Journal of Applied Ecology

Journal of Applied Ecology doi: 10.1111/j.1365-2664.2011.01971.x

Quantifying variance components in ecological models based on expert opinion

Christina A. Czembor*, William K. Morris, Brendan A. Wintle and Peter A. Vesk

School of Botany, The University of Melbourne, Parkville, Vic. 3010, Australia

Summary

1. Expert opinion is often relied on to build ecological models when empirical data are absent. Despite widespread use, expert models often ignore uncertainty though it may affect predictions. We assess the importance of such uncertainty for modelling forest restoration.

2. Of an initial 19 experts found using a literature search and peer recommendations, five parameterized models predicting ecological responses to proposed restoration actions for a degraded dry woodland system in Victoria, Australia. We incorporated uncertainty from three sources: disagreement (between-expert uncertainty), self-assessed imprecision (within-expert uncertainty) and modelled system stochasticity. These sources of uncertainty were quantified as variance components in hierarchical models.

3. The between-expert variance component contributed more to overall model uncertainty than both within-expert variance and modelled system stochasticity. The estimate of between-expert variance also had the greatest parameter uncertainty of the three components.

4. *Synthesis and applications*. We present a method to decompose variance in model predictions. We suggest that modelling strategies relying on a single expert opinion, consensus between experts, or that only incorporate uncertainty due to system stochasticity could produce biased models and over-confident predictions. Management decisions based on biased and over-confident predictions lead to inefficient conservation investments and poor outcomes. Our research highlights the importance of seeking multiple expert opinions to fully characterize uncertainty and make robust decisions.

Key-words: expert opinion, forest management, hierarchical model, state-and-transition, uncertainty, variance components analyses, vegetation dynamics development tool

Introduction

Expert opinion has often been used to parameterize ecological models when other data are scarce or unavailable (e.g. Johnson & Gillingham 2004; Forbis *et al*. 2006; O'Neill *et al*. 2008). Empirical data may be limited because complex ecosystem interactions are difficult to measure (Ferguson *et al*. 2008), or because data collection is limited by time and money (Johnson & Gillingham 2004; Murray *et al*. 2009). While expert opinion is not equivalent to empirical data (Pearce *et al*. 2001), sometimes it may be the only information available to construct models and inform management decisions (Johnson & Gillingham 2004; O'Neill *et al*. 2008). Expert opinion may also add value to existing data within Bayesian modelling and decision frameworks (McCarthy 2007).

Failing to account for model uncertainty can lead to ineffective management (Johnson & Gillingham 2004; Buckley *et al*. 2005). In expert-opinion models, predictive uncertainty arises from multiple sources (Regan, Colyvan & Burgman 2002). The relative magnitude of these various forms of uncertainty can affect the strategy of model formulation and decisions based on those models. Yet, these sources of uncertainty and their relative magnitudes have to date provoked little discussion. In this article, we describe an approach to estimating the variance components of vegetation dynamics model predictions.

We consider three components that cause predictive uncertainty: system stochasticity, between-expert uncertainty and within-expert uncertainty. System stochasticity arises because complex biological processes, such as disturbances, vary randomly in time (e.g. Raulier, Pothier & Bernier 2003; Kangas & Kangas 2004) and space (e.g. McCarthy & Burgman 1995). Uncertainty due to natural stochasticity is commonly incorporated into ecological models by repeatedly sampling from distributions for each model parameter, using Monte Carlo simulations and related approaches (e.g. McCarthy &

*Correspondence author. E-mail: c.czembor@gmail.com

© 2011 The Authors. Journal of Applied Ecology © 2011 British Ecological Society

Burgman 1995; Buckley *et al.* 2005; Nicholson & Possingham 2007).

Within-expert uncertainty occurs when an expert is unsure of a parameter value; this is sometimes known as imperfect knowledge. The degree of confidence an expert has in their estimate can be characterized with the use of credible intervals, imprecise probabilities, p-bounds or other bounding methods (Ferson 1996; Walley 2000; McCarthy 2007; Speirs-Bridge *et al.* 2010). Often only a point estimate is used for parameters, rather than a distribution or range. Point estimates do not account for the imperfect knowledge of experts and decisions relying on predictions from such models are not robust to within-expert uncertainty (Ben-Haim 2001). There are few studies that have addressed within-expert uncertainty (e.g. Alho, Kangas & Kolehmainen 1996; Walker, Evans & MacIntosh 2001) and techniques for eliciting uncertainty intervals for expert parameter estimates are rarely applied in ecology and conservation (but see Speirs-Bridge *et al.* 2010). We approach this problem by conducting a bounded sensitivity analysis to test the importance of within-expert uncertainty.

Between-expert uncertainty occurs because of disagreement among experts. Disagreement can reflect spatial heterogeneity in processes (Martin *et al.* 2005), a dearth of knowledge about processes (Morgan, Pitelka & Shevliakova 2001), or cognitive biases (Burgman 2005). Like within-expert uncertainty, between-expert uncertainty is also only rarely addressed or quantified in expert-opinion-based modelling (known examples are Alho, Kangas & Kolehmainen 1996; Czembor & Vesk 2009; Hurley, Rapaport & Johnson 2009; Johnson & Gillingham 2004; O'Neill *et al.* 2008). Commonly, between-expert uncertainty is avoided entirely by consulting only one expert (Sutherland 2006; O'Neill *et al.* 2008) or forcing expert groups to reach consensus (Morgan, Pitelka & Shevliakova 2001; Irvine *et al.* 2009) prior to parameterization or after using mathematical averaging, Delphi methods, or other approaches (e.g. Kangas *et al.* 1998; Leskinen & Kangas 2001; Sutherland 2006; Murray *et al.* 2009). Consensus methods produce parameters with only single values (e.g. Forbis *et al.* 2006; Vavra, Hemstrom & Wisdom 2007). Estimates made from forced consensus can be susceptible to over-confidence (Cooke 1991) because some participants dominate the group or participants submit to the opinion of the majority (Clemen & Winkler 1999; Burgman 2005; Sutherland 2006).

When multiple experts construct models of ecosystem dynamics in isolation, there can be large differences in the predictions of the different expert models (Czembor & Vesk 2009). If we consider each expert model to be equally plausible, it is difficult to make management decisions based on diverse model predictions (Czembor & Vesk 2009). The difficulty is compounded when we consider within-expert uncertainty and system stochasticity. Currently, little is known about the relative contributions of system stochasticity, within- and between-expert uncertainty to overall model uncertainty (but see Alho, Kangas & Kolehmainen 1996); such information is necessary for determining the robustness of predictive models used to inform restoration decisions.

Here, we present an approach to quantifying the components of uncertainty that cause variation in expert model predictions for use in decision-making. We illustrate our approach using a state-and-transition modelling framework to predict the dynamics and restoration trajectory of overstorey *Eucalyptus* vegetation in a dry woodland ecosystem of Victoria, Australia. We use information elicited from a small number of experts in a field that has relatively few experts and almost no relevant data for the restoration decisions at hand.

Materials and methods

CASE STUDY: THE RESTORATION OF DRY WOODLANDS IN VICTORIA, AUSTRALIA.

