Bayesian Analysis (2007) **2**, Number 1, pp. 99–136

Geographically Assisted Elicitation of Expert Opinion for Regression Models

Robert Denham* and Kerrie Mengersen†

Abstract. One of the perceived strengths of Bayesian modelling is the ability to include prior information. Although objective or noninformative priors might be preferred in some situations, in many other applications the Bayesian framework offers a real opportunity to formally combine data with information available from experts. The question addressed in this paper is how to elicit this information in a form suitable for prior modelling. Particular attention is paid to geographic data for which maps might be used to assist in the elicitation. Two case studies are used to illustrate the methodology: estimation of city house prices and prediction of presence of a rare species.

Keywords: elicitation, expert opinion, regression

1 Introduction

One perceived strength of the Bayesian approach to data analysis is that it provides a formalised method of combining prior knowledge with data. The specification of the prior knowledge does however pose some problems. One criticism of using prior knowledge in the form of *informative* priors is that subjectivity is introduced into the model. A different prior may produce a different posterior, and hence different interpretation of the results. This is why many workers propose the use of non-informative priors. Here we use the term non-informative to describe any of the priors which do not take into account any partial knowledge that might exist, and so include reference priors (Kass and Wasserman 1996), Jeffreys' priors (Bernardo 1979; Bernardo and Smith 1994) and though usually not strictly non-informative, vague or flat priors. The issue of prior specification becomes more important as the amount of data decrease. Since the posterior can be considered a compromise between the likelihood and the prior, the use of prior knowledge effectively increases the amount of data available. Models using scarce data are also the sort of situation in which the value of prior knowledge becomes more valuable, since a model relying on data alone can sometimes lack the precision to be useful, or may contradict so clearly the source of the data that the model is rejected. For example, the modelling of a rare species' distribution might proceed something like this: a) collect data on presences and absences of the species, and using a set of spatial explanatory variables use logistic regression to produce a map providing the probability of presence across a region of interest; b) show this map to a group of experts, who will then modify the map to be more in line with their knowledge. This seems counter-intuitive at first, as clearly the expert knowledge here takes precedence

*School of Mathematical and Physical Sciences, University of Newcastle, Australia
†School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia, mailto:k.mengersen@qut.edu.au

over the data. Subjectivity in this scenario then enters *after* the statistical modelling has been done, and enters in a non-formal manner, so many of the useful properties of a statistical model are lost. How, for example, does one estimate the uncertainty of a prediction at a given point if an expert has modified the prediction surface? This appears to be an unusual way to conduct an analysis, but in many complex situations these sort of compromises are often made. Gavaris and Ianelli (2002) give a similar example in fisheries data, where some parameters, such as M, the instantaneous rate of mortality are given arbitrary values because of difficulties in estimation due to lack of data. The simplification of models by including arbitrary values for important parameters can introduce bias and make realistic statements about model uncertainty difficult. Other parameters, such as the catchability coefficient q in the fisheries example have so few data to estimate that the estimates are found to be *unreasonable*, again suggesting that the existing knowledge overrides the data. Subjectivity often creeps into statistical models when data is scarce, and rather than trying to ignore it, it would be preferable to include it in a formalised manner. This is possible in the Bayesian approach.

There is another possible reason that informative priors are used less often than expected. This has to do with the difficulty in choosing the prior distributions, and this is especially the case when the priors are chosen or *elicited* based on expert opinion. Experts are rarely accustomed to quantifying their beliefs, and there are a number of psychological stumbling blocks that make the task difficult.

Wolfson (1995) discusses some of the recognised psychological issues that commonly occur in the elicitation of probabilities. Two of the most well known are the heuristic of *availability* (Tversky and Kahneman 1974) and the heuristic of *adjustment and anchoring*. Availability describes the situation in which assessors link their probabilities to the frequency with which they can recall the event. Cooke (1991) provides the example in which the probability of dying from a well publicised disease such as botulism is overestimated, while the probability of death by more common, but less publicised, diseases such as stomach cancer are often underestimated. According to Wolfson, adjustment and anchoring relates to the behaviour in which an assessor anchors their judgment at some starting point and adjusts outwards; often the adjustment is insufficient, and results in a credible interval which contains less probability than was asked. These, and other sources of bias and error in converting an expert's beliefs to probabilistic statements will not be addressed in this paper. Here we wish to concentrate on a setting in which elicitations will take place, and assume that the issues above can be surmounted at least to some extent.

Provided we are satisfied that it is possible to elicit probabilities from an expert, and we are prepared to put the effort into doing the task, we need some structured approach to convert a set of elicited probabilities into prior distributions to suit a Bayesian model. Consider the normal linear regression model in which we wish to use expert knowledge to assist in the modelling of a response Y according to some set of explanatory variables X.

Kadane et al. (1980) and Kadane (1980) distinguish two elicitation approaches for such a situation. The first is predictive, in which the expert is asked to quantify his opinion of the value of the response variable at various fixed values of the explanatory variables. The second approach is termed structural, and refers to procedures that specifically elicit information on parameters. The appropriateness of either of these methods depends on both the problem to be solved and the type of expert providing information. As an example, consider the problem of predicting a species' geographic distribution. This is frequently achieved using a regression based approach such as a generalised linear model (GLM) or a generalized additive model, with species' presences or absences as the response variable, and a suite of environmental variables as the explanatory variables. Of the available GLM techniques it is logistic regression which is used most frequently. These approaches are well documented in the ecological literature; see Manel, Williams, and Ormerod 2001 and Guisan and Zimmermann (2000) for recent reviews. If we were to undertake the statistical modelling of such data from a Bayesian perspective and wish to make use of expert opinion, we might consider two distinct types of experts. The first, which we describe as the *physiologist*, may have a good understanding of the physical requirements of the species. This may be through experimental studies such as germination trials, glass house or laboratory trials. The second, which we term the *field ecologist*, has good knowledge of the places in which the species is likely to be found, though not necessarily expertise in the reasons why. The physiologist is likely to respond well to a structural elicitation procedure, while the field ecologist is likely to prefer a predictive approach. We also believe that the expertise of many will fall somewhere between the two extremes, and elicitation procedures should reflect this.

The value of expert opinion, particularly when data is scarce has also been recognised by the ecology community, although usually not explicitly incorporated into a GLM. When data are scarce or unreliable, models based completely on expert knowledge have been proposed (Store and Kangas 2001; Tamis and Van 't Zelfde 1998; Pearce et al. 2001), although Pearce et al. (2001) also show how a GLM can be modified by using expert opinion. In their work, the incorporation took place during pre-modelling, model-fitting and post-modelling stages of the analysis. The pre-modelling incorporation involved the creation of a new explanatory variable, based on fine-scale vegetation and growth-stage information. The model-fitting stage involved experts choosing an appropriate subset of explanatory variables, and at the post-modelling stage, maps produced from GLMs were modified or refined based on expert opinion. While this is one of the few papers that explicitly discuss and compare the way that expert opinion is used in combination with statistical models, it is likely that most modelling exercises routinely use some combination of the these approaches. These methods don't directly modify the statistical estimation of parameters. The first two approaches merely modify the inputs to the statistical model, and the third approach merely modifies the predictive surface to better fit what the experts believe to be reality. Another way of incorporating expert opinion into a statistical model not considered by Pearce et al. (2001) falls somewhere between their *model fitting* stage and the *post-modelling* stage, in that it directly modifies the model parameter estimates. It requires the GLM to be formulated in the Bayesian context, with a prior provided for the parameter estimates. This is the approach considered in the present paper.

Zellner and Rossi (1991); West (1985); West et al. (1985); Albert (1988); Zeger and Karim (1991); Albert and Chib (1993); Gelfand et al. (1996); Chen et al. (2003, 1999) have all described methods which allow informative priors to be included in a logistic regression, although the emphasis has been on computational methods rather than the prior specification itself. Others (Chen and Dey 2003) have shown how historical data can be used for analysis and variable selection for a current study. The use of expert opinion in logistic regression doesn't appear to have been documented until Bedrick et al. (1996, 1997) who describe a method in which informative priors for logistic regression can be specified and elicited, and illustrate how an expert might provide sufficient information to specify the prior. The conditional means prior (CMP) and data augmentation prior (DAP) of Bedrick et al. (1996) share many things in common with some of the earlier regression elicitation work (Kadane et al. 1980; Garthwaite and Dickey 1988) in that they recognise the difficulties in eliciting directly a prior for β. To address this difficulty, it is generally suggested that the elicitation procedure ask the expert questions about the response for given values of the explanatory variables, and convert these into a prior for β. The CMP approach requires an expert to provide their estimate of the probability of a success $\tilde{p}_i$ at $i = 1 \dots k$ carefully chosen points in explanatory variable space $\tilde{x}_i$. Bedrick et al. (1997) show how this can be applied to a logistic regression model in which the response is the survival after a injury, with explanatory variables injury severity score, revised trauma score, age and predominant type of injury.

This approach can be adapted to questions of species distribution modelling. Taking one recent example of species distribution (Gibson et al. 2004), if expert opinion were to be used in the CMP framework, an expert would be asked to provide their estimate of the probability of presence of the rufous bristlebird at six combinations of *elevation*, *distance to creek*, *distance to coast*, *sun index* and *habitat complexity*. Again, the ability to answer these questions may depend to a large extent on the type of expert we are questioning. Other serious concerns are that some variables, particularly indices, may become more difficult to understand as the complexity of the model increases, and that the method is perhaps testing the expert's knowledge of the distribution of the explanatory variables, rather than what the expert is an expert in.

This paper explores flexible elicitation procedures which make use of both structural and predictive processes, and ways the elicitation procedure can be improved in geographic settings. The method is designed to acknowledge different types of experts such as *physiologist* and the *field ecologist* and to provide appropriate resources to assist them in better translating their expertise into a formal prior. In Section 2 we briefly examine standard elicitation approaches for regression problems, then describe how elicitation could be approached in a geographic setting. A more flexible option can be achieved by combining the predictive method with a structural method, and this is described in Section 3.

Kadane and Wolfson (1998) make the important observation that *the goal of elicitation ... is to make it as easy as possible for subject-matter experts to tell us what they believe, in probabilistic terms, while reducing how much they need to know about probability theory to do so.* With this in mind, careful consideration must be given to the tools used to elicit expert opinion. In almost all elicitation applications, computer software is used since this is the simplest way to provide feedback to the expert on their choices and allow visualisation of responses while simultaneously handling any necessary computation. Visualisation of the geographic nature of the data and of the elicited expert information is an integral part of the methodology proposed in this paper. To further facilitate ease of use, the software developed for the case studies described below built on existing discipline-specific software familiar to the experts and linked to third party statistical and graphical software.

Craig et al. (1998) provide an example of an elicitation project which is spatial in nature, and wrote software that included an interactive map which was used to select points to gather expert opinion, much like our approach, although their elicitation procedure is quite different.

Two experiments were conducted to test the proposed methodology. The first case study, modelling the median house prices on a suburb basis for the city of Brisbane, Australia was used to evaluate the various approaches as described in Section 4. The second case study, described in Section 5 extends this evaluation to the more interesting problem of predicting the distribution of a rare species, the Australian brush-tailed rock-wallaby. A discussion completes the paper.

