experts of the variance of their residuals and $\left(\mathbf{X}^T\mathbf{X}\right)^{-1}$ represents the covariance between the parameters induced by the design matrix $\mathbf{X}$.

With all these manipulations, we may rewrite the likelihood for the restricted model in (8) as

$$P(\mathbf{Y}\mid\mathbf{X},\beta,\mathbf{\Sigma})\propto|\mathbf{\Sigma}|^{-\frac{N}{2}}\exp\left\{-\tfrac{1}{2}tr\left(\mathbf{V}\mathbf{\Sigma}^{-1}\right)\right\}\exp\left\{-\tfrac{1}{2}\left(\left(\beta-\overline{\mathbf{B}}\right)^T\left(\sigma_\Sigma^2\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\right)^{-1}\left(\beta-\overline{\mathbf{B}}\right)\right)\right\}. \tag{A.6}$$

Appendix B - Prior Distributions

Assume that $\mathbf{\Sigma}$ is a priori distributed according to an inverse Wishart distribution with parameters $\mathbf{G}$ and m, and the single covariate effect vector β conditional on $\mathbf{\Sigma}$ is a priori distributed according to a multivariate normal with mean vector ϕ and covariance matrix $\sigma_\Sigma^2 \mathbf{A}$. Again, ϕ, $\mathbf{A}$, $\mathbf{G}$ and m are arbitrary prior hyperparameters determined by the decision maker.

Thus, the prior distribution on $\mathbf{\Sigma}$ can be written

$$p(\mathbf{\Sigma} \mid \mathbf{G}, m) \propto |\mathbf{\Sigma}|^{-m/2} \exp\left\{-\frac{1}{2} tr\left(\mathbf{G}\mathbf{\Sigma}^{-1}\right)\right\}, \tag{B.1}$$

while the prior distribution of $\underline{\boldsymbol{\beta}}$ given $\mathbf{\Sigma}$ can be written

$$\mathbf{p}(\beta \mid \mathbf{X}, \phi, \mathbf{A}, \mathbf{\Sigma}) \propto \sigma_\Sigma^{-1} \mathbf{exp}\left\{-\frac{\mathbf{1}}{\mathbf{2}\sigma_\Sigma^2} (\beta - \phi)^T \mathbf{A} (\beta - \phi)\right\}. \tag{B.2}$$

Thus the joint prior distribution obtained by multiplying (B.1) by (B.2) is

$$p(\beta, \Sigma \mid \mathbf{X}, \mathbf{A}, \mathbf{G}, \phi, m) \propto |\Sigma|^{-m/2} \exp\left\{-\frac{1}{2} tr\left(\mathbf{G}\Sigma^{-1}\right)\right\} \sigma_\Sigma^{-1} \exp\left\{-\frac{1}{2\sigma_\Sigma^2} (\beta - \phi)^T \mathbf{A}(\beta - \phi)\right\}. \tag{B.3}$$

Multiplying (B.3) by the likelihood in (A.6) gives

$$\begin{aligned} p(\beta, \Sigma \mid \mathbf{Y}, \mathbf{X}, \mathbf{A}, \mathbf{G}, \phi, p, m) \propto |\Sigma|^{-m/2} \exp\left\{-\frac{1}{2} tr\left(\mathbf{G}\Sigma^{-1}\right)\right\} \sigma_\Sigma^{-1} \exp\left\{-\frac{1}{2\sigma_\Sigma^2} (\beta - \phi)^T \mathbf{A}(\beta - \phi)\right\} \times \\ |\Sigma|^{-\frac{N}{2}} \exp\left\{-\frac{1}{2} tr\left(\mathbf{V}\Sigma^{-1}\right)\right\} \exp\left\{-\frac{1}{2\sigma_\Sigma^2} \left(\beta - \overline{B}\right)^T \mathbf{X}^T \mathbf{X}\left(\beta - \overline{B}\right)\right\}. \end{aligned}$$

and combining like terms yields

$$\begin{aligned} p(\beta, \Sigma \mid \mathbf{Y}, \mathbf{X}, \mathbf{A}, \mathbf{G}, \phi, p, m) \propto |\Sigma|^{-(m+N)/2} \exp\left\{-\frac{1}{2} tr\left((\mathbf{G} + \mathbf{V})\Sigma^{-1}\right)\right\} \times \\ \sigma_\Sigma^{-1} \exp\left\{-\frac{1}{2\sigma_\Sigma^2} (\beta - \phi)^T \mathbf{A}(\beta - \phi) - \frac{1}{2\sigma_\Sigma^2} \left(\beta - \overline{B}\right)^T \mathbf{X}^T \mathbf{X}\left(\beta - \overline{B}\right)\right\}. \end{aligned} \tag{B.4}$$

The first term in (B.4) indicates that the posterior distribution of $\mathbf{\Sigma}$ is again an inverse Wishart distribution with posterior parameters $\mathbf{G} + \mathbf{V}$ and $m + q$. The second term in (B.4) needs further simplification, so we consider just the exponent to yield

$$- \frac{1}{2\sigma_\Sigma^2}\left(\left(\beta - \overline{B}\right)^T \mathbf{X^T X}\left(\beta - \overline{B}\right) + \left(\beta - \phi\right)^T \mathbf{A}\left(\beta - \phi\right)\right).$$