We focus on a heavily degraded ecosystem in Victoria, Australia, the Box-Ironbark forests and dry woodlands, which consist primarily of ironbark (*Eucalyptus tricarpa*) and box (*E. microcarpa, E. polyanthemos, E. melliodora*) eucalypts. These woodlands have been harvested for timber since the 1830s (Forests Commission of Victoria 1928; Newman 1961; Kellas 1991). Long-term harvesting has changed them from open woodlands with heterogeneous stand structure and many large – 120 to 150 cm d.b.h. – hollow-bearing trees (Newman 1961; Kellas 1991) to woodlands without large, hollow-bearing trees and relatively homogeneous structure dominated by dense stands of small coppice stems (Environment Conservation Council 2001). Loss of area and change in vegetation structure, including the absence of hollows required by many fauna species (Kellas 1991; Alexander 1997; Soderquist 1999), have contributed to severe population declines, extirpations and extinctions of over 350 species (Muir, Edwards & Dickens 1995; Alexander 1997; Environment Conservation Council 2001).

Management authorities recently initiated three strategies to restore Box-Ironbark stand structure to more closely resemble pre-harvesting conditions: new ecological thinning prescriptions, modified timber harvesting regulations, and stand development without harvesting. Ecological thinning reduces stem density to a level similar to an older stand. The aim of thinning is to promote increased growth rates of remaining trees and accelerate the development of large trees (Environment Conservation Council 2001). Modified harvesting regulations stipulate that hollow-bearing and large trees, > 60 cm DBH, are not harvested and that higher densities of medium-sized trees are left for recruitment to larger-size classes (Department of Natural Resources and Environment 1998; Sutton 2000). Stand development without harvesting constitutes ceasing all harvesting but allowing fire management.

Management authorities recognized that direct evidence for an effect of different strategies would take many years to eventuate due to the slow growth of trees. To compensate for a lack of field data an expert elicited model was built; this is described in detail by Czembor & Vesk (2009). The present study complements this real-world example of forest restoration and illustrates how to quantify the different sources of variance in model predictions.

SELECTING EXPERTS

A lack of existing field data necessitated the use of expert opinion to derive model parameters. We defined an expert as someone with:

1. A minimum of 5 years of education, research experience or technical training in Box-Ironbark vegetation dynamics.

2. Very high or high (self-assessed) levels of theoretical and/or practical experience working in Box-Ironbark woodlands.
3. Published research on Box-Ironbark dynamics or management in peer-reviewed journals or government agency-endorsed reports.
4. Peer nomination of being an expert.

We located candidate experts by consulting the park management authority's steering committee for Box-Ironbark restoration and conducting a literature search for all published articles on Box-Ironbark woodlands. We contacted all identified candidates to ask for recommendations of other people who might be eligible to participate. This resulted in a total of 48 candidates, 19 of which passed our expert selection guidelines (wildlife biologists, social scientists and non-local or inexperienced candidates were removed at this stage). We were able to contact 14 of these experts, of whom nine agreed to participate. Those unable to participate cited a lack of time as the primary reason they were unavailable. We sent electronic surveys to these experts; eight of which were returned. We conducted personal interviews with these eight experts and could construct models for five. Models were not constructed when experts could not provide the full suite of parameters required to generate a simulation model.

Table 1. Transition agents for vegetation states and their associated management scenarios

Transition agent	Natural disturbance	Current harvesting	Ecological thinning
Growth	✓	✓	✓
Coppice response	✓	✓	✓
Dodder laurel	✓	✓	✓
Drought	✓	✓	✓
Drought + mistletoe	✓	✓	✓
Insect attack	✓	✓	✓
Wildfire	✓	✓	✓
Wind-throw	✓	✓	✓
Wind-throw + wildfire	✓	✓	✓
General timber harvesting	–	✓	–
Firewood harvesting	–	✓	–
Sawlog harvesting	–	✓	–
Ecological thinning	–	–	✓
Ecological thinning + poison	–	–	✓
Fuel reduction burn	–	✓	✓

MODELLING VEGETATION DYNAMICS

State-and-transition simulation models (STSMs) are used to characterize vegetation dynamics. An STSM defines a set of discrete vegetation states and associated transition agents such as vegetation growth, natural disturbances or management actions (Westoby, Walker & Noy-Meir 1989). An algorithm applies the transition agents, which have associated probabilities of occurring, to non-spatial vegetation units, or cells, iteratively over a series of timesteps and, according to these probabilities, the cells change from one state to another.

Here, we used a computer software implementation of STSMs, the vegetation dynamics development tool (VDDT; ESSA Technologies Ltd 2007), to predict the long-term vegetation dynamics of Victorian Box-Ironbark woodlands. The models were parameterized using the information elicited from five experts. The electronic survey results were used to determine the four vegetation states included in the VDDT models: high-density regrowth, low-density regrowth, high-density mature and low-density mature (Fig. 1). In subsequent interviews, experts assisted in estimating the initial conditions of the models and specified how each transition agent (Table 1) would alter vegetation states. A detailed list of transition agent probabilities for each model is available in Appendix S1 (Supporting information) and a comprehensive description of the models used can be found in Czembor & Vesk (2009). A brief outline is given here.

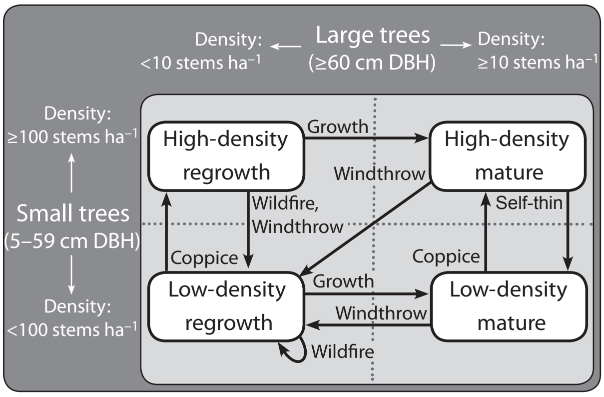

Fig. 1. The four vegetation states used in state-and-transition simulation models defined by the density and size of *Eucalyptus* trees. Arrows describe some of the common transition agents and how they affect the vegetation states.

The models simulated the dynamics of 1000 independent cells. Models began with a set of initial conditions that dictated the proportion of cells in each vegetation state. These conditions reflected the probable current state of real-world Box-Ironbark woodlands and were generated by averaging estimates provided by experts during interviews. We used the average of all expert estimates rather than different initial condition estimates for each expert's model to allow model predictions to be compared through time. Experts agreed that current stands typically have high densities of small stems and lack large trees, meaning the current woodlands are on average 71·2% high-density regrowth, while the remaining proportions are 10·3% low-density regrowth, 12·5% high-density mature and 6% low-density mature. We also considered two other sets of initial conditions that reflected potential historical conditions, but previous analyses indicated that initial conditions have relatively little effect on long-term model predictions (Czembor 2009) and will not be pursued further here.