2 Geographically assisted expert elicitation

Consider first the formulation for the Bayesian linear regression model given by Kadane et al. (1980),

$$\begin{aligned} Y|X,\beta,\sigma^2 &\sim N(X\beta,\sigma^2), \\ \beta|\sigma^2 &\sim N(b,\sigma^2 R^{-1}), \\ \tfrac{1}{\sigma^2} &\sim \tfrac{\chi^2_\delta}{w\delta}, \end{aligned}$$

where Y is a vector of observations and X a matrix of explanatory variables. The distribution of Y then is multivariate normal with mean $X\beta$ and variance σ^2. The model requires prior distributions for β and σ. Kadane et al. chose a conjugate prior structure so that β is multivariate normal and $1/\sigma^2$ is a scaled χ^2. The hyperparameters which need to be estimated in the elicitation procedure are b, δ, R and w.

To estimate the hyperparameter b, the expert is asked to provide their median value $y_{i_{0.5}}$ at each of x_i, $i = 1,\ldots,m$, and b is estimated using weighted least squares, with weights

$$v_i = \frac{y_{i_{0.9}} - y_{i_{0.5}}}{y_{i_{0.5}}}.$$

with $y_{i_{0.5}}$, $y_{i_{0.9}}$ the expert's median and 90th percentile at site x_i.

The b parameter is then estimated by

$$\hat{b} = (X^T Q_y^{-1} X)^{-1} X^T Q_y^{-1} y_{0.5}$$

where $y_{0.5}$ is the vector of elicited median values and $Q_y^{-1} = \mathrm{diag}(v_1,\ldots,v_m)$. We use weighted least squares rather than the unweighted least squares as in Kadane et al., since it seems difficult to accept that the elicited errors would have equal variances. Two points far apart geographically may in fact be close together in variable space, and we want to ensure that the point with which the expert is more familiar carries more weight than the other.

The degrees of freedom parameter, δ, can be estimated taking

$$a(x_i) = \max\left(\frac{y_{i_{0.90}} - y_{i_{0.50}}}{y_{i_{0.75}} - y_{i_{0.50}}}, \frac{T_\infty(.90)}{T_\infty(.75)}\right),$$

where $T_\delta(p)$ is the pth quantile from the standard t distribution with δ degrees of freedom, y_{i_p} is the expert's p'th quantile at x_i.

The hyperparameter δ is then chosen to satisfy

$$\frac{T_\delta(0.90)}{T_\delta(0.75)} = \frac{\sum_{i=1}^{m} a(x_i)}{m}$$

The estimation of w and R is a little more complicated, and follows what Kadane et al. describe as a *conditional elicitation procedure*. This requires further elicitation for a subset of the values already provided. From the set of elicited values, $y_1, \cdots, y_m$, we choose m' of these, with $p+2 \le m' \le m$ to perform further elicitations. The additional elicitations are as follows:

```
for i in 1,...m'-1 {
        generate y_i^0
        conditional on y_1^0 ··· y_i^0 expert provides y_{i+1_{0.5}}, y_{i+1_{0.75}}
        conditional on y_1^0 ··· y_i^0 expert provides y_{i_{0.5}}, y_{i+2_{0.5}} ··· y_{m'_{0.5}}
    }
for i = m' {
        generate y_i^0
        conditional on y_1^0 ··· y_i^0 expert provides y_{i+1_{0.5}}, y_{i_{0.5}}
}
```

Kadane and Wolfson (1998) describe use of the conditional assessments to construct a symmetric positive definite matrix U, where the ith diagonal element is the spread of y_i, given by

$$S(y_i|x_i) = \frac{(y_{i_{0.75}} - y_{i_{0.50}})^2}{T_\delta^2(0.75)}, \tag{1}$$

and the remaining $i-1$ elements of the ith row (and thus also of the ith column) are the co-spreads of y_i with $(y_i, \ldots, y_{i-1})$. As explained by Kadane and Wolfson, w_i can be obtained as an estimate of w using the fact that at each stage of the construction of this matrix, conditionally on $y_1^0, \ldots, y_i^0, y_{i+1}$ and y_{i+1}^* have a joint bivariate t-distribution.

Thus there will be m' estimates of w, which can be averaged to obtain the final estimate. Then the relationship that $U - wI$ is an elicited version of the matrix

$$S(X^T\beta|X) = \frac{w}{n}X^{-1}X^T, \tag{2}$$

where $n = \delta + m$ can be exploited to obtain the estimate of R^{-1}.

In the context of species modelling, y_i is taken to be the number of observations of a species at site i, X is the $n \times k$ matrix of explanatory variables and $\eta = X\beta$ is the vector of linear predictor values. Under the logistic regression model $y_i \sim \mathrm{Bin}(n_i, \mu_i)$. With the logit as the link function, $g(\mu_i) = \log(\mu_i/(1-\mu_i))$, with a prior for β ideally given by

$$p(\beta) = \mathrm{MVN}(b, \Sigma).$$

It is through this prior that we can include expert opinion.

The methods of Kadane et al (1980), Kadane and Wolfson (1996), Wolfson (1995), and Garthwaite and Dickey (1988, 1992), are well suited to many applications. However, when applying such a technique to predictions of a spatial nature, there could be some loss in translation from the expert to a prior. Applications of this nature could include the modelling of disease rates, the logit of a species presence, or the house prices of a suburb. Under the standard elicitation procedures, to estimate the hyperparameter b, an expert would be asked to provide an estimate, perhaps including some quantile information, for the value of y at particular values of x.

This task can be difficult if the expert's knowledge is more related to their knowledge of *locational* values of y, rather than the relationship between y and the explanatory variables. The expert might, for example, quantify their beliefs by first trying to imagine the locations that correspond to x, and then providing a typical value of y at these points. The expert is effectively being asked about their knowledge of the distribution of x across the landscape, a task which becomes increasingly difficult as the number of explanatory variables increases. It seems that in many cases it would be more efficient, and more consistent for the expert, if they were asked to provide information on their knowledge of the geographic distribution of y directly.

To provide this information, an expert can be asked to map their estimates of the response across the landscape. The detail of the map may vary from a complete coverage of the area of interest, to a number of individual points with accompanying estimates. The complete map is likely to be time consuming to achieve, and unless there is a specific need for complete coverage (one possible scenario in which this would be useful is discussed below), this would rarely be attempted. More common would be a map consisting of areas of approximately equal response.

This essentially is equivalent to Kadane et. al. (1980), except that instead of choosing a systematic set of design points, the design points are chosen by the expert based on location, without specific reference to any explanatory variables. This allows the expert to provide information on those areas he is most familiar with, although to achieve sufficient coverage of the variable space the expert may be required to include information in areas with which he is less familiar. This increased uncertainty should be reflected in the expert's quantiles for y. As an alternative, design points could be chosen in the same manner as in Kadane, and displayed on the map. The expert is then asked to provide quantiles for a number of these design points. Many more design points can be provided than is required for the elicitation, giving the expert the opportunity to concentrate on those areas in which he is most familiar.

There are a number of ways in which the expert might be able to provide the elicitations. The simplest would be to annotate a hard copy map of the area of interest. This does however limit the expert to reporting on values at a fixed scale. For example, if aspect (the direction in which a slope faces) is found to be an important predictor of a species presence/absence, it would be difficult for an expert to select sites with a given aspect using a map with a scale of say 1:25 000. Variables which are not discernible at the scale of the experts map would need to be given a non–informative prior, or elicited separately.

A more flexible procedure would make use of a Geographic Information System (GIS). The two case studies described below provide examples of how a GIS can be used to aid in the elicitation procedure. A GIS approach allows the expert to access information at any point in a convenient manner. For example, the expert can have several layers of data available at one time, each providing information of a different feature (rainfall+temperature). The user can also build queries to determine relationships between points and nearby features (distance to water). The user can also zoom in or out, removing the scale dependency of the hard copy maps. Taking the example of aspect again, if an expert believes a species is likely to be present in a particular area, but only on sites with a given aspect, they can zoom in until aspect becomes discernible and choose a location, or alternatively could select all sites within the area with suitable aspect. The interactivity of the GIS also means that the expert can easily revisit and modify any elicitations they have made.

3 Combination Approach

3.1 Methodology

Predictive elicitation procedures are often preferred over structural approaches since the expert need not understand the statistical model and is only asked about observable quantities. There are, however, situations in which a structural approach may be appropriate. In the experience of Kadane and Wolfson (1998), for example, economists are used to thinking in terms of parameters, and respond well to a structural procedure. Garthwaite (1998, Queensland Department of Natural Resources report) presented a structural elicitation procedure for habitat distribution models. While this problem is spatial in nature, it does not explicitly take this into account in the elicitation procedure, except perhaps providing some information on the distribution of explanatory variables in the form of paper maps. An expert who is asked for information on the relationship between a species presence and average annual rainfall, for example, would be provided with a map of rainfall over the region of interest.

However, an elicitation procedure need not exclusively be predictive or structural. Kadane and Wolfson (1998) present an example which they describe as a hybrid approach. In this case, some variables were elicited in the predictive manner, and others in the structural manner. Our proposal can also be considered a hybrid approach, but differs from that of Kadane and Wolfson in that it offers the opportunity to use either method *simultaneously* for a single variable. It is the expert's preference which determines which is used more.

The basis for this procedure is the combination of the map prior of Section 2 with a structural elicitation procedure. To illustrate, consider initially the procedure for eliciting the b parameters in a linear regression. The expert uses the map elicitation procedure to derive a first pass estimate of b. Univariate graphs for each of the p variables are presented. In each, we fix each of the remaining $p-1$ variables at some value, such as the mean or median. That is, for variable j, $j = 1, \ldots p$, we display the graph of

$$y = b_0 + b_j X_j + \sum_{k=1, k \neq j}^{p} b_k \bar{X}_k \ .$$

As the expert updates the map, by either adding points or editing values, the univariate graphs are automatically updated. Additionally, the user is able to manipulate the graphs, which in turn will update the map. The user continues adding points and adjusting values until a consensus between the map and the univariate graphs is reached.