Expanding the quadratic forms and grouping like terms yields

$$- \frac{1}{2\sigma_\Sigma^2}\left(\beta^{\mathbf{T}}\left(\mathbf{X^T X} + \mathbf{A}\right)\beta - \beta^{\mathbf{T}}\left(\mathbf{X^T X}\overline{B} + \mathbf{A}\phi\right) - \left(\overline{B}^T \mathbf{X^T X} + \mathbf{A}\phi^T\right)\beta\right) + \text{constant}(\beta, \mathbf{\Sigma}),$$

where $\mathbf{X^T X}$ and $\mathbf{A}$ are symmetric, yielding

$$- \frac{1}{2\sigma_\Sigma^2}\left(\beta^{\mathbf{T}}\left(\mathbf{X^T X} + \mathbf{A}\right)\beta - \beta^{\mathbf{T}}\left(\mathbf{X^T X}\overline{B} + \mathbf{A}\phi\right) - \left(\mathbf{X^T X}\overline{B} + \mathbf{A}\phi\right)^{\mathbf{T}}\beta\right) + \text{constant}(\beta, \mathbf{\Sigma}),$$

Completing the square to form a quadratic form and using the symmetry of $\mathbf{X^T X}$ and $\mathbf{A}$ we obtain

$$-\frac{1}{2\sigma_\Sigma^2}\left\{\left(\beta - \left(\mathbf{X}^T\mathbf{X} + \mathbf{A}\right)^{-1}\left(\mathbf{X}^T\mathbf{X}\overline{B} + \mathbf{A}\phi\right)\right)^T \left(\mathbf{X}^T\mathbf{X} + \mathbf{A}\right)\left(\beta - \left(\mathbf{X}^T\mathbf{X} + \mathbf{A}\right)^{-1}\left(\mathbf{X}^T\mathbf{X}\overline{B} + \mathbf{A}\phi\right)\right)\right\} \tag{B.5}$$

plus term that is constant with respect to β conditional on $\mathbf{\Sigma}$ and so can be dropped while maintaining proportionality. (B.5) implies that β conditional on $\mathbf{\Sigma}$ and the design matrix $\mathbf{X}$ is a posteriori distributed according to a multivariate normal with mean vector $\left(\mathbf{X^T X} + \mathbf{A}\right)^{-1}\left(\mathbf{X^T X}\overline{B} + \mathbf{A}\phi\right)$ and precision $\sigma_\Sigma^2\left(\mathbf{X^T X} + \mathbf{A}\right)$.

References

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

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

Clemen, R. T., R. L. Winkler. 1985. Limits for the precision and value of information from dependent sources. *Operations Research* **33**(2) 427-442.

Clemen R. T. 1986. Calibration and aggregation of probabilities. *Management Science* **32**(3) 312-314.

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.

Cox, D. R. 1972. Regression models and life tables. Journal *of the Royal Statistical Society Ser. B* **34** 187-220.

Dawid, A. P. 1981. Some matrix-variate distribution theory: notational considerations and a Bayesian application. *Biometrika* **68**(1) 265-274.

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

Finn, J. D. 1974. *A general model for multivariate analysis*. New York: Holt, Rinehart & Winston.

French, S. 1980. Updating belief in the light of someone else's opinion. *Journal of the Royal Statistical Society Series A* **143** 43-48.

French, S. 1981. Consensus of opinion. *European Journal of Operations Research* **7** 332-340.

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.

Jouini, M. N. R. T. Clemen. 2002. Copula models for aggregating expert opinion. *Operations Research* **44**(3) 444-457.

Kadane, J. B., J. M. Dickey, R. L. Winkler, W. S. Smith, S. C. Peters. 1980. Interactive elicitation of opinion for a normal linear model. *Journal of the American Statistical Association* **75**(372) 845-854.

Lindley, D. 1983. Reconciliation of probability distributions. *Operations Research* **31** 866-880.

Lindley, D. 1985. Reconciliation of discrete probability distributions. In *Bayesian Statistics 2*, J. Bernado et al. (Eds.), North Holland, Amsterdam 375-390.

Mastrangelo, C.M., G.C. Runger, D.C. Montgomery. 1996. Statistical process monitoring with principal components. *Quality and Reliability Engineering International* **12** 203-210.

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.

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.

Newbold, P., C. W. J. Granger. 1974. Experience with forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistical Society: Series A* **137** 131-149.

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

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

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

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

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

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

Szwed, P., J. Rene van Dorp, J. R. W. Merrick, T. A. Mazzuchi, A. Singh. 2004. A Bayesian paired comparison approach for relative accident probability assessment with covariate information. Accepted by *European Journal of Operations Research*.

Tatsuoka, M. M. 1971. *Multivariate analysis*. New York: Wiley.

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.

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

Winkler, R. L., R. T. Clemen. 1992. Sensitivity of Weights in Combining Forecasts. *Operations Research* **40**(3) 609-614.

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