Each cell was randomly assigned an initial age sampled from a uniform distribution. Each expert defined the minimum and maximum ages so that the models reflected the multi-age structure of Box-Ironbark stands. Here, age is defined as the number of timesteps (years) a cell has been in its current state. Cell age is important for the VDDT models as some transitions can only occur within specific age ranges.

To model the alternative management strategies, we asked experts to determine the effects of each management strategy on vegetation states. Cells in each model were apportioned equally between management strategies. The VDDT model output includes the proportion of the simulated landscape in each of four vegetation states. However, managers are interested in maximising the amount of Box-Ironbark woodland that is considered 'low-density woodland' because this vegetation state is believed to provide high-quality habitat for many fauna species. Therefore, we only use the proportion of cells in low-density woodland for the current analyses. We ran each model for 150 timesteps because preliminary analyses showed most models stabilized after approximately 100 years.

INCORPORATING UNCERTAINTY

Vegetation dynamics development tool (VDDT) models mimic system stochasticity by varying the order that transitions occur over multiple Monte Carlo simulations. The list of possible transition agents with their corresponding probabilities is specified according to a cell's current vegetation state and age. At one timestep, for a given Monte Carlo simulation, VDDT randomizes the list of possible transition agents and arrays their probabilities consecutively, such that each probability covers a distinct interval. VDDT then sums the probabilities of the possible transition agents and draws a random number from a uniform distribution between zero and one. If the number drawn is greater than the summed probabilities of transition agents possible in that state, then no transition occurs. If the draw is less than the summed probabilities, then the transition agent whose distinct probability interval corresponds to the random number is selected and a transition occurs, changing the cell to the appropriate state (ESSA Technologies Ltd 2007). Cell ages may also be adjusted after some transitions. This process is repeated for every cell and iterated over each timestep for the specified number of Monte Carlo simulations; we relied on 100 simulations. With this process, no more than one transition agent may be selected for a cell at one timestep and variability in the rates of disturbances is not incorporated.

To incorporate between-expert uncertainty, each of the five experts independently provided their own parameter estimates for each transition and specified how each transition agent would affect vegetation change.

To account for within-expert uncertainty, 25 replicate model parameterizations – called projects – were generated for each expert. A complete set of transition agent parameters for each project was randomly drawn from beta distributions that summarized the uncertainty in the estimates for each transition probability (Vose 1996; Burgman 2005). Though we assumed that uncertainty would follow a beta distribution, the experts themselves did not directly provide these probability densities. People are often comfortable providing a best guess and extreme values (Burgman 2005) that can later be converted to probability distributions (Regan, Colyvan & Burgman 2002; Dorazio & Johnson 2003). Therefore, the experts provided all transition probabilities in the form of three-point parameter estimates, where the upper and lower values represented the plausible bounds and the middle values represented an estimate of the mean (Morgan & Henrion 1990a; Stainforth *et al.* 2005). However, we found fitting beta distributions to all three values was often computationally difficult: the upper estimates on some experts' models were orders of magnitude larger than the middle estimate and, when used to generate beta distributions, resulted in non-real numbers using our algorithm. As a compromise, we calculated the variance of the distributions using normal bounds of the middle and lower estimates:

$$\sigma^2 = \left(\frac{\mu - l}{z}\right)^2, \qquad \text{eqn 1}$$

where, σ^2 is the variance, μ is the middle estimate, l is the lower estimate, and z is a z-value derived from a normal distribution depending on the nominal (subjective) probability that the true parameter lies between the prescribed limits.

When calculating the variance, we used a z-value assuming that the prescribed limits provided by experts represented 90% confidence bounds. To test for sensitivity to this assumption, we repeated the analyses assuming 80% and 95% confidence bounds. The sensitivity to these assumptions was of interest because there is substantial evidence in the literature to suggest that experts commonly overestimate the probability that the true value lies within the intervals they provide (e.g. Teigen & Jørgensen 2005). Shape parameters (α and β) for the beta distributions were calculated using the mean, μ, and the variance imputed for a given confidence level, σ^2, for each parameter (McCarthy 2007):

$$\begin{aligned}\alpha &= \mu\left[\frac{\mu(1-\mu)}{\sigma^2} - 1\right]\\ \beta &= (1-\mu)\left[\mu\frac{(1-\mu)}{\sigma^2} - 1\right].\end{aligned} \qquad \text{eqn 2}$$

We completed one trial where 100 replicate model parameterizations were constructed for each expert using the 90% confidence level. We visually inspected probability density functions showing the proportion of cells in low-density mature for 100 replicates vs. 25 replicates and found that the distributions were virtually identical. Because the time required to construct and simulate replicates was extremely onerous, only 25 replicates were constructed for each expert.

VARIANCE COMPONENTS ANALYSIS

The VDDT model predictions for the proportion of cells in the low-density mature state were logit transformed for data analysis that required normally distributed data. We completed data analyses on VDDT model predictions using R version 2.6.2 (R Development Core Team 2008). Variance components analyses (VCAs) were used to quantify the amount of the total variance that could be explained by a particular part of the model error structure (Alho, Kangas & Kolehmainen 1996; Quinn & Keough 2002). Here, we used linear mixed-effect models in a maximum likelihood framework (Faraway 2006; Gelman & Hill 2007; Baayen 2008) to determine the relative contribution to the variation in the proportion of cells in the desired vegetation state, low-density mature, of the three sources of uncertainty. The R packages nlme (Pinheiro *et al.* 2008) and lme4 (Bates 2008) were used to fit the linear mixed-effect models:

$$\begin{aligned}\text{Area}_i &= \beta \cdot \text{init}_i + \eta^{\text{expt}}_{j[i]} + \eta^{\text{proj}}_{jk[i]} + \eta^{\text{sims}}_i, \quad \text{for } i = 1,\ldots,n\\ \eta^{\text{expt}}_j &\sim \text{N}\left(0, \sigma^2_{\text{expt}}\right), \quad \text{for } j = 1,\ldots,5\\ \eta^{\text{proj}}_{jk} &\sim \text{N}\left(0, \sigma^2_{\text{proj}}\right), \quad \text{for } j = 1,\ldots,5, \; k = 1,\ldots,25\\ \eta^{\text{sims}}_i &\sim \text{N}\left(0, \sigma^2_{\text{sims}}\right), \quad \text{for } i = 1,\ldots,n\end{aligned} \qquad \text{eqn 3}$$

Here, Area_i is the logit transformed proportion of cells in low-density mature for the i^{th} simulation, where $n = 37\,500$ simulations. The term β is a coefficient vector, length three, which combined with a 3 by n matrix of binary values init_i, indicating which of the initial conditions the simulation began with, forms the fixed part of the model. The random part of the model is made up of three sets of error terms, η_j^{expt}, η_{jk}^{proj} and η_i^{sims}. The parameter η_j^{expt} consists of $J = 5$ terms while η_{jk}^{proj} consists of $K = 25$ projects nested within j. The term η_i^{sims} is the overall model residuals due to variation between replicate simulations of each model parameterization. Each set of estimates for the error terms are normally distributed, centred on zero, with corresponding variances σ^2_{expt}, σ^2_{proj} and σ^2_{sims}.