The procedure for estimating the hyperparameters R and w could continue as described in Section 2, although under the combined approach it may be the case for some experts to prefer to provide a more direct prior for β, such that

$$\beta \sim N(b_0, \Sigma_b)$$

The expert, for example, can provide a 95% *envelope* around the displayed regression lines. This information can be used to determine Σ_b. For each variable X_p, the expert is asked to provide upper and lower 95% quantiles for $\hat{y}$ at either end of the range of X_p. These points create a quadrilateral region, which describes the distribution of β. The derivation of Σ_b is as follows:

If

$$\hat{y}_i \sim N(bX_i^*, \sigma^2_{\hat{y}_i})$$

where X^* is the matrix of the end points . For the simple case with two explanatory variables ($P = 3$), then

$$X^* = \begin{pmatrix} 1 & X_{11} & \bar{X}_2 \\ 1 & X_{12} & \bar{X}_2 \\ 1 & \bar{X}_1 & X_{21} \\ 1 & \bar{X}_1 & X_{22} \end{pmatrix}$$

then $b = Q\hat{y}$, where $Q = ((X^*)^T X^*)^{-1}(X^*)^T$ and $\mathrm{Var}(b_i) = \sum_{k=1} q_{ik}^2 \sigma_{\hat{y}_k}^2$ since $\mathrm{Cov}(y_i, y_j) = 0$, and $\mathrm{Cov}(b_i, b_j) = \frac{1}{2}\left(\sum_{k=1}(q_{ik} + q_{jk})^2 \sigma_{\hat{y}_k}^2 - \mathrm{Var}(b_i) - \mathrm{Var}(b_j)\right)$

To clarify this, in the two variable case, if

$$Q = \begin{pmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ q_{31} & q_{32} & q_{33} & q_{34} \end{pmatrix}$$

then

$$\begin{aligned} b_0 &= q_{11}\hat{y}_1 + q_{12}\hat{y}_2 + q_{13}\hat{y}_3 + q_{14}\hat{y}_4 \\ \mathrm{Var}(b_0) &= q_{11}^2\sigma_{\hat{y}_1}^2 + q_{12}^2\sigma_{\hat{y}_2}^2 + q_{13}^2\sigma_{\hat{y}_3}^2 + q_{14}^2\sigma_{\hat{y}_4}^2 \end{aligned}$$

and similarly for b_1 and b_2. In addition,

$$b_0 + b_1 = (q_{11} + q_{21})\hat{y}_1 + (q_{12} + q_{22})\hat{y}_2 + (q_{13} + q_{23})\hat{y}_3 + (q_{14} + q_{24})\hat{y}_4$$

so

$$\begin{aligned} &\mathrm{Var}(b_0) + \mathrm{Var}(b_1) + 2\mathrm{Cov}(b_0, b_1) = \\ &\quad (q_{11} + q_{21})^2\sigma_{\hat{y}_1}^2 + (q_{12} + q_{22})^2\sigma_{\hat{y}_2}^2 + (q_{13} + q_{23})^2\sigma_{\hat{y}_3}^2 + (q_{14} + q_{24})^2\sigma_{\hat{y}_4}^2 \end{aligned}$$

3.2 Software Development

For the specific case of species' distribution modelling, we needed software that allowed the expert to explore the statistical relationship between the species and a number of environmental variables. Initial prototypes were constructed using the R language (R Development Core Team 2004). It quickly became apparent that this approach was limited since the ability to display and query the spatial data was a crucial component of the elicitation exercise. Rather than write functions replicating the functionality of a geographic information system (GIS) into a statistical package, we integrated statistical methodology into an existing GIS.

Our software thus has three components: a GIS, statistical routines and graphing functions. This combination means the expert is free to explore the spatial data using all the functionality of the GIS and to answer statistical questions with the assistance of interactive graphs. The statistics are largely hidden from the expert. The construction of these three components is briefly discussed below in the context of the species' distribution case study, but the concepts are transferable to any generalised linear modelling situation.

Many ecologists are familiar with GIS software, and this often forms part of their standard research toolbox. This familiarity means that the expert is not confronted with a new and unfamiliar package in the elicitation procedure. A GIS in its simplest sense, is a software package that allows the user to interactively represent and query spatial data. From our point of view this is important since while we may be interested in gaining prior information on a number of explanatory variables, say X_1, X_2, X_3, we believe that the expert is best able to provide their prior on the variable by making use of extra, contextual information. This may mean simply allowing the expert to see the spatial distribution of X_1, but it may involve more complex displays such as overlaying towns, roads or reserves.

An example situation using the Brush-Tailed Rock Wallaby elicitation example (See Section 5) might begin by showing the region of interest with biogeographical regions (see below), then labelling major towns in the area, national parks and other reserves, a digital terrain model (DTM) as well as showing the variables of interest to the statistical model such as slope, aspect and geology; examples screenshots are given in Figures 1, 2 and 3.

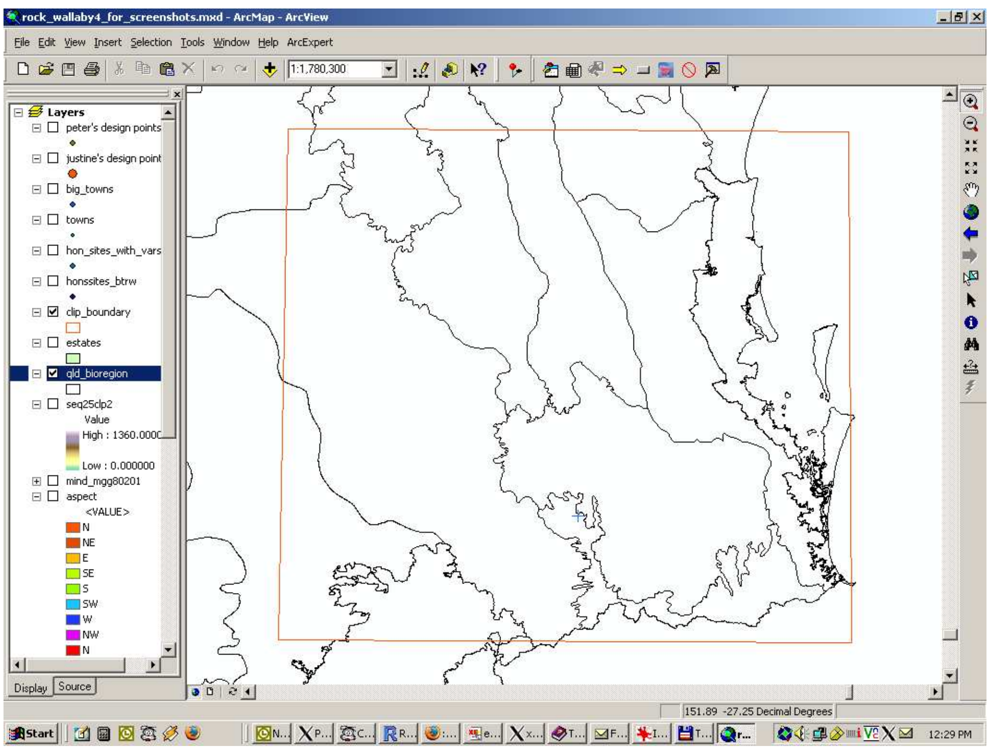

Figure 1: GIS showing region of interest (delineated in red) with biogeographic regions.

Additional functions were written to provide simple buttons and menus on the ArcGIS interface for the expert to access the elicitation functions (Figure 4).

ArcGIS has only limited statistical functionality, but it is possible to customise its

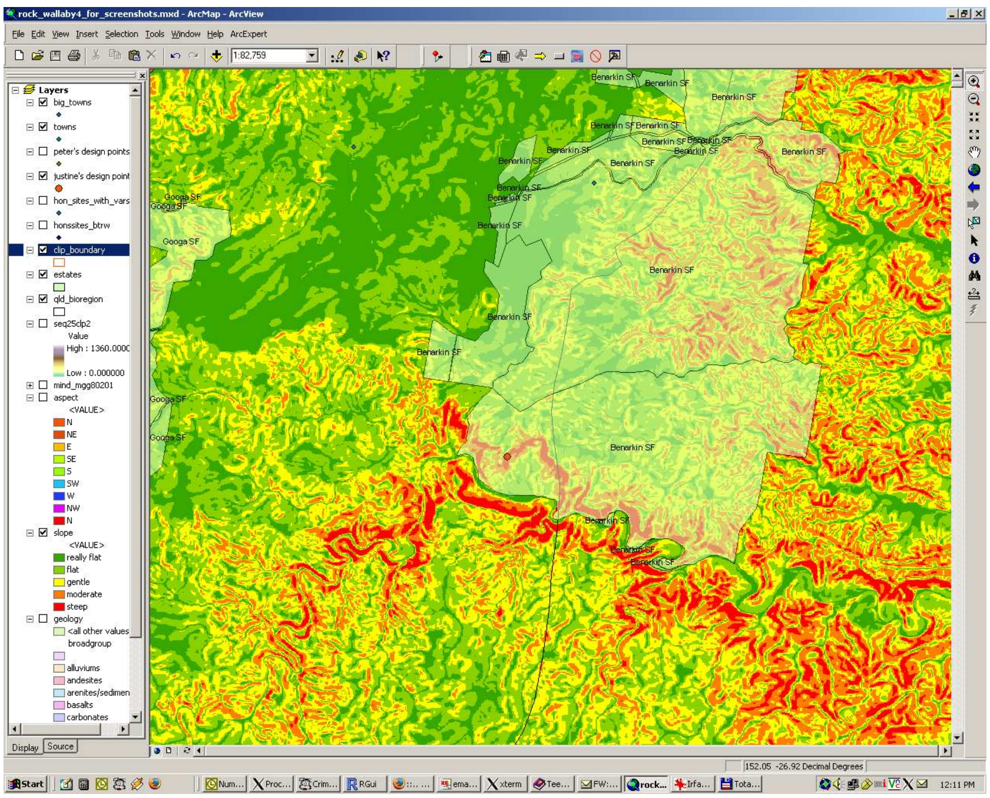

Figure 2: GIS showing the zoomed in section of the region of interest with slope

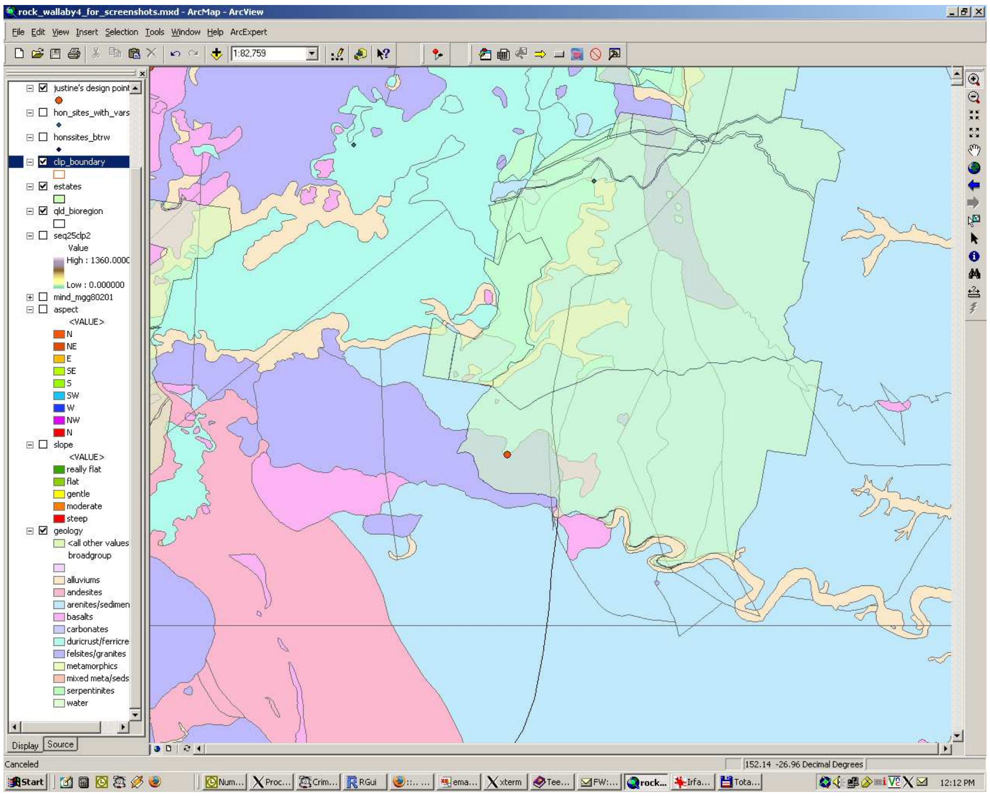

Figure 3: GIS showing the zoomed in section of the region of interest with geology

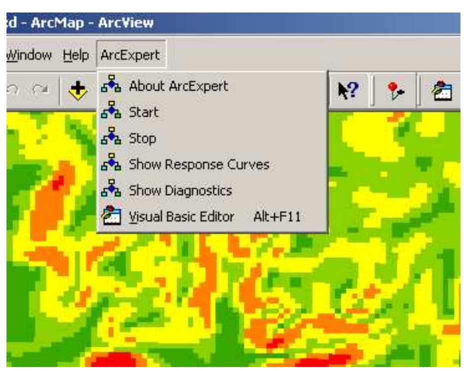

Figure 4: Menu for elicitation functions and point locator button (red push-pin icon) as seen on the ArcGIS interface

environment by scripting with Visual Basic for Applications (VBA), and by communicating with other applications written for Windows. Statistical calculations required for the estimation of a number of parameters was achieved by linking ArcGIS to R using the R(D)Com package. Some examples of the specialist statistical routines constructed in VBA and the functions used to call and use R are given in the Appendix.

