

Research, part of a Special Feature on Assessing Risks to Wildlife

Characterizing Species at Risk II: Using Bayesian Belief Networks as Decision Support Tools to Determine Species Conservation Categories Under the Northwest Forest Plan

Bruce G. Marcot¹, Paul A. Hohenlohe², Steve Morey³, Russ Holmes¹, Randy Molina¹, Marianne C. Turley⁴, Mark H. Huff³, and John A. Laurence¹

ABSTRACT. We developed a set of decision-aiding models as Bayesian belief networks (BBNs) that represented a complex set of evaluation guidelines used to determine the appropriate conservation of hundreds of potentially rare species on federally-administered lands in the Pacific Northwest United States. The models were used in a structured assessment and paneling procedure as part of an adaptive management process that evaluated new scientific information under the Northwest Forest Plan. The models were not prescriptive but helped resource managers and specialists to evaluate complicated and at times conflicting conservation guidelines and to reduce bias and uncertainty in evaluating the scientific data. We concluded that applying the BBN modeling framework to complex and equivocal evaluation guidelines provided a set of clear, intuitive decision-aiding tools that greatly aided the species evaluation and conservation process.

Key Words: *Bayesian belief networks; decision models; expert panels; risk analysis; Northwest Forest Plan; species conservation.*

INTRODUCTION

Species conservation under the Northwest Forest Plan

The Northwest Forest Plan (NWFP) is a multiagency land and resource use plan covering over 9.7×10^6 ha (24×10^6 acres) of public lands, which are administered by federal agencies in the Pacific Northwest United States (USDA and USDI 1994). The NWFP was established in 1994 in part to conserve old forest ecosystems, including specific guidelines for protecting a diverse set of 403 rare and little-known species associated with late-successional and old-growth (LSOG) conifer and mixed-conifer hardwood forests.

The future viability of these species was found to be potentially at risk from multiple stressors, particularly disruption of their habitat and reduction and isolation of their populations caused by sundry land management activities on Federal public lands. The risks were seen to be potentially exacerbated due to the species' overall scarcity, adverse

demographic or genetic effects of isolation of small populations, narrow ecological amplitude, i.e., habitat specialization, and high sensitivity to adverse environmental change. At best, risk levels were poorly known because some species are difficult to locate and survey. Any of these conditions, many of which may be compounding stressors, could put a species at risk and in need of specific conservation management beyond the basic guidelines of the NWFP.

The daunting task of evaluating the status of, and effects of multiple stressors on, so many species of fungi, lichens, bryophytes, vascular plants, mollusks, and vertebrates associated with LSOG forest was given to the Survey and Manage (SM) species mitigation program under NWFP. The SM program was instituted to determine the status and conservation needs of each of these species in an adaptive management process (Molina et al. 2006). The SM program's components included designing and implementing field surveys (Molina et al. 2003, Edwards et al. 2004), developing species-habitat predictive models (Leshner 2005, Marcot 2006),

¹USDA Forest Service, ²Oregon State University, ³USDI Fish and Wildlife Service, ⁴USDI Bureau of Land Management

preparing and implementing conservation guidelines, and holding formal annual species reviews (ASRs).

Use of Decision Support Systems

To address the difficult task of evaluating effects of multiple stressors on multiple species, and of recommending appropriate conservation actions, we developed a decision support system codifying the mostly qualitative evaluation guidelines for the ASR. "Decision support system" is a general term for various sorts of models that can be used to inform managers on the implications of alternative actions. The models can be developed using simulations, decision trees, expert systems, fuzzy logic, or many other forms. Decision support systems using various kinds of models have been developed and applied to help optimize forest silviculture (Rojo and Orois 2005), assess forest fire risk (Iliadis 2005), and plan restoration of ecosystems (Reynolds and Hessburg 2005, Pieterse et al. 2002). Many other examples are available in the literature. Decision support systems used for environmental management are often developed in an ad hoc way without rigorous testing.

Decision support systems typically include models where a set of management guidelines or decisions and their effects, including costs or benefits, i.e., "utilities," are represented explicitly as probabilities. With a decision support system, the manager can evaluate the probabilities of effects of a decision. Whether the action is desirable or not depends in part on the management objectives and the decision maker's risk attitude, that is, risk-seeking, risk-neutral, or risk-intolerant. In general, decision support systems can work well to represent quantitative and qualitative evaluation or decision guidelines, and are particularly useful in difficult resource management contexts with multiple issues, assessment criteria, stakeholders, and values (Cleaves 1995, Zhu and Dale 2000). This was the circumstance we faced in the ASRs.

Decision support systems can be based on predicting ecological outcomes from conditions influenced by management decisions, or can focus primarily on representing and structuring a decision process itself. The former, ecological prediction models, were developed under the SM program for selected species and are discussed in the companion article (Marcot 2006). The latter, structuring the ASR decision process, is presented here.

Two main questions arose as we developed and applied a decision support system to help structure the ASR decision process: (1) How can a structured decision support system help scientific experts and managers implement a set of mostly qualitative assessment and decision guidelines, to produce coherent and well-documented recommendations for management? (2) Can structured models help scientists and managers evaluate rule-based criteria for classification and management decisions, to determine if they are consistent, coherent, and unbiased? Our objectives in this paper are to answer these questions by explaining the decision models that we created and used in the ASR process, and by discussing their potential use in related species conservation programs.