We fit variance components models at timesteps between and including 10 and 150 years at 10-yearly intervals. To test for sensitivity to the assumption that within-expert uncertainty was described by 90% confidence bounds, all 15 different timestep models were also fit assuming 80% and 95% confidence bounds.

Posterior probability densities for each of the variance components parameters were estimated with Markov Chain Monte Carlo (MCMC) methods, using a Gibbs Sampler implemented in the R package coda (Plummer *et al.* 2008). The posterior density distributions describe the uncertainty in estimated magnitudes of the variance components. MCMC analyses were performed for the 150-year model. The samplers were run for 110 000 iterations. Convergence was assessed using the potential scale reduction factor based on three chains and a 10 000 iteration burn-in (Gelman *et al.* 2000). Posterior density distributions from one of the three chains were used to determine the uncertainty in variance parameter estimates. MCMC analyses were repeated for each of the three confidence bound levels separately.

Results

MODEL PREDICTION

Most expert models predicted a slight increase from the initial 6% of the landscape in the low-density mature to an average of 7·5–11·9% after 150 years (Fig. 2; Experts 3, 4 and 5). There was large variation between-expert model parameterizations, which was highly influenced by Expert 2 and, to a lesser degree, Expert 1. Within-expert variation was relatively constant over timesteps for most expert models. Variation due to system stochasticity differed between-expert models and increased over time especially for Expert models 1 and 5.

VARIANCE COMPONENTS ESTIMATES

The overall variance in the proportion of low-density mature cells and the magnitude of each variance component was similar for each of the three within-expert confidence levels (Fig. 3). Total variance increased over time, peaked between 80 and 100 years, and then declined slightly as the models equilibrated near the end of simulations. After equilibrium, the majority of the overall variance estimate could be explained by the between-expert component ($\sigma^2_{\text{expt}} = 0{\cdot}73$). The variance due to system stochasticity ($\sigma^2_{\text{sims}} = 0{\cdot}12$) accounted for less than between-expert variance, but was greater than the within-expert variance component ($\sigma^2_{\text{proj}} = 0{\cdot}02$).

The estimate of the between-expert variance was the most uncertain component according to the posterior density distributions generated by MCMC sampling (Fig. 4a,b). The 95% Bayesian credible interval for between-expert variance using the 90% confidence bound was 0·261–6·013 (Fig. 4a,b), while the posterior density distribution for within-expert variance was more certain with a 95% credible interval of 0·010–0·018 (Fig. 4a,c). The variance due to system stochasticity had the tightest posterior density (Fig. 4a,d), as the uncertainty around each variance component reflected their relative sample sizes. The posterior density distribution for between-expert variance arises from the estimates of the intercepts for each expert. These intercepts appear to cluster for three experts (Experts 3–5), from which the intercepts for Experts 1 and 2 are offset (see Fig. 2). Interestingly, Experts 1 and 2 self-identified as having expertise primarily in ecology, while Experts 3–5 self-identified as having expertise in natural resource and forest management. Estimates for all variance components parameters were little changed whether expert transition agent estimates represented 80%, 90% or 95% confidence bounds (Table 2).

Discussion

CHARACTERIZING AND DEALING WITH UNCERTAINTY IN RESTORATION DECISIONS

Large between-expert variance was observed. This could be attributed to different professional experience in Box-Ironbark woodlands, such as seen between academic researchers and forestry professionals. This variation can be avoided in management decisions by subsampling similar individuals from the available pool of experts or by aggregating opinions about restoration actions through consensus, as is the norm with forest management steering committees. Yet, we would suggest that this could lead to biased predictions and thus potentially unreliably decisions.

BETWEEN-EXPERT UNCERTAINTY

Differences between experts

The large between-expert variance component reflects the disagreement about how modes of transitions and their probability of occurring affect vegetation states. The estimated

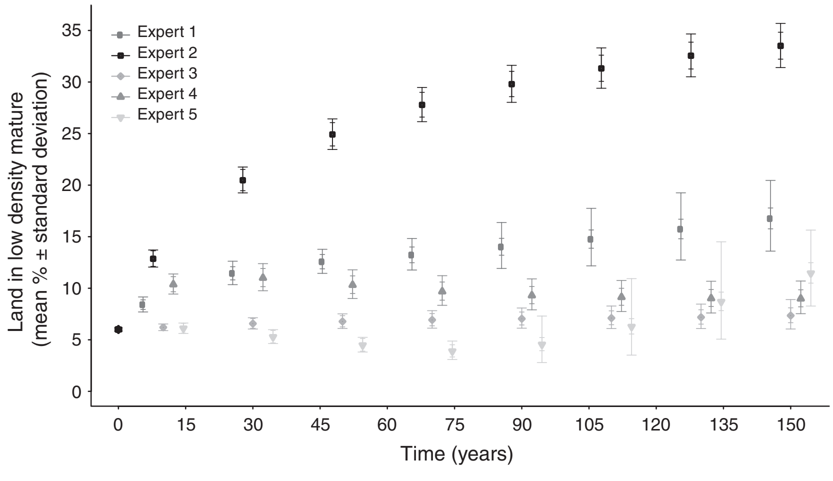

Fig. 2. The mean percent of cells in Low-density mature over time for each expert. Error bars represent standard deviations for within-expert variation (inner lines) and system stochastic simulations (outer lines).

© 2011 The Authors. Journal of Applied Ecology © 2011 British Ecological Society, *Journal of Applied Ecology*