The graphing requirements of the software were facilitating by sending the results of the statistical analysis to a number of interactive graphing functions using the TeeChart ActiveX control. These were displayed as forms in ArcGIS and were used in the first stage to allow an interactive selection of hyperparameters for a Beta distribution and at the second stage to display univariate interpretations of the expert's prior as described above. An example of the code used to convert expert elicited data to graphable elements is given in the Appendix.

4 Case Study 1: Estimating House Prices

Brisbane, with a population of 1.6 million people, is Australia's third largest city. It is a city experiencing considerable economic growth, and this is reflected in increases in median property prices throughout the city. This growth has led to a widespread interest in house prices amongst many residents, and for the purposes of this example we consider anybody who has recently bought or sold a house, or who is looking to do so as an *expert*. A general trend in Brisbane, as with many cities, is that suburbs close to the city centre tend to have higher median house prices than suburbs further out. We thus use distance to city centre as the first explanatory variable to predict median house price. The Brisbane River flows through the city, and out to Moreton Bay on the cities eastern side. There is some suggestion that proximity to either the river or the bay also affects house price, so this formed our second explanatory variable.

There is no doubt that there are a number of other variables which determine a suburb's median house price, such as topography, access to public transport and historical features. These variables were not considered in the model, for a number of reasons. First, this was an exercise in elicitation, not an attempt to find the best model for house price prediction. Secondly, most regression models for physical systems, such as species distribution models, have some explanatory variables that are difficult to capture. An important variable determining the presence or absence of many plant species for example relates to soil moisture content, a value difficult to determine on the regional scales necessary for distribution models. We are often restricted to the use of available explanatory variables, even when we are aware that there are other, important variables missing from the model.

To achieve flexibility while maintaining the simplicity of linear regression, we have chosen to use piecewise–linear regression. Knots were chosen to be at the 0.33 and 0.66 quantiles for each of the explanatory variables, though it would also have been an option to include the knots as additional parameters in the model.

The formulation of the model then is

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x'_{i1} + \beta_3 x''_{i1} + \beta_4 x_{i2} + \beta_5 x'_{i2} + \beta_6 x''_{i2}$$

Where X_1 represents distance from city centre in kilometers and X_2 represents distance from the Brisbane River/Moreton Bay in kilometres, and

$$x'_{ij} = \begin{cases} x_{ij} - x_{0.33j} & \text{if} \quad x_{ij} > x_{0.33j} \\ 0 & \text{otherwise} \end{cases}$$

and

$$x''_{ij} = \begin{cases} x_{ij} - x_{0.66j} & \text{if} \quad x_{ij} > x_{0.66j} \\ 0 & \text{otherwise} \end{cases}$$

We also choose to simplify the Bayesian representation of this model when compared with either Kadane et al. (1980) or Garthwaite and Dickey (1992). Specifically, we use an inverse gamma prior for σ, requiring the estimation of two hyperparameters ν_0 and S_0, and we specify a prior for the regression parameters β as described in Section 3, giving

$$\begin{aligned} Y|X, \beta, \sigma^2 &\sim N(X^{\mathsf{T}}\beta, \sigma^2), \\ \beta &\sim N(b, \Sigma_b), \\ \sigma^2 &\sim \mathrm{IG}(\nu_0/2, \nu_0 S_0/2). \end{aligned} \tag{3}$$

Volunteers were broken into two groups. The first used the map elicitation procedure first, followed by the combined approach. The second group used the standard elicitation procedure without geographical assistance, followed by the combined approach. For the standard elicitation procedure, a simplified approach based on Kadane et al. (1980) was used. In this case, design points were chosen to be the four knots for each variable, with the possible option of a further eight design points. The typical way to choose the design points is to find the minimum and maximum values for each of the variables and divide this into say 4 equal parts. Then there will be p^4 design points. This doesn't always work very well, since not all combinations will make sense. For example, in Brisbane since the river flows through the city centre, there are no locations which are both close to the city and far from the water. Design points were instead chosen to cover the explanatory variable space.

The software developed for the combined elicitation approach for this case study is illustrated in the screenshot in Figure 5. The bottom right-hand graph depicts the suburbs of Brisbane and the river, with one suburb, Ashgrove, highlighted. The expert's median house price and limits are displayed above the graph and a corresponding slide bar allows these to be altered easily. To the left of the plot the visual realisation of the expert's information is given in the form of the relationship between house price and distance to city centre and distance to water. Options to 'Dump to R', 'Reset' or 'Quit' are also given.

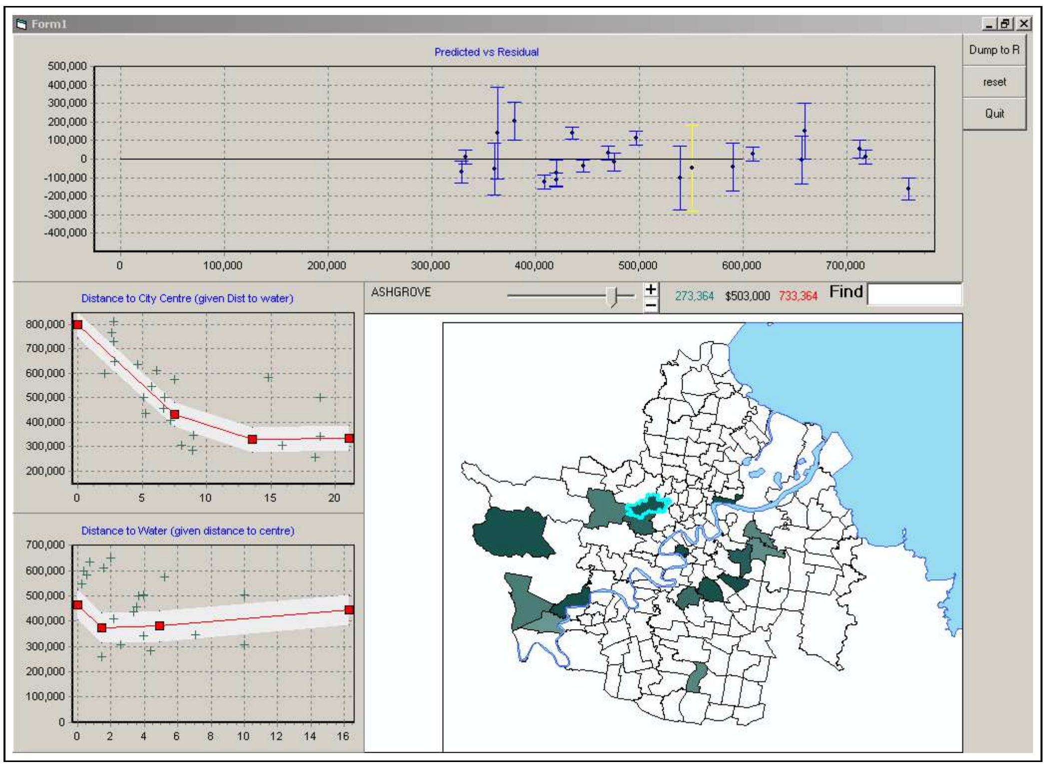

Figure 5: Screenshot of Elicitation Program

	Standard			Map		
	annie	sam	brendan	mark	shane	robert
b_0	659312	354691	459242	926968	520043	867269
b_1	-29891	-4446	3644	-24027	-19240	-49349
b_2	10178	12081	-18670	-17212	8939	43272
b_3	29476	-25357	2366	11322	21008	411
b_4	-113534	-44714	-61887	-57078	-65588	-35217
b_5	99456	41435	33292	60914	55118	17284
b_6	19906	3219	33035	-24665	13184	32458

Table 1: Elicitation results for b (see Equation 3) using the combination approach. The first three experts used a standard elicitation approach, followed by the combined approach, the last three experts used a map only approach followed by the combined approach.

The results from the experts using the combined approach are given in Table 1. For this experiment we had observed median house prices for each suburb in the Brisbane city area, and comparisons between the elicited and the actual values can be made. Figure 6 and Figure 7 shows this comparison.

These results suggest that in general the experts were able to provide quantifications of their beliefs of the distribution of house prices accross Brisbane which were consistent with actual house prices. Each session with the experts was conducted individually, but the priors appear to be relatively consistent. Only one of the experts provided a prior markedly different to the actual relationship between house prices and the two explanatory variables. This, of course, doesn't necessarily imply that the procedure failed to accurately elicit that expert's opinion. It is in fact difficult to assess the quality of an elicitation procedure, since differences between procedures may merely reflect differences in opinions between experts, rather than a differences due to the methodology. We also have no measurable way of determining how close the quantification of opinion matches an expert's true beliefs. Some work has been done in the area of validation of elicitation procedures. Kadane and Wolfson (1998) in their survey of the literature cite relevant measurements of reliability (Wallsten and Budescu 1982), coherence (Lindley et al. 1979) and calibration (Morgan and Henrion 1990). All participants, however, reported that they preferred the combined approach over the map or traditional approach. This, coupled with the agreement the experts priors had with the actual relationship provides some support for the method.

The experiences of the experts in providing information under these experimental conditions were documented and compared. In both groups, the exploitation of the geographic nature of the data in the elicitation process was unanimously preferred by the experts. The overall indication was that elicitation without distributional information was at least as good as the elicitation without geographical assistance but not as good as the combination approach. Moreover, experts reported that the combination approach reduced the time required to provide quantitative information with which they were satisfied, due in particular to the ease of submission of the information and the available

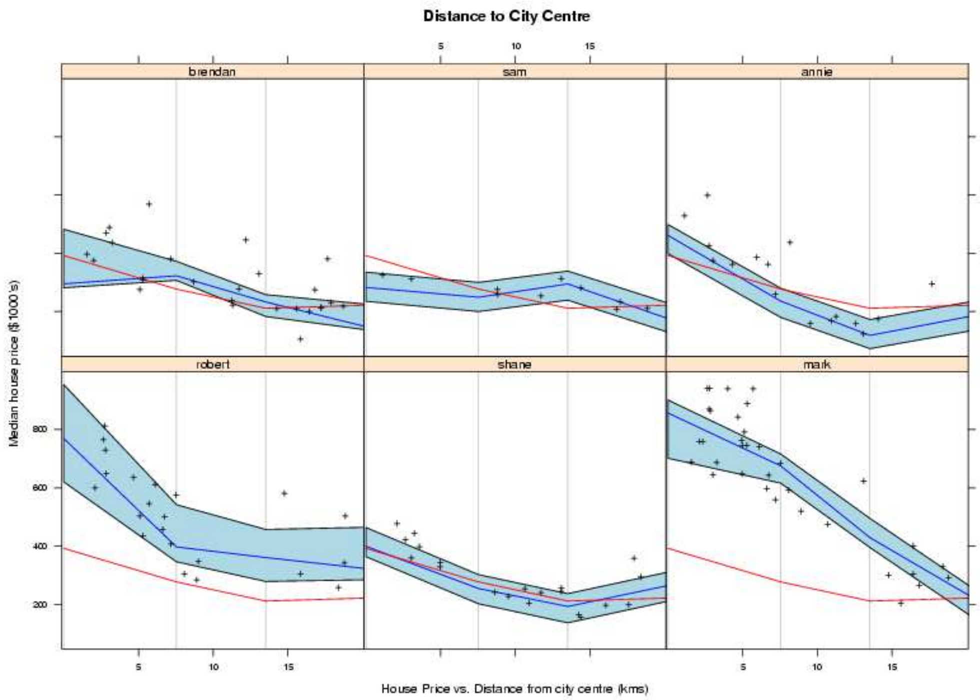

Figure 6: Elicited relationship between distance from city centre and house price. Experts on the top row used a standard elicitation procedure followed by the combined approach. Experts on the bottom row used a map only procedure followed by the combined approach. The blue line represents the expert's mean estimate ($\hat{b}$) and the light blue region represents the expert's 95% envelope for the prior distribution of β. The red line is the estimated relationship using the known median house price of all Brisbane suburbs. Crosses represent the expert's selected design points.