USING DECISION MODELS TO IMPLEMENT DECISION GUIDELINES

The annual species reviews (ASR) species evaluation process under the Survey and Manage (SM) program entailed applying a set of standards and guidelines (USDA and USDI 2001) to determine the appropriate conservation management category for each species. The guidelines consisted of an extensive and complex set of criteria used to determine persistence likelihood, i.e., viability, and rarity of each species. The guidelines specified 6 conservation management categories (A-F) for the SM species (Table 1), with a 7th possible outcome being to take the species off the SM list.

We crafted a set of decision-aiding models to help structure the interpretation and use of the evaluation guidelines in a consistent and repeatable way. The modeling structure we used was that of Bayesian belief networks (BBNs), which we discuss by way of a specific example. BBNs are models in which variables are connected by probability relations, and are displayed as nodes in a network diagram. The basic structure and statistical underpinning of BBNs are described by Marcot (2006) and Marcot et al. (*in press*). We used the BBN modeling shell Netica™ (Norsys, Inc.). Hereafter, these models are referred to as "BBN decision models."

The annual species review process

The ASR process was designed to compile, document, and evaluate new scientific information on rare and little-known, late-successional and old-

Table 1. Criteria and management of conservation categories of Survey and Manage species (USDA and USDI 2001). N/A = not applicable. The “three basic criteria” refer to occurrence in the Northwest Forest Plan area, association with late-successional and old-growth forests, and plan provision for persistence. “Off” means not to be included under the Survey and Manage species mitigation. Also see Table 2 for further criteria.

Species category	Criteria			Management		
	Meet three basic criteria?	Relative rarity	Predisturbance surveys	Management of known sites	Predisturbance surveys	Strategic surveys
A	yes	rare	practical	manage all known sites	required	required
B	yes	rare	not practical	manage all known sites	not required	required
C	yes	uncommon	practical	manage high-priority sites	required	required
D	yes	uncommon	not practical	manage high-priority sites	not required	required
E	information insufficient	rare	N/A	manage all known sites	not required	required
F	information insufficient	uncommon	N/A	not required	not required	required
Off	no	---	---	N/A	N/A	N/A

growth- (LSOG) associated species. The ASRs were conducted as a sequence of evaluation steps.

First, taxa experts compiled data on natural history, occurrence, and ecology of each SM species. Species were chosen for assessment in the ASR based on those determined to have substantial new information that changed our scientific understanding and might change management of the species. This information included, where known, species taxonomic status, geographical distribution, environmental and biophysical correlates, habitat associations, abundance, habitat and population trends, and major stressors and threats to persistence. Information was summarized for each species according to the complex set of evaluation parameters and was applied to the BBN decision models if it was determined that substantial new information had been documented since the last review.

Next, the taxa experts presented each species' information to an ASR evaluation panel consisting of eight members, i.e., four natural resource managers and four natural resource specialists. Separate panels were convened on faunal and floral species. The evaluation panels deliberated over the information, models, and guidelines, and used a Delphi paneling approach (Ayyub 2001) to recommend to decision makers the appropriate conservation management category for each species. Each panelist provided their individual beliefs via several stages of “voting” for conservation categories for each species, discussed further below.

The paneling process was designed to permit anonymity in the final voting. Overall, this structured decision approach helped optimize ASR management decisions by reducing bias and lending to legality, repeatability, thorough deliberation and

documentation, and efficiency of the evaluation process. In the final step, results of the votes from the evaluation panel were presented to decision makers who made final choices on the disposition of the conservation status for each species.

Applying the guidelines to determine species conservation categories

The guidelines (Table 2) were applied by the panelists, and programmed into the BBN decision models, in the following way to determine each species' appropriate conservation management category. A species was to be included under the SM mitigation if it met the first three evaluation categories, that is, if (1) the species' geographic range or habitat is within the Northwest Forest Plan (NWFP) area, (2) the species is closely associated with LSOG forest, and (3) the NWFP does not otherwise provide for the species' persistence. If any of these criteria failed, the species was not included under the SM mitigation. Next, to determine the specific conservation management category (Table 1) for each species, it was determined if (4) data are sufficient to determine what management is needed for species persistence, (5) surveys are practical to conduct, and (6) the species is relatively rare or uncommon.

Building the Bayesian belief network decision-aiding models from evaluation guidelines

To build the BBN decision-aiding models, we first constructed the overall model (Fig. 1) that expressed the application of the six evaluation categories and their management implications (Table 2). If a species failed to meet any of the first three evaluation categories, the outcome node, i.e., "SM species category" (Fig. 1) was forced to "off." That is, the species failed the initial screening to be placed under the SM mitigation, and no further evaluation of the remaining categories was necessary. If at least some nonzero probability of not being off the SM mitigation resulted, then the remaining evaluation categories applied.

In this way, we essentially translated the written guidelines as Boolean criteria, linked by standard "and/or/not/Xor" Boolean functions, and provided those as deterministic outcomes in the BBN conditional probability tables. A deterministic

outcome is one in which only a single condition results from a combination of inputs. The alternative is chance outcomes in which nonzero probabilities are denoted for >1 possible result.

For example, in the overall outcome model (Fig. 1), if geographic range was "in," LSOG association was "yes," and plan provision for persistence was "no," then the species category was not, i.e., anything but "off" the list. Continuing this example, if data were "insufficient," surveys were "practical," and the species was "rare," then the outcome would be species category E (see line 6 in Table 3). A portion of the extensive conditional probability table for this node is shown in Table 3. Each of the noninput nodes in all the BBN models had similarly structured deterministic outcome tables.

We used deterministic outcomes in the BBN nodes because the guidelines did not specify otherwise. However, the power and utility of such a decision model came when there was uncertainty about any of the evaluation criteria, and in combining effects of the