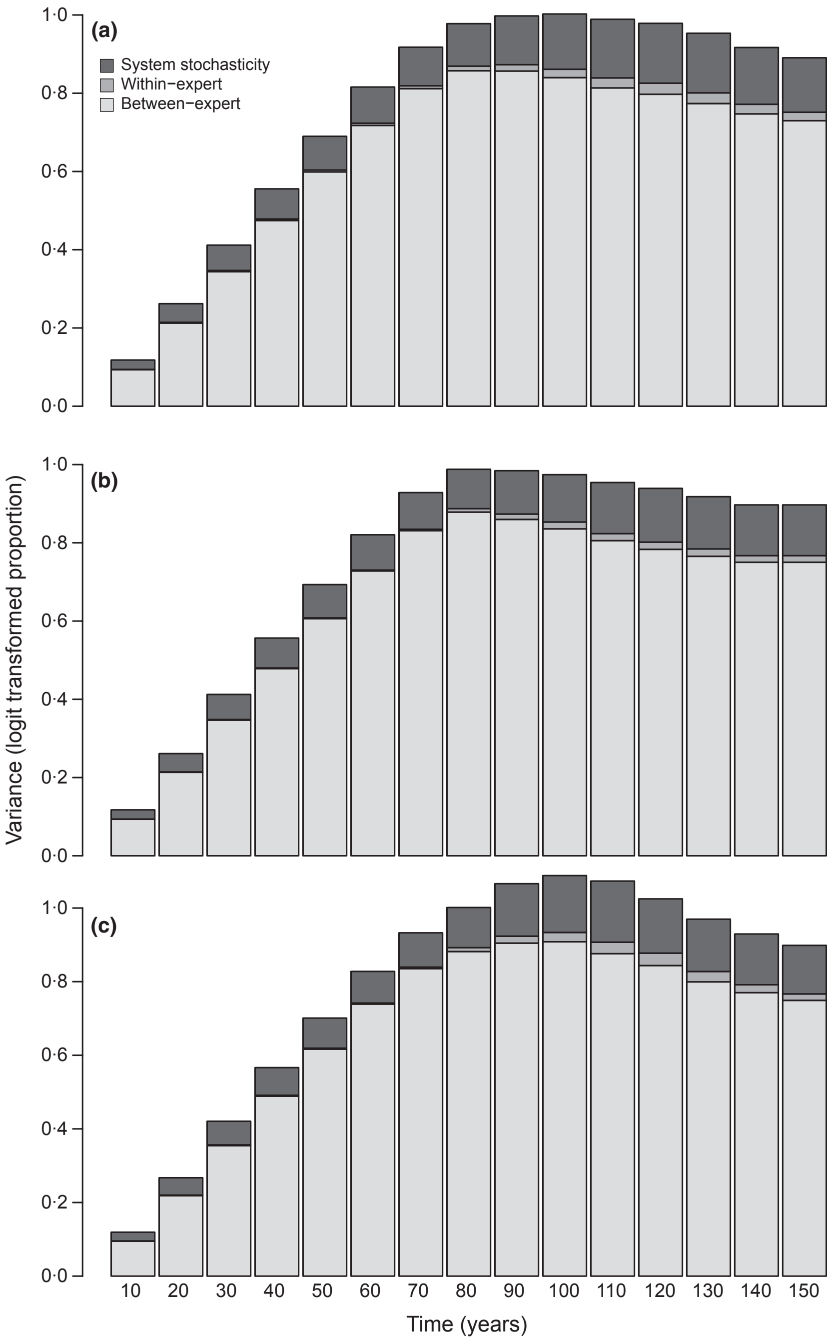

Fig. 3. Barplots showing the total variance due to between-expert uncertainty, within-expert uncertainty, and system stochasticity every 10 years for 150 years. Results are for the (a) 80%, (b) 90% and (c) 95% confidence bounds (The 95% confidence interval that corresponds to a logit transformed variance of 0·73 – between-expert estimate for the (a) 80% confidence bounds at 150 years – spans proportions from 0·03 to 0·46).

between-expert variance was less certain than other components because it was based on only five experts, while replication at the other levels of the hierarchical VCA was much greater. Nonetheless, the lower bound of the 95% credible interval for the between-expert parameter was greater than the upper bound of the other two parameters.

Large and uncertain between-expert uncertainty highlights the importance of allowing experts to provide different opinions. This conflicts with the practice of a majority of previously implemented expert opinion STSMs. Typically, modellers using STSMs employ consensus methods to achieve a single model parameterization (e.g. Hemstrom *et al.* 2002; McIntosh *et al.* 2003; Forbis *et al.* 2006; Wales, Suring & Hemstrom 2007; Bestelmeyer *et al.* 2009). Here, our experts provided very different parameter estimates and we had no *a priori* reason to weight the information given by experts.

Differences between experts are common, and can be due to complex system processes (Kangas *et al.* 1998). We found that differences between experts were caused by variation in both the rates and expected effects of transition agents on vegetation states (Appendix S1 Supporting information; Czembor & Vesk 2009). In a similar study using VCA to quantify sources of model uncertainty, Alho, Kangas & Kolehmainen (1996) found much variation between experts but, unlike our results, there was also substantial variance within-experts. Morgan, Pitelka & Shevliakova (2001) found that experts provided a diversity of opinions about the effects of climate change on forests.

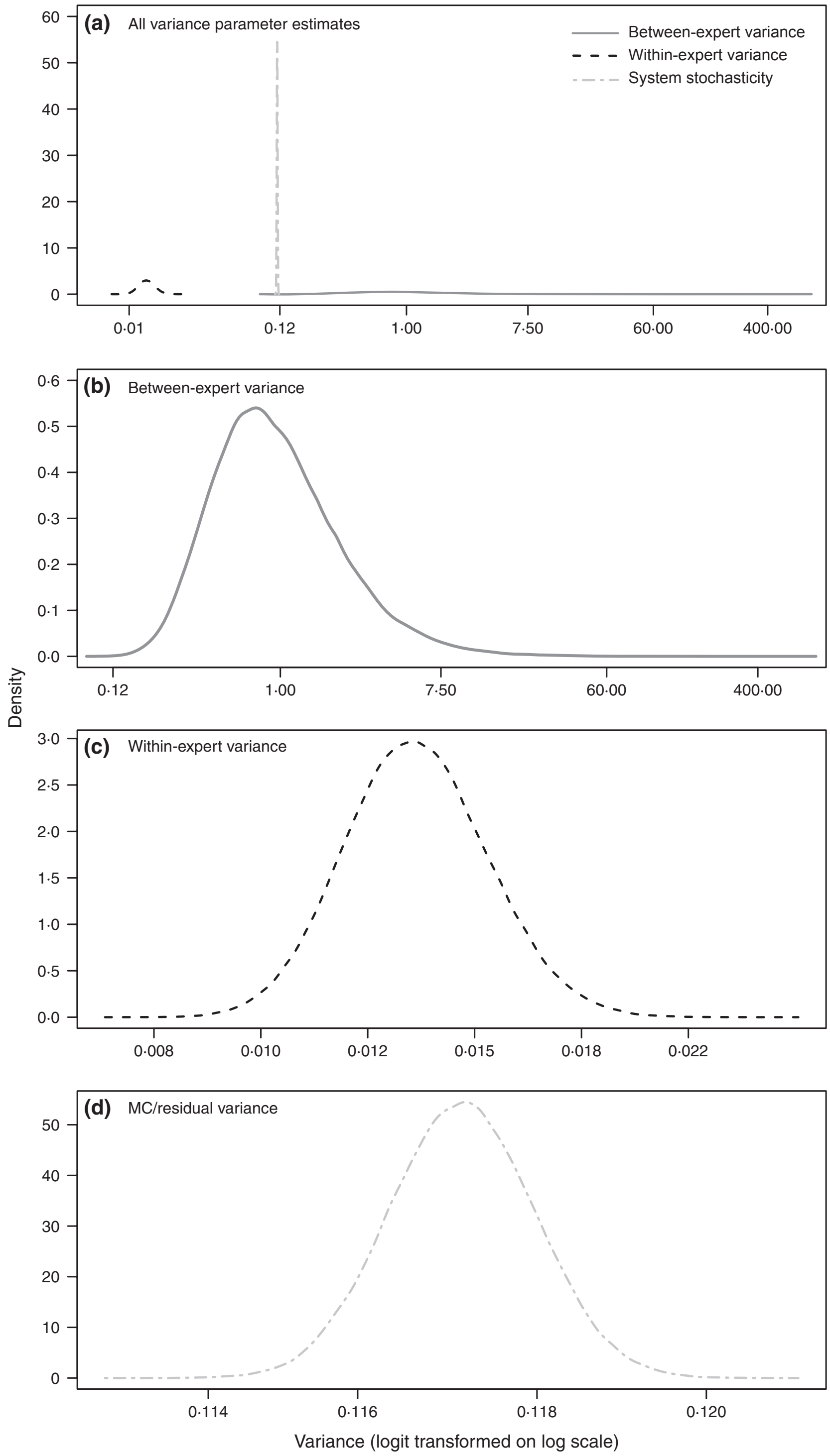

Fig. 4. Bayesian posterior probability density functions for (a) all VCA parameters, (b) between-expert variance, (c) within-expert variance and (d) system stochasticity. Results are for the 150-year timestep of the 90% confidence bound project.