.

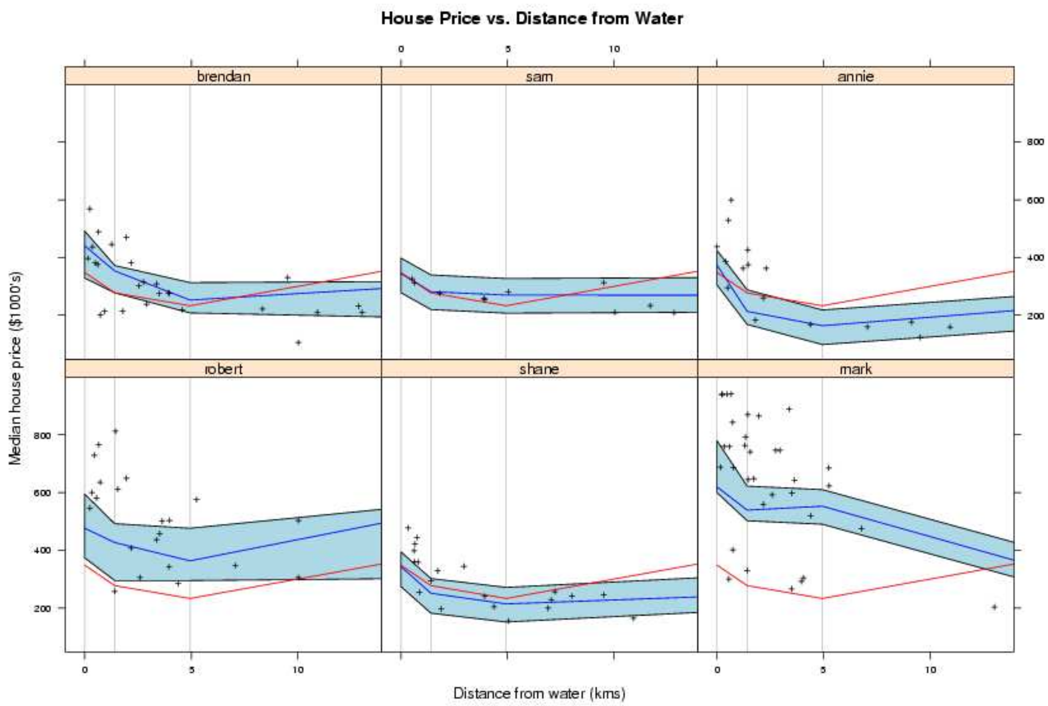

Figure 7: Elicited relationship between distance from water (bay or river) and house price. Experts on the top row used a standard elicitation procedure followed by the combined approach. Experts on the bottom row used a map only procedure followed by the combined approach. The blue line represents the expert's mean estimate ($\hat{b}$) and the light blue region represents the expert's 95% envelope for the prior distribution of β. The red line is the estimated relationship using the known median house price of all Brisbane suburbs. Crosses represent the expert's selected design points.

graphical feedback.

Most experts provided slightly different priors under the different elicitation methods to which they were exposed. Almost all expressed greater satisfaction with the results of the graphical approach, but it was difficult to determine whether this was because of the accuracy of the result or the quality of the process of elicitation. This is further addressed in the Discussion below. Graphical representation of priors obtained under the combination approach are shown as part of the next case study.

5 The Brush-Tailed Rock Wallaby

The brush-tailed rock-wallaby (*Petrogale penicillata*) is a medium sized wallaby, about 1.2m in total length. The Brush-tailed Rock-wallaby has a coastal to sub-coastal distribution, ranging from just north of Brisbane to western Victoria (Maynes and Sharman 1983). Its range has declined substantially, particularly in the west and south. There is still much to be learnt about the ecology and habitat requirements of the species, including habitat requirements (Eldridge 1997). Distribution mapping for this species will add to our knowledge of this species, and ultimately assist in its conservation and management (Carter and Goldizen 2003).

Predicting the distribution of this species with the assistance of expert opinion was the subject of a workshop organised at the Queensland University of Technology. Two experts were recruited. Peter Jarmin (Expert P) is a highly respected researcher with many years experience in macropods but no local knowledge of the species of interest and no experience in GIS. In contrast, Justine Murray (Expert J) has excellent recent, local knowledge about the species of interest and good GIS skills. We can consider expert J as the *field ecologist* and expert P the *physiologist*.

We chose a large area in south-east Queensland (Figure 8) because it is known to contain populations of the species of interest and is large enough to provide a reasonable coverage of environmental variables but small enough that useful information could be obtained from the experts in the limited time available in the workshop.

Expert J made suggestions as to which explanatory variables may be appropriate. In her view, *aspect* was the most important explanatory variable, although *geology* and terrain are also likely to be important. We chose *slope* as a measure of terrain. An additional variable, annual mean moisture index (*mind*), was also chosen to represent a variable which would be more difficult to elicit information on. Moisture index is an indication of the moisture in the soil, and is a function of the precipitation, evaporation and soil type. The index ranges from 0 (dry) to 1.0 (saturated). This parameter was included to allow us to see if useful prior information could be elicited even when an understanding of the variable is limited.

The variables *slope* and *mind* included quadratic terms in our model. The *geology* variable was classified into 11 broad groups based on dominant rock type with the assistance of a geologist. These 11 broad groups were later reduced to four groups, based on the geology types for which the experts provided information and the groups that were featured in the observed data.

The model in which we are interested is thus

$$
\begin{aligned}
y_i &\sim \text{Bernoulli}(p_i) \\
\text{logit}(p_i) &\sim \text{N}(\mu_i, \sigma^2) \\
\mu &= \beta_0 + \beta_1 x_{slope,i} + \beta_2 x_{mind,i} + \beta_3 x^2_{mind,i} + \\
& \quad \beta_4 x_{aspect,i} + \beta_5 x^2_{aspect,i} \\
& \quad \beta_6 x_{geol1,i} + \beta_7 x_{geol2,i} + \beta_8 x_{geol3,i} \\
\beta &\sim \text{MVN}(b, \Sigma)
\end{aligned}
$$

where the subscript i represents the ith observation. Expert J provided a dataset, of which we used a subset of 38 observations. We wish to elicit from the expert their prior for β, that is a multivariate normal distribution with mean vector b and variance-covariance matrix Σ.

Each expert was questioned separately. During each session, the procedure was explained to the expert. Although each expert was reasonably well versed with concepts of probability and statistics, a short discussion on the concepts of medians and quantiles was also included here in an attempt to calibrate the user's quantification approach. We used the double lottery system to help the expert quantify quantiles. We were specifically interested in four explanatory variables (*slope*, *mind*, *aspect* and *geology*), although a number of additional layers were available in the GIS to help the expert obtain contextual information when providing their estimates. These included layers to help the expert locate themselves, such as towns, national parks and state forests, and a 25m digital elevation model to help the expert visualise the terrain. Both the broad group and the detailed description was available to the expert.

We allowed the expert to choose the design points rather than choosing them in advance. Experience has shown us that this is preferable since when working with rare or uncommon species, the vast majority of sites will have a low probability of presence. Even a carefully preselected set of design points often provide a poor range in the probability of presence. We have essentially used expert opinion in selection of the design points by asking the experts to choose sites that would cover the range of probabilities of species occurrence, from highly unlikely to be present, to highly likely to be present. Further restrictions were also constructed to ensure that a reasonable coverage of the study area was created, and that the no design point was allowed to be located in an area in which the expert had undertaken field work.

The expert chooses a design point or a *virtual field site* by clicking on a point in the map, upon which an interactive dialog pops up (Figure 9). This dialog includes the plot of a beta distribution with three adjustable points located at the median and at the 0.05 and 0.95 quantiles. The expert is asked to provide their best estimate of the probability of presence at that site by clicking on the median point, and sliding it left or right until its position matches their belief. They were then questioned about the possible range of values of the probability of presence. To help estimate the quantiles,

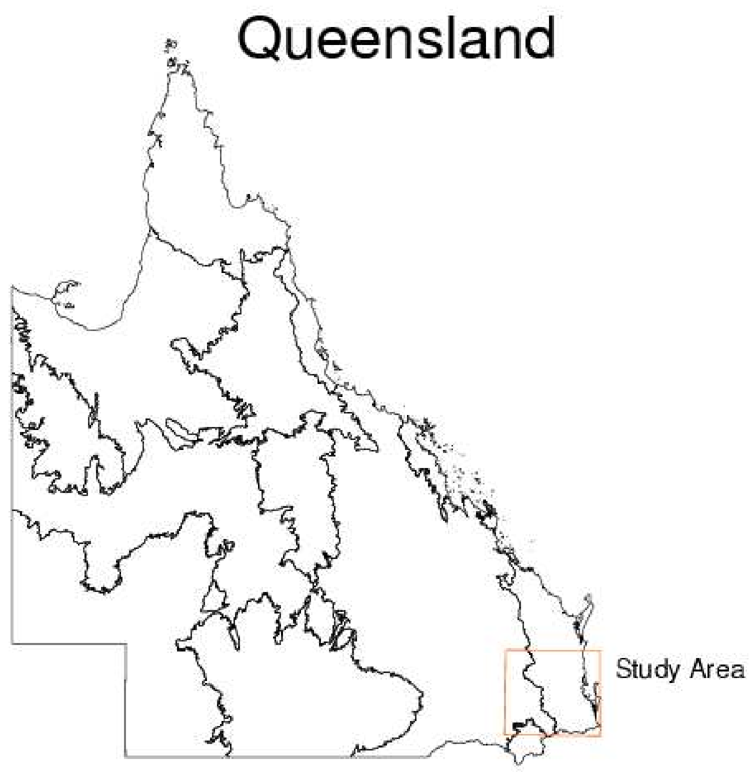

Figure 8: Location of study area

the expert was asked to consider placing 100 field sites at locations like the one chosen, in the same vicinity. Their median was the value in which the expert believed would be the proportion of sites in which the species would be found, given an exhaustive search each time. The upper and lower 0.05 quantiles were chosen by asking the higher and lower number which the expert thought was believable, but surprising. The value of *surprising* was set based on the 0.05 and 0.95 quantiles from the lottery exercise mentioned earlier. These upper and lower quantiles were set by the two points on the graph. Since we were fitting a beta distribution, these points are not independent, which occasionally caused some frustration for the expert, suggesting their prior for the probability of occurrence didn't necessarily match a beta distribution. We preferred to force them to compromise, as the compromise would have been made later by choosing the best beta distribution fit to their chosen points anyway. The expert also had the option to adjust the parameters for the beta distribution directly, which automatically updates the plot of the beta distribution. This was found to be often convenient for either fine tuning their chosen distribution, or for quickly choosing specific distributions such as the uniform distribution if they had no prior opinion for the point (which never actually happened), or for the extremes (almost no chance of occurrence at that point, or almost definitely present). Below the graph are the values of the four explanatory variables and the detailed description of the geology.