Estimating variance

The large magnitude of between-expert uncertainty we estimated may be due to the small number of experts used. Estimating variance using a small sample size can be problematic because of the inability to estimate population parameters using few samples. Minimum sample sizes have been set at 30, 50 or anywhere from 4 to 25 times the number of predictors/variables (Green 1991; MacCallum *et al.* 2001). Often, the sample size is relatively small in published studies that rely on Bayesian statistics (e.g. 1 in O'Hagan 1998; 8 in Pellikka *et al.* 2005; 9 in Murray *et al.* 2009). Interestingly, many previous studies that rely on expert opinion do not report the number of experts consulted (e.g. Alho & Kangas 1997; O'Hagan 1998; Forbis *et al.* 2006; Ferguson *et al.* 2008; Bashari, Smith & Bosch 2009).

We acknowledge that the estimated variance for between-expert uncertainty is poorly determined as a result of the small

Table 2. The mean variance estimates and 95% Bayesian credible intervals (in square brackets) for the VCA parameters derived from posterior probability density distributions

Parameter	Confidence bound (%)	Variance estimate
	80	0·957 [0·263, 6·070]
Between-expert (σ^2_{expt})	90	0·955 [0·261, 6·013]
	95	0·983 [0·269, 6·234]
	80	0·023 [0·018, 0·030]
Within-expert (σ^2_{proj})	90	0·013 [0·010, 0·018]
	95	0·019 [0·014, 0·024]
System stochasticity/	80	0·140 [0·138, 0·142]
residual (σ^2_{sims})	90	0·117 [0·115, 0·119]
	95	0·132 [0·130, 0·134]

sample size, as illustrated by the wide 95% Bayesian credible intervals for this parameter. Increasing the number of experts may reduce between-expert uncertainty and increase our confidence in the parameter estimate if additional experts agreed, thus reducing the influence of outlier experts (Kangas *et al.* 1998). Methods have also been formulated to combine divergent expert models using Bayesian priors (e.g. McCarthy & Masters 2005; McAllister, Stanley & Starr 2010).

Due to problems estimating variance with small samples sizes, we do not endorse a sample size of five as a minimum number of experts for all forest restoration case studies. Unfortunately, as we exhausted the limited pool of available experts, we have no evidence to illustrate the minimum number of experts, and can only speculate about the effects of increasing the sample size in this case study. We believe that it will often be the case in real restoration examples that information is limited to few empirical data and/or appropriate experts. Most importantly, we find that when experts provide individual estimates, rather than being forced into consensus, their models present substantial differences in the effects of the transition agents and large uncertainty in the outcomes of different restoration strategies. Since differences between experts contribute most to model variance in this example, it is important to ascertain which parameters were most different between models so that additional data can be collected to determine the true rate and effect of parameters.

WITHIN-EXPERT AND STOCHASTIC UNCERTAINTY

We expected within-expert uncertainty to be caused by experts being uncertain about the true values of parameters. During interviews, experts who were highly uncertain provided very large bounds on estimates while those who were more confident provided smaller bounds. However, some of the variation in within-expert uncertainty may also be due to the conflation of within-expert uncertainty with variance in rates of transitions. We intended the within-expert variance component to describe only expert's cognitive uncertainty about fixed parameter values, but we cannot rule out the possibility that experts were also providing an estimate of the variability in parameter rates.

Despite the potential for conflation of parameter stochasticity and imperfect expert knowledge, the within-expert uncertainty component was relatively unimportant compared with between-expert uncertainty. However, experts can be poor judges of their own uncertainty (Cooke 1991) and can be prone to overconfidence, a common bias in subjective assessments (Morgan & Henrion 1990b; Ayyub 2001; Soll & Klayman 2004). We treated experts' upper and lower estimates as 80%, 90% or 95% confidence bounds. However, reasonable bounds are often much narrower than absolute bounds (Stainforth *et al.* 2005). For example, when asked for 98% or 90% confidence intervals, people often provide 60% and 50% confidence intervals respectively (Yaniv & Foster 1995; Teigen & Jørgensen 2005). This occurs because there is a cognitive trade-off between providing wide, less informative intervals and providing precise, informative estimates that may not include the true estimate (Yaniv & Foster 1995). We did not explicitly ask experts to provide estimates that reflected a specific confidence interval. Instead, we asked for the 'most reasonable' upper and lower limits. The small within-expert variance component may have been due to this linguistic ambiguity (Burgman 2005). However, the within-expert variance component was so small relative to between-expert uncertainty that even if we had assumed much wider beta distributions, the within-expert uncertainty would still be trivial compared with the between-expert variation. Perhaps, what is most interesting in this study is that experts were so confident in their own estimates, even though those estimates diverged so substantially from those of their peers.

The magnitude of the system stochasticity variance component varied between experts. Though this at first seems unintuitive it can be easily explained by a simple relationship with a given model parameterization and model stochasticity. When experts specify models with high transition rates, there can be a greater difference between simulations after the same number of timesteps than a model with a low transition rates.

OTHER SOURCES OF UNCERTAINTY

There are additional sources of uncertainty that can contribute to overall model uncertainty. Underspecificity is a type of linguistic uncertainty that occurs when there is unwanted generality (Regan, Colyvan & Burgman 2002), which might occur because the vegetation states we defined were very broad. For example, within the high-density mature vegetation state, small trees could have any diameter between 5–59 cm and any density greater than 100 stems ha^{-1}. This broad classification was necessary to minimize the number of vegetation states and the confusion in estimating parameters for many vegetation states. However, some experts found that these broad categories made it difficult to estimate the rates and effects of disturbances. Asking experts for unspecified bounds (discussed above) is also an example of underspecificity. Structural model uncertainty is another potential source that occurs when a model structure is employed while other equally plausible models are not accounted for (Wintle *et al.* 2003). Here, we ignored this type of uncertainty and only used a single model structure with multiple parameterizations. Structural uncertainty could have been incorporated by constructing different model structures and asking experts to parameterize each one, or by allowing each expert to construct their own model using unique vegetation states, similar to the approach taken by Smith, Parrott & Robertson (2008). However, parameterizing multiple model structures would have been too onerous for experts and using unique model structures for each expert would have compromised our analysis of within-expert uncertainty.

Conclusions

Here, we have successfully classified and quantified three important sources of uncertainty in STSMs based on expert opinion. Our variance components analyses revealed that estimates of between-expert uncertainty were larger and more uncertain than the other sources of uncertainty analysed. This emphasizes the importance of maximizing the sample of experts within available resources – noting that the numbers of appropriate experts to estimate the efficacy of restoration actions will often be small – and carefully selecting experts from different backgrounds. Our results highlight the potential for over-confidence if consensus methods are used for decision-making. Characterizing this important aspect of model uncertainty is a first step towards developing methods to reduce overall uncertainty. Reducing the impacts of inter-expert uncertainty may be achieved by better characterization, by exploring the reasons for divergence between experts, refinement of expert estimation via calibration, and training and weighting of experts with gold-standard data sets.

Acknowledgements

We would like to thank our experts: A. Bennett, J. Cuddy, R. Hateley, M. Keatley, J. Kellas, R. Price, A. Tolsma and L. Vearing. We thank J. Moore and three anonymous referees for constructive comments and criticisms. C. Czembor was supported by an Australian Postgraduate Award, the Ecological Society of Australia, and the Botany Foundation at the University of Melbourne. Ethics approval was granted on HREC application 0717119.

References

Alexander, J.S.A. (1997) *Threatened Hollow-Dependent Fauna in Box-Ironbark Forests of Victoria*. Department of Natural Resources and Environment, Victoria.

Alho, J.M. & Kangas, J. (1997) Analyzing uncertainties in experts' opinions of forest plan performance. *Forest Science*, **43**, 521–528.

Alho, J.M., Kangas, J. & Kolehmainen, O. (1996) Uncertainty in expert predictions of the ecological consequences of forest plans. *Applied Statistics*, **45**, 1–14.

Ayyub, B.M. (2001) *Elicitation of Expert Opinions for Uncertainty and Risks*. CRC Press, Boca Raton.

Baayen, R.H. (2008) *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge University Press, New York.

Bashari, H., Smith, C. & Bosch, O. (2009) Developing decision support tools for rangeland management by combining state and transition models and Bayesian belief networks. *Agricultural Systems*, **99**, 23–34.

Bates, D.M. (2008) *lme4: Linear Mixed-Effects Models using S4 Classes*. R package version 0.99875-9.

Ben-Haim, Y. (2001) *Information-Gap Decision Theory: Decisions under Severe Uncertainty*, 1st edn. Academic Press, San Diego.

Bestelmeyer, B.T., Tugel, A.J., Peacock Jr, G.L., Robinett, D.G., Shaver, P.L., Brown, J.R., Herrick, J.E., Sanchez, H. & Havstad, K.M. (2009) State-and-transition models for heterogeneous landscapes: a strategy for development and application. *Rangeland Ecology and Management*, **62**, 1–15.

Buckley, Y.M., Brockerhoff, E., Langer, L., Ledgard, N., North, H. & Rees, M. (2005) Slowing down a pine invasion despite uncertainty in demography and dispersal. *Journal of Applied Ecology*, **42**, 1020–1030.

Burgman, M. (2005) *Risks and Decisions for Conservation and Environmental Management*. Cambridge University Press, Cambridge.

Clemen, R.T. & Winkler, R.L. (1999) Combining probability distributions from experts in Risk Analysis. *Risk Analysis*, **19**, 187–203.

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

Czembor, C.A. (2009) *Incorporating Uncertainty into Expert Models for Management of Box-Ironbark Forests and Woodlands in Victoria, Australia*. Master's thesis, The University of Melbourne, Australia.

Czembor, C.A. & Vesk, P.A. (2009) Incorporating between-expert uncertainty into state-and-transition simulation models for forest restoration. *Forest Ecology and Management*, **259**, 165–175.

Department of Natural Resources and Environment. (1998) *Box-Ironbark Timber Assessment Project: Bendigo Forest Management Area and Pyrenees Range*. Department of Natural Resources and Environment, East Melbourne.

Dorazio, R.M. & Johnson, F.A. (2003) Bayesian inference and decision theory – A framework for decision making in natural resource management. *Ecological Applications*, **13**, 556–563.

Environment Conservation Council. (2001) *Box-Ironbark Forests and Woodlands Investigation*. Environment Conservation Council, Victoria.

ESSA Technologies Ltd. (2007) *Vegetation Dynamics Development Tool User Guide, Version 6.0*. ESSA Technologies Ltd., Vancouver.

Faraway, J.J. (2006) *Extending the Linear Model with R: Generalized Linear, Mixed Effects, and Nonparametric Regression Models*. Chapman and Hall/CRC, Boca Raton.

Ferguson, C.A., Carvalho, L., Scott, E.M., Bowman, A.W. & Kirika, A. (2008) Assessing ecological responses to environmental change using statistical models. *Journal of Applied Ecology*, **45**, 193–203.

Ferson, S. (1996) What Monte Carlo methods cannot do. *Human and Ecological Risk Assessment: An International Journal*, **2**, 990–1007.

Forbis, T.A., Provencher, L., Frid, L. & Medlyn, G. (2006) Great Basin Land Management planning using ecological modeling. *Environmental Management*, **38**, 62–83.

Forests Commission of Victoria. (1928) *Handbook of Forestry in Victoria*. Forests Commission of Victoria, Melbourne.

Gelman, A. & Hill, J. (2007) *Data Analysis using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York.

Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (2000) *Bayesian Data Analysis*, 2nd edn. Chapman & Hall/CRC, Boca Raton.

Green, S.B. (1991) How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research*, **26**, 499–510.

Hemstrom, M.A., Wisdom, M.J., Hann, W.J., Rowland, M.M., Wales, B.C. & Gravenmier, R.A. (2002) Sagebrush-Steppe vegetation dynamics and restoration potential in the interior Columbia Basin, U.S.A. *Conservation Biology*, **16**, 1243–1255.

Hurley, M.V., Rapaport, E.K. & Johnson, C.J. (2009) Utility of expert-based knowledge for predicting wildlife-vehicle collisions. *Journal of Wildlife Management*, **73**, 278–286.

Irvine, R.J., Fiorini, S., Yearley, S., McLeod, J.E., Turner, A., Armstrong, H., White, P.C.L. & van der Wal, R. (2009) Can managers inform models? Integrating local knowledge into models of red deer habitat use. *Journal of Applied Ecology*, **46**, 344–352.

Johnson, C.J. & Gillingham, M.P. (2004) Mapping uncertainty: sensitivity of wildlife habitat ratings to expert opinion. *Journal of Applied Ecology*, **41**, 1032–1041.

Kangas, A.S. & Kangas, J. (2004) Probability, possibility and evidence: approaches to consider risk and uncertainty in forestry decision analysis. *Forest Policy and Economics*, **6**, 169–188.

Kangas, J., Alho, J.M., Kolehmainen, O. & Mononen, A. (1998) Analyzing consistency of experts' judgments – case of assessing forest biodiversity. *Forest Science*, **44**, 610–617.

Kellas, J.D. (1991) Chapter 9. Management of the dry sclerophyll forests in Victoria. 2. Box-Ironbark forests. *Forest Management in Australia* (eds F.H. McKinnell, E.R. Hopkins & J.E.D. Fox), pp. 163–169. Beatty and Sons, Surrey.

Leskinen, P. & Kangas, J. (2001) Modelling future timber price development by using expert judgments and time series analysis. *Silva Fennica*, **35**, 93–102.

MacCallum, R.C., Widaman, K.F., Preacher, K.J. & Hong, S. (2001) Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*, **36**, 611–637.