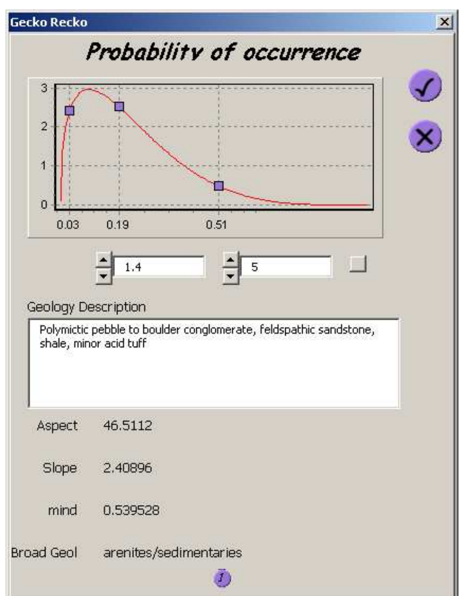

Figure 9: Beta distribution elicitation dialog.

The expert can accept this distribution if they are satisfied it matches their prior belief, or cancel and move to another point.

Once a minimum number of points have been selected the expert is able to view the univariate response curves (Figure 10) created by fitting a logistic regression model to the design points provided by the expert. These curves show the univariate relationship between the probability of presence and each of the explanatory variables in turn. They are created plotting the predicted probability of occurrence over a range of values for a particular variable, with all other variables fixed at some value, typically the mean. A categorical variable such as *geology*, is represented by boxplots rather than a curve. Many ecologists use these curves both as an aid to understanding the relationship a species has with a particular environmental gradient, and also as a check that a statistical model makes sense from a physical point of view.

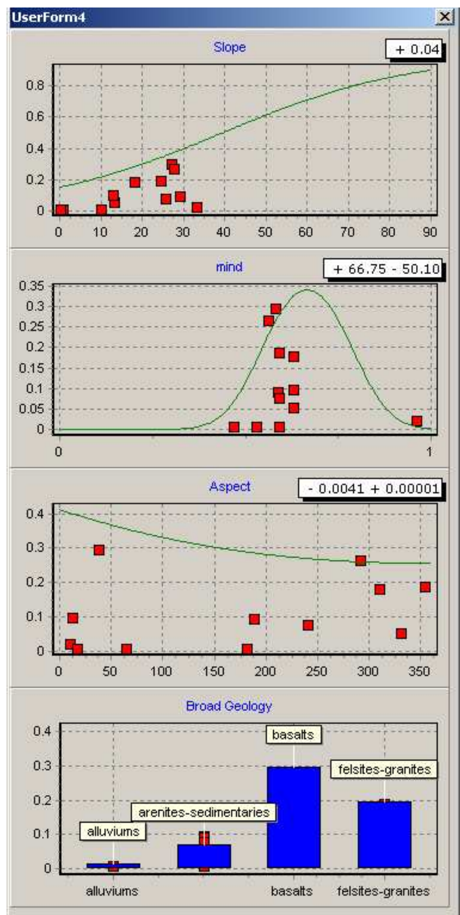

Figure 10: Univariate response curves

At each design point $\mathbf{x}_i$ the expert has provided a prior for the probability of presence of the species, $p(p_i) = \text{Beta}(\alpha_i, \beta_i)$. Using the mean for the estimate of the probability of presence, $\hat{p}_i = \frac{\alpha_i}{\alpha_i+\beta_i}$, we determine the pseudo observations and binomial sample size such that the binomial variance is approximately equivalent to the variance of $p(p_i)$,

$$\hat{n}_i = \frac{\hat{p}_i(1-\hat{p}_i)(\alpha_i+\beta_i)^2(\alpha_i+\beta_i+1)}{\alpha_i\beta_i} \tag{4}$$

$$\hat{y}_i = \hat{n}_i\hat{p}_i \tag{5}$$

This approach is equivalent to a weighted logistic regression, with each design point weighted by the variance of $p(p_i)$.

The expert is now able to add and review design points. The response curves displayed are interactive, in that the user may click on any point, which will zoom the map to that point and display the beta fitting form. The user can then modify $p(p_i)$, and immediately see the effect the adjustment has on the response curves. The expert also has the option of reviewing those points which are influential, or appear to be outliers. An interactive graph shows a plot of the residuals, and a plot of Cook's distances (Figure 11). Again the expert can click on a point and review the design point. The graphs are linked, so that a selected point in one graph will cause the corresponding point in each of the other graphs to be highlighted.

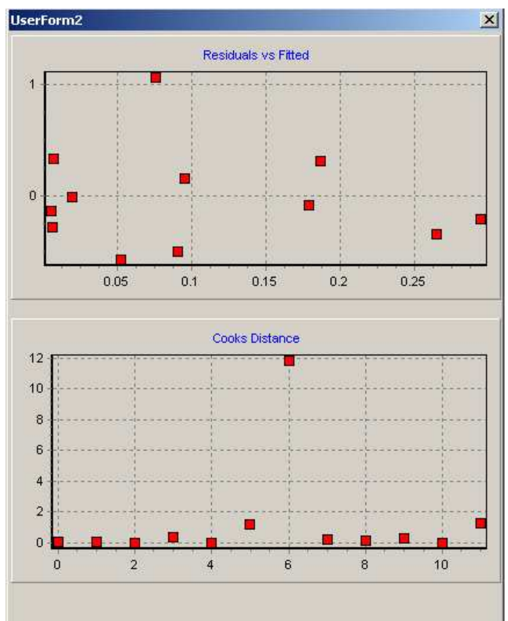

Figure 11: Diagnostics can be displayed and queried. The graph is linked both to the design point and to the response curves, so any change in the value of the expert's $p(p_i)$ will automatically update all graphs.

The expert continues the procedure, getting feedback from the response curves until the data they have provided matches what they believe is the relationship between the species and its environment.

Each elicitation procedure took approximately one hour, and resulted in 10 design points from expert J and 13 from expert P. Ideally more points would have been selected, but due to time constraints the elicitation exercise finished when each expert was satisfied with the agreement between their virtual field sites and the response curves.

Once the expert has provided a prior distribution for each design point, we convert these into a prior for the regression parameter β. This could be done using the same weighted regression approach using Equation 5, but we prefer to use a simulation based approach. While this takes a little more computational time, it makes no assumptions to derive pseudo sample sizes, or asymptotic normality assumptions in the GLM estimation stage.

At each j iteration and for each of the expert's design points, we sample a probability of occurrence $\tilde{p}_{ij}$ from the distribution chosen by the expert, that is Beta(α_i, β_i). We then fit a linear regression model to the logit of these observations, using least squares to estimate b_j in

$$\begin{aligned}\text{logit}(\tilde{p}_{ij}) \quad = \quad & b_{0j} + b_{1j}x_{slope,i}b_{2j}x_{mind,i} + b_{3j}x^2_{mind,i} + \\ & b_{4j}x_{aspect,i} + b_{5j}x^2_{aspect,i} \\ & b_{6j}x_{geol1,i} + b_{7j}x_{geol2,i} + b_{8j}x_{geol3,i}.\end{aligned}$$

Here the subscript i represents the ith design point provided by the expert, and j the jth simulated value. b_j is thus one draw from the expert's prior for β. This sample from the prior can then be used in a subsequent Bayesian analysis when the analysis employs a Markov Chain Monte Carlo estimation step. Alternatively, we can summarise the prior by assuming it comes from a known multivariate distribution, such as multivariate normal.

The results of the elicitation exercise indicated that in general P's prior was less disperse than J's, although both priors include 0 in the 95% credible interval. Table 2 and Figure 12 compare the two priors.

To investigate the effect on the posterior, we summarised the prior from the experts by assuming they were multivariate normally distributed, and used the function `MCMClogit` from the `R` package `MCMCpack` (Martin et al., 2003) to generate the posteriors.

The priors elicited from the expert were reasonably informative, with the posteriors using these priors clearly different from the posterior using a uniform improper prior. This is particularly noticeable with the *geology* variable. Figure 13 illustrates the differences.

The reason for the large effect of geology can be seen by examining how the observations fall into each of the *geology* classes. As shown in the table below, only two groups had observed absences, which explains the large negative coefficients for these parameters in the posterior with the non-informative priors. Both experts were surprised at this effect, and this is shown in their posteriors for that parameter (see Figure 13). To investigate this a little further, we regenerated the priors and posteriors, omitting the

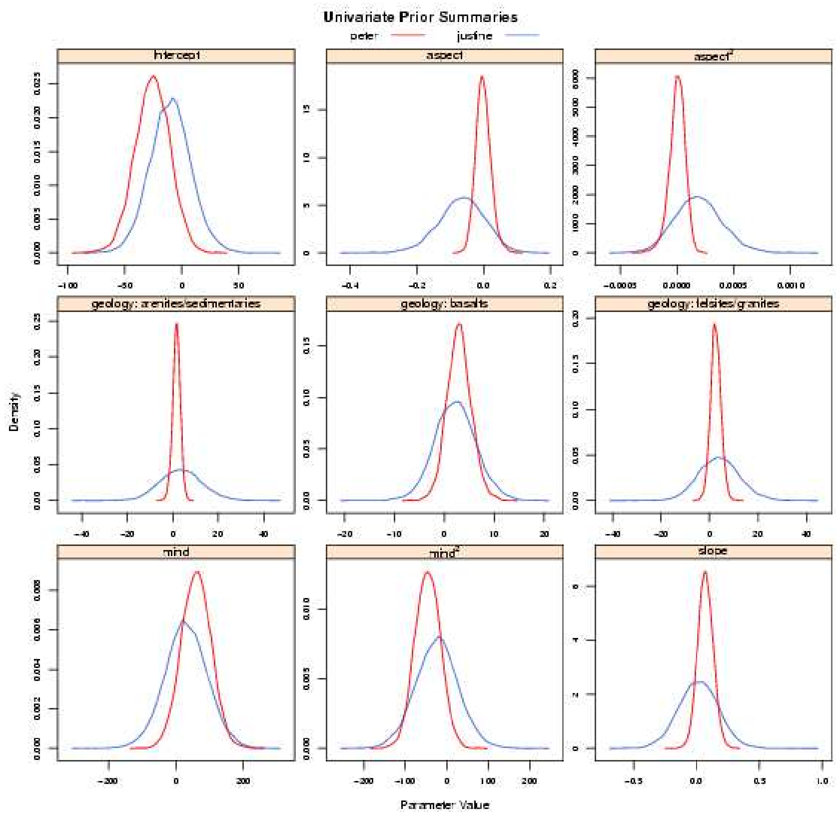

Figure 12: Comparison of priors for J (blue line) and P (red line).

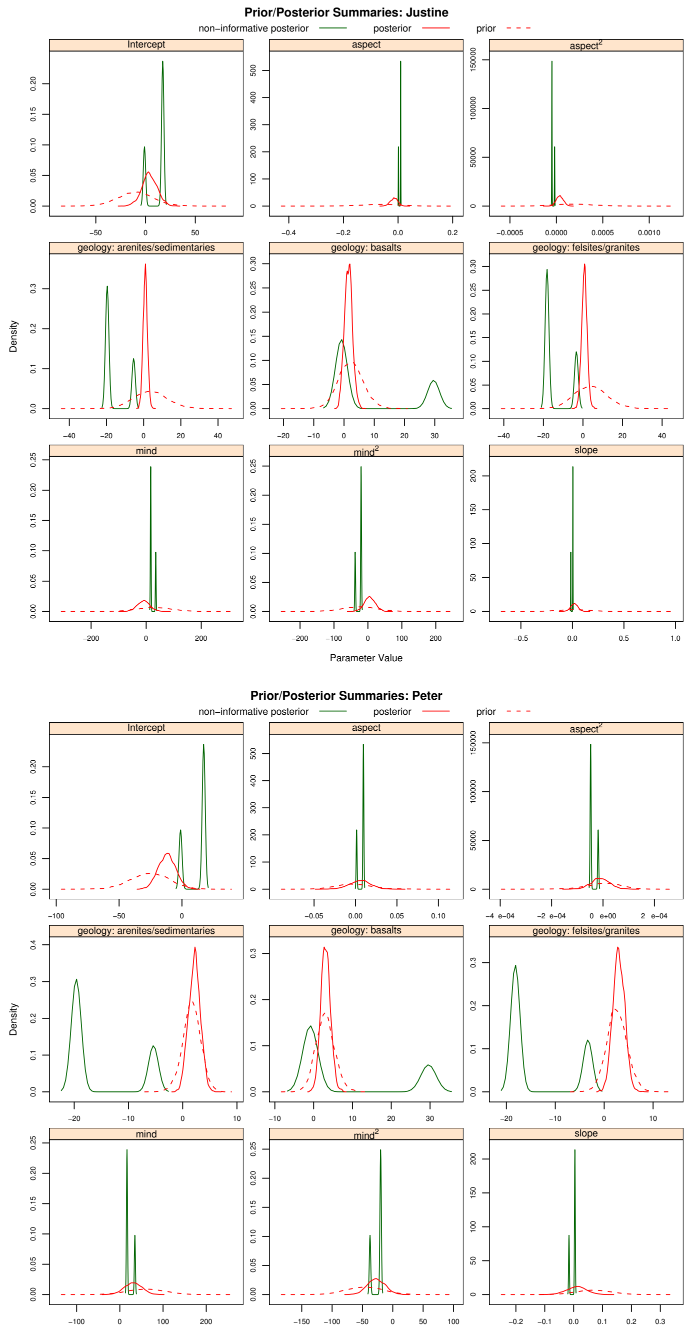

Figure 13: Informative prior (broken red line), posterior using the informative prior (solid red line), and posterior using a non-informative prior (solid green line), for expert J (top) and expert P (bottom). Note the discrepancies for the *geology* parameters.

	Expert J	Expert P
b_0	−10.08 (−46.28,25.33)	−25.67 (−56.49,3.27)
b_1	0.02 (−0.29, 0.33)	0.07 (−0.05,0.19)
b_2	30.79 (−96.76, 159.69)	61.70 (−25.28,148.70)
b_3	−23.66 (−129.95, 80.79)	−46.80 (−107.30,13.19)
b_4	−0.06 (−0.20, 0.07)	−0.002 (−0.04,0.05)
b_5	1.48e-4 (−2.07e-4,6.04e-4)	5.57e-4 (−1.53e-04,1.23e-04)
b_6	3.61 (−14.04, 21.78)	1.69 (−1.86,4.82)
b_7	2.20 (−5.81, 10.36)	3.01 (−1.83,7.88)
b_8	3.89 (−12.19, 20.59)	2.61 (−1.45,6.82)

Table 2: Comparison of the median values of each expert's prior, with a 95% credible interval.

geology variable. From Figure 15 we see that the two experts' priors become much more similar, with only a noticeable difference for the *aspect* parameter. Expert J had fairly firm beliefs that a northerly aspect would increase the probability of occurrence. This is reflected in the posterior response curve for that variable (Figure 14). The posterior response curve is constructed in the same manner as the response curves used in the elicitation procedure (see Section 5).

This is at odds with the posterior from the non-informative prior, which suggests higher probabilities at sites with southerly aspects. Although expert P agreed with expert J with respect to the relationship between aspect and species presence, his prior was not sufficiently informative to overwhelm the data. The nature of the response curves become a clearer in the absence of the *geology* variable. From Figure 16, we see that both experts expect to see an increase in probability of occurrence as slope increases. This is not supported by the data with a non-informative prior.

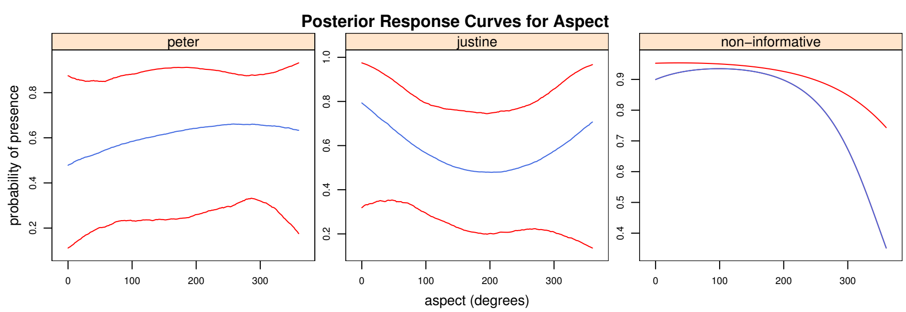

Figure 14: Expert J expressed a belief that a northerly aspect would increase the probability of occurrence. This is reflected in the posterior, but is not apparent when a non-informative prior is used. Although Expert P agreed with expert J in the nature of the response between aspect and probability of occurrence, their prior was not sufficiently informative to reverse the trend suggested by the data.

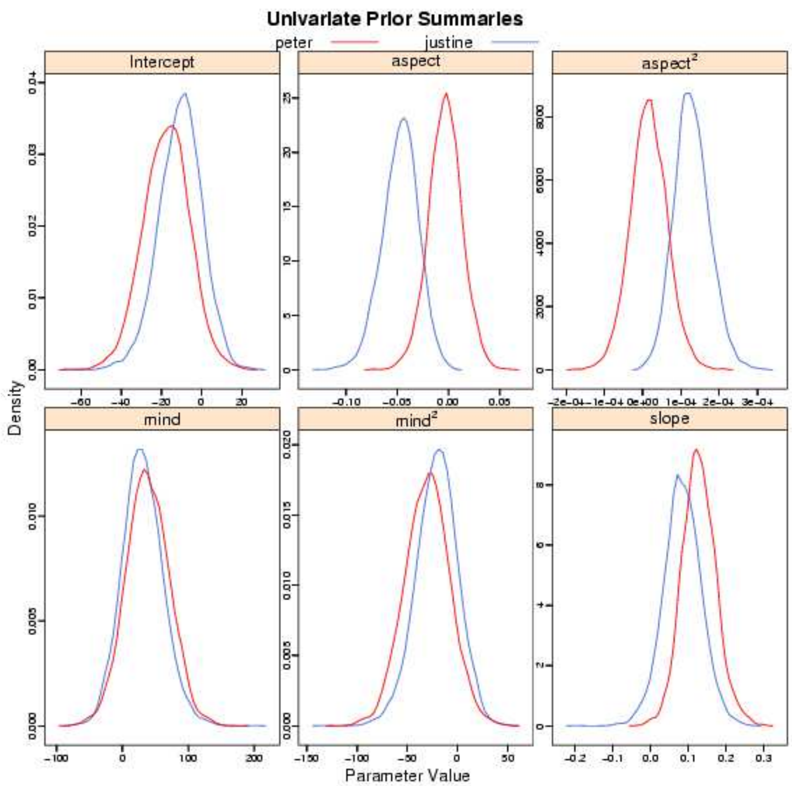

Figure 15: Comparison of priors for J (blue line) and P (red line), with the variable *geology* omitted

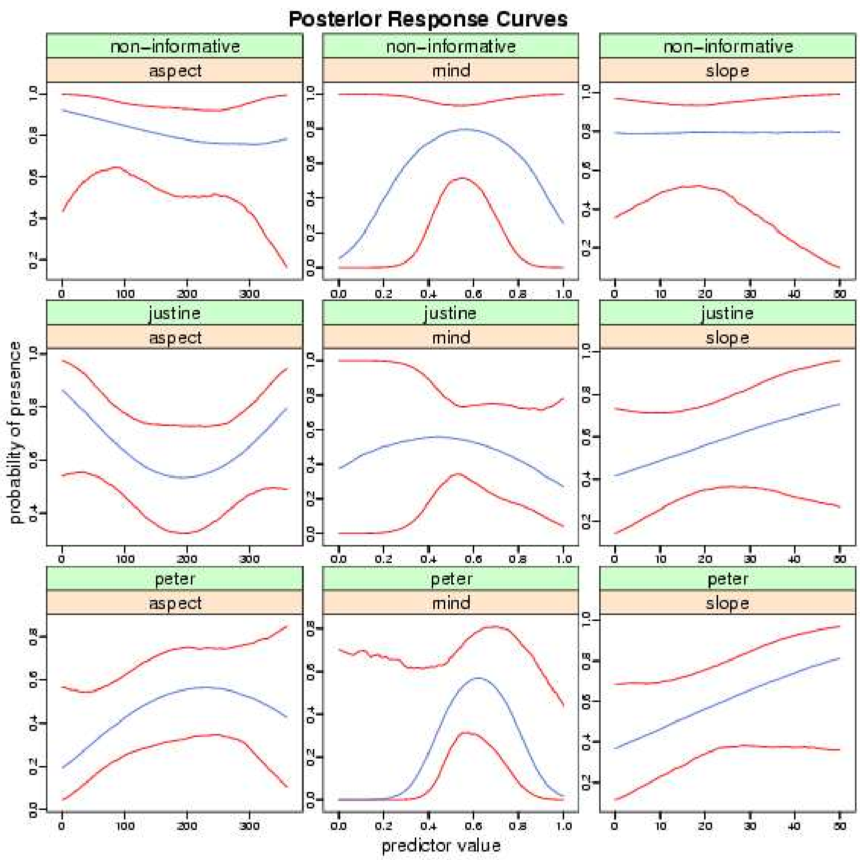

Figure 16: Response curves for the model omitting *geology*. Note that both models with informative priors suggest an increase in probability as slope increases. The posterior using the non-informative prior does not support this.

Geology class	Absence	Presence
alluviums	0	3
arenites/ sedimentaries	7	13
basalts	0	6
felsites/granites	2	7

Table 3: Distribution of observed presences and absences into each of the four broad geology groups used in the model.

6 Discussion

This paper addresses the problem of eliciting expert information in the context of geographic data by proposing a new combination approach that acknowledges the different structural and predictive skills of experts and their different needs in the elicitation process. Specialist software was developed that complements familiar workplace GIS tools used for geographic data. The proposed methodology was used in two quite different experimental situations and statistical models. Although the case studies are admittedly limited in generality, the strong indication in both experiments was that the exploitation of the experts' skills in geographic thinking, and the combination of both structural and predictive elicitation, was beneficial in terms of ease of use by the experts, timeliness of elicitation, feedback and satisfaction with the resultant quantification of information.