Martin, T.G., Kuhnert, P.M., Mengersen, K. & Possingham, H.P. (2005) The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. *Ecological Applications*, **15**, 266–280.

McAllister, M.K., Stanley, R.D. & Starr, P. (2010) Using experiments and expert judgment to model catchability of Pacific rockfishes in trawl surveys, with application to bocaccio (*Sebastes paucispinis*) off British Columbia. *Fishery Bulletin*, **108**, 282–304.

McCarthy, M.A. (2007) *Bayesian Methods for Ecology*. Cambridge University Press, Cambridge.

McCarthy, M.A. & Burgman, M.A. (1995) Coping with uncertainty in forest wildlife planning. *Forest Ecology and Management*, **74**, 23–36.

McCarthy, M.A. & Masters, P. (2005) Profiting from prior information in Bayesian analyses of ecological data. *Journal of Applied Ecology*, **42**, 1012–1019.

McIntosh, B.S., Muetzelfeldt, R.I., Legg, C.J., Mazzoleni, S. & Csontos, P. (2003) Reasoning with direction and rate of change in vegetation state transition modelling. *Environmental Modelling and Software*, **18**, 915–927.

Morgan, M.G. & Henrion, M. (1990a) Chapter 7. Perfoming probability assessment. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. (eds M.G. Morgan & M. Henrion), pp. 141–168. Cambridge University Press, Cambridge.

Morgan, M.G. & Henrion, M. (1990b) Chapter 6. Human judgement about and with uncertainty. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. (eds M.G. Morgan & M. Henrion), pp. 102–136. Cambridge University Press, Cambridge.

Morgan, M.G., Pitelka, L.F. & Shevliakova, E. (2001) Elicitation of expert judgments of climate change impacts on forest ecosystems. *Climatic Change*, **49**, 279–307.

Muir, A.M., Edwards, S.A. & Dickens, M.J. (1995) *Description and Conservation Status of the Vegetation of the Box-Ironbark Ecosystem in Victoria*. Department of Conservation and Natural Resources, East Melbourne.

Murray, J.V., Goldizen, A.W., O'Leary, R.A., McAlpine, C.A., Possingham, H.P. & Choy, S.L. (2009) How useful is expert opinion for predicting the distribution of a species within and beyond the region of expertise? A case study using brush-tailed rock-wallabies *Petrogale penicillata*. *Journal of Applied Ecology*, **46**, 842–851.

Newman, L.A. (1961) *The Box-Ironbark Forests of Victoria, Australia*. Forest Commission Victoria, Melbourne.

Nicholson, E. & Possingham, H.P. (2007) Making conservation decision under uncertainty for the persistence of multiple species. *Ecological Applications*, **17**, 251–265.

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

O'Neill, S.J., Osborn, T.J., Hulme, M., Lorenzoni, I. & Watkinson, A.R. (2008) Using expert knowledge to assess uncertainties in future polar bear populations under climate change. *Journal of Applied Ecology*, **45**, 1649–1659.

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

Pellikka, J., Kuikka, S., Linden, H. & Varis, O. (2005) The role of game management in wildlife populations: uncertainty analysis of expert knowledge. *European Journal of Wildlife Research*, **51**, 48–59.

Pinheiro, J., Bates, D.M., DebRoy, S., Sarkar, D. & R Development Core Team. (2008) *nlme: Linear and Non-Linear Mixed-Effects Models*. R package version 3.1-86.

Plummer, M., Best, N., Cowles, K. & Vines, K. (2008) *Coda: Output Analysis and Diagnostics for Markov Chain Monte Carlo Simulations*. R package version 0.13-1.

Quinn, G. & Keough, M. (2002) *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, Cambridge, UK.

R Development Core Team. (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raulier, F., Pothier, D. & Bernier, P. (2003) Predicting the effect of thinning on growth of dense balsam fir stands using a process-based tree growth model. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **33**, 509–520.

Regan, H.M., Colyvan, M. & Burgman, M.A. (2002) A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications*, **12**, 618–628.

Smith, G.C., Parrott, D. & Robertson, P.A. (2008) Managing wildlife populations with uncertainty: cormorants *Phalacrocorax carbo*. *Journal of Applied Ecology*, **45**, 1675–1682.

Soderquist, T.R. (1999) *Tree Hollows of the Box-Ironbark Forest: Analyses of Ecological Data from the Box-Ironbark Timber Assessment in the Bendigo Forest Management Area and Pyrenees Range*. Department of Natural Resources and Environment, East Melbourne, Victoria.

Soll, J.B. & Klayman, J. (2004) Overconfidence in interval estimates. *Journal of Experimental Psychology*, **30**, 299–314.

Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G. & Burgman, M. (2010) Reducing overconfidence in the interval judgments of experts. *Risk Analysis*, **30**, 512–523.

Stainforth, D.A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D.J., Kettleborough, J.A., Knight, S., Martin, A., Murphy, J.M., Piani, C., Sexton, D., Smith, L.A., Spicer, R.A. & Allen, A.J.T.M.R. (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406.

Sutherland, W.J. (2006) Predicting the ecological consequences of environmental change: a review of the methods. *Journal of Applied Ecology*, **43**, 599–616.

Sutton, M. (2000) *Box-Ironbark Timber Resource Analysis*. Department of Natural Resources and Environment, Victoria.

Teigen, K.H. & Jørgensen, M. (2005) When 90% confidence intervals are 50% certain: On the credibility of credible intervals. *Applied Cognitive Psychology*, **19**, 455–475.

Vavra, M., Hemstrom, M.A. & Wisdom, M. (2007) Modeling the effects of herbivores on the abundance of forest overstory states using a state-transition approach in the upper Grande Ronde River Basin, Oregon, USA. *Landscape and Urban Planning*, **80**, 212–222.

Vose, D. (1996) *Quantitative Risk Analysis: A Guide to Monte Carlo Simulation Modelling*. John Wiley & Sons Ltd, West Sussex.

Wales, B.C., Suring, L.H. & Hemstrom, M.A. (2007) Modeling potential outcomes of fire and fuel management scenarios on the structure of forested habitats in northeast Oregon, USA. *Landscape and Urban Planning*, **80**, 223–236.

Walker, K.D., Evans, J.S. & MacIntosh, D. (2001) Use of expert judgment in exposure assessment: Part 1. Characterization of personal exposure to benzene. *Journal of Exposure Analysis and Environmental Epidemiology*, **11**, 308–322.

Walley, P. (2000) Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, **24**, 125–148.

Westoby, M., Walker, B. & Noy-Meir, I. (1989) Opportunistic management for rangelands not at equilibrium. *Journal of Range Management*, **42**, 266–274.

Wintle, B.A., McCarthy, M.A., Volinsky, C.T. & Kavanagh, R.P. (2003) The use of Bayesian model averaging to better represent uncertainty in ecological models. *Conservation Biology*, **17**, 1579–1590.

Yaniv, I. & Foster, D.P. (1995) Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology*, **124**, 424–432.

Received 17 October 2010; accepted 21 January 2011
Handling Editor: Andre Punt

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Transition agents specified by each expert.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.