In this approach, the structural feedback with the predictive questions suits both types of experts, so that a pure physiologist will use the points as a means to get to the curves he wants, a pure field ecologist will not really be interested in the curves at all, since he has no opinion on the relationship between the species and its environment, he just knows where it lives. Obviously, most people will be between these extremes, and this approach caters for these. Moreover, it is better than just adjusting response curves, since the predictive questions allows us to get a handle on the expert's prior uncertainty. It is also better than a solely predictive approach, since it has all the spatial context, and avoids the expert having to answer questions on esoteric explanatory variables (but the response curves allow them to check that their answers are consistent).

It is difficult to assess the quality of any elicitation procedure, since differences between procedures may merely reflect differences in opinions between experts, rather than a differences due to the methodology. We also have no measurable way of determining how close the quantification of opinion matches an expert's true beliefs. Some work has been done in the area of validation of elicitation procedures. Kadane and Wolfson (1998) in their survey of the literature cite relevant measurements of reliability (Wallsten and Budescu 1982), coherence (Lindley et al. 1979) and calibration (Morgan and Henrion 1990).

In this paper, the aim of the case studies was not to create the most accurate prediction of house prices or the distribution of the brush-tailed rock wallaby, but rather provide an example of how expert opinion can be combined with real data to produce a statistical model that balances our prior beliefs with what the data tell us. Had we wished to create an accurate map of either distribution, we may have needed to spend a considerably longer time collecting data, educating experts and choosing and organising a suite of explanatory variables. There are very real problems with the sets of observations, and we have made no mention of how they were sampled, or their representativeness. That said, the datasets are not atypical of the sort of data faced in practice, especially by ecologists and particularly when the species in question is rare or poorly studied. The challenge, then, is not to reject the data as inadequate, but to make the best of the situation. This work, for example, shows that when the data suggest a relationship between the response and the environmental variables which does not make physical sense, the expert opinion can be used to force a compromise between a believable model and a purely data driven model. The level of compromise depends on the strength of the prior. When there is less conflict between the prior and the likelihood, such as when we considered the model with *geology* omitted in the species distribution study, then the informative prior acts to increase the precision of the posterior.

We believe that one key advantage of the approach we have outlined in this paper is that it allows the expert to make use of contextual, spatial information when asking questions about the response at various levels of the explanatory variables. The usefulness of this contextual information depends largely on the expert. In our second case study, for example, one expert, expert J, with extensive local knowledge, used this contextual information extensively when developing her prior. Expert P, in contrast, relied more heavily on the values of the explanatory variables. Our approach is designed to allow the full spectrum of experts to use our GIS based approach.

O'Hagan (1998) and Craig et al. (1998) make the point that most elicitation exercises will involve computer assistance, but also that if statisticians are required to write specialist software each time they wish to elicit expert opinion, then elicitation will remain a rarely used procedure. More work needs to be done in the development of general purpose software to encourage others to conduct serious elicitation. The work we show is in one respect fairly specialised, dealing only with logistic regression in a spatial context, and showing how it works for one particular species and a given set of explanatory variables, each of which have been coded in explicitly. We didn't set out to create a general tool, but we believe that it would be a suitable model for further, more generalised approaches. It would be very simple, for example, to use the basis of this example for different species and a different set of explanatory variable. There is plenty of scope for further research here, notably how to make it flexible enough to deal with different statistical models, more than a handful of explanatory variables (since many variables will present problems for display and interpretation), but perhaps most interestingly and challenging, how to include variable selection.

Throughout this paper we have referred to the problem as one of regression with geographic data, rather than using the term *spatial regression*, since this implies a model which takes into account the relationship nearby sites may have with each other. Our description refers to a simpler regression model, where although the variables of interest are distributed throughout space we make no attempt to describe the spatial autocorrelation. We have chosen the simpler model to illustrate an elicitation technique, but there is no reason why this approach can not be applied to more complex regression approaches, such as spatial regression.

A final remark, though, is that if expert opinion is going to be useful, then it implies a limited amount of data, which also suggests that complex models would be unlikely to be supported. We suggest that every effort be made to keep the model as simple as possible. The expert model could be considered to be an early stage in the modelling of the species' distribution, useful for highlighting areas to collect more data or consider other approaches. Our aim is to derive the most from our existing data, which includes expert opinion.

References

Albert, J. H. (1988). "Computational methods using a Bayesian hierarchical generalized linear model." *Journal of the American Statistical Association*, 83: 1037–1044. 102

Albert, J. H. and Chib, S. (1993). "Bayesian analysis of binary and polychotomous response data." *Journal of the American Statistical Association*, 88: 669–679. 102

Bedrick, E. J., Christensen, R., and Johnson, W. (1996). "A new perspective on priors for generalized linear models." *Journal of the American Statistical Association*, 91: 1450–1460. 102

— (1997). "Bayesian binomial regression: Predicting survival at a trauma center." *The American Statistician*, 51: 211–218. 102

Bernardo, J. M. (1979). "Reference posterior distributions for Bayesian inference." *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 41: 113–147. 99

Bernardo, J. M. and Smith, A. F. M. (1994). *Bayesian Theory*. Chichester: Wiley. 99

Carter, K. and Goldizen, A. W. (2003). "Habitat choice and vigilance behaviour of brush-tailed rock-wallabies (*Petrogale penicillata*) within their nocturnal foraging ranges." *Wildlife Research*, 30: 355–364. 118

Chen, M. H. and Dey, D. K. (2003). "Variable selection for multivariate logistic regression models." *Journal of Statistical Planning and Inference.*, 111: 37–55. 102

Chen, M. H., Ibrahim, J. G., Shao, Q. M., and Weiss, R. E. (2003). "Prior elicitation for model selection and estimation in generalized linear mixed models." *Journal of Statistical Planning and Inference.*, 111: 57–76. 102

Chen, M. H., Ibrahim, J. G., and Yiannoutsos, C. (1999). "Prior elicitation, variable selection and Bayesian computation for logistic regression models." *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 61: 223–242. 102

Cooke, R. M. (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science*. New York: Oxford University Press. 100

Craig, P. S., Goldstein, M., Seheult, A. H., and Smith, J. A. (1998). "Constructing partial prior specifications for models of complex physical systems (Disc: p55-68)." *The Statistician*, 47: 37–53. 103, 131

Eldridge, M. D. B. (1997). "Rock-wallaby conservation: essential data and management priorities. Proceedings of the 1994 National Rock-wallaby Symposium workshops." *Australian Mammalogy*, 19: 325–330. 118

Garthwaite, P. H. and Dickey, J. M. (1988). "Quantifying expert opinion in linear-regression problems." *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 50: 462–474. 102

Gavaris, S. and Ianelli, J. (2002). "Statistical issues in fisheries' stock assessments (with discussion)." *Scandinavian Journal of Statistics*, 29: 245–271. 100

Gelfand, A. E., Sahu, S. K., and Carlin, B. P. (1996). "Efficient parametrisations for generalized linear mixed models (with discussion)." In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M. (eds.), *Bayesian Statistics*, volume 5, 165–180. Oxford: Oxford University Press. 102

Gibson, L. a., Wilson, B. A., Cahill, D. M., and Hill, J. (2004). "Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach." *Journal of Applied Ecology*, 41: 213–223. 102

Guisan, A. and Zimmermann, N. E. (2000). "Predictive habitat distribution models in ecology." *Ecological Modelling*, 135: 147–186. 101

Kadane, J. (1980). "Predictive and structural methods for eliciting prior distributions." In Zellner, A. (ed.), *Bayesian Analysis in Econometrics and Statistics – Essays in Honor of Harold Jeffreys*. Amsterdam: North Holland: North Holland Publishing Company. 100

Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. S., and Peters, S. C. (1980). "Interactive Elicitation of Opinion for a Normal Linear Model." *Journal of the American Statistical Association*, 75(372): 845–854. 100, 102, 103

Kadane, J. B. and Wolfson, L. J. (1998). "Experiences in elicitation." *Journal of the Royal Statistical Society. Series D. The Statistician*, 47: 3–19. 102, 104, 115, 130

Kass, R. E. and Wasserman, L. (1996). "The selection of prior distributions by formal rules." *Journal of the American Statistical Association*, 91: 1343–1370. 99

Lindley, D. V., Tversky, A., and Brown, R. V. (1979). "On the reconciliation of probability assessments (with discussion)." *Journal of the Royal Statistical Society. Series A. Statistics in Society*, 142: 146–180. 115, 130

Manel, S., Williams, H. C., and Ormerod, S. J. (2001). "Evaluating presence-absence models in ecology: the need to account for prevalence." *Journal of Applied Ecology*, 38: 921–931. 101

Maynes, G. M. and Sharman, G. B. (1983). "Brush-tailed Rock-Wallaby." In Strahan, R. (ed.), *The Complete Book of Australian Mammals*. Sydney: Angus and Robertson. 118

Morgan, M. G. and Henrion, M. (1990). *Uncertainty: a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. New York: Cambridge University Press. 115, 130

O'Hagan, A. (1998). "Eliciting expert beliefs in substantial practical applications." *The Statistician*, 47: 21–35. 131

Pearce, J. L., Cherry, K., Drielsma, M., Ferrier, S., and Whish, G. (2001). "Incorporating expert opinion and fine-scale vegetation mapping into statistical models of faunal distribution." *Journal of Applied Ecology*, 38: 412–424. 101

R Development Core Team (2004). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org 108

Store, R. and Kangas, J. (2001). "Integrating spatial multi-criteria evaluation and expert knowledge for GIS-based habitat suitability modelling." *Landscape and Urban Planning*, 55: 79–93. 101

Tamis, W. L. M. and Van 't Zelfde, M. (1998). "An expert habitat suitability model for the disaggregation of bird survey data - Bird counts in the Netherlands downscaled from atlas block to kilometre cell." *Landscape and Urban Planning*, 40: 269–282. 101

Tversky, A. and Kahneman, D. (1974). "Judgement under uncertainty: heuristics and biases." *Science*, 185: 1124–1131. 100

Wallsten, T. S. and Budescu, D. V. (1982). "Encoding subjective probabilities: a psychological and psychometric review." *Management Science*, 29: 151–173. 115, 130

West, M. (1985). "Generalized linear models: scale parameters, outlier accommodations and prior distributions." In Bernardo, J. M., DeGroot, M. H., Lindley, D. V., and Smith, A. F. M. (eds.), *Bayesian Statistics*, volume 2. Amsterdam: North-Holland. 102

West, M., Harrison, P. J., and Migon, H. S. (1985). "Dynamic generalized linear models and Bayesian forecasting (with discussion)." *Journal of the American Statistical Association*, 80: 73–97. 102

Wolfson, L. J. (1995). "Elicitation of priors and utilities for Bayesian analysis." Ph.D. thesis, Department of Statistics, Carnegie Mellon University, Pittsburgh. 100

Zeger, S. L. and Karim, M. R. (1991). "Generalized linear models with random effects; a Gibbs sampling approach." *Journal of the American Statistical Association*, 86: 79–86. 102

Zellner, A. and Rossi, P. E. (1991). "Bayesian analysis of dichotomous quantal response models." *Journal of Econometrics*, 25: 365–393. 102

Acknowledgments

We are grateful to the case study experts for their time and enthusiasm, patience and expertise. We would particulary like to thank Justine Murray from the University of Queensland and Peter Jarmin from the University of New England for their input into the brush tailed rock wallaby modelling. We are also grateful to Rebecca O'Leary for organising the expert elicitation workshop at QUT and to the participants of this workshop for their positive input.

This research was supported by the Australian Research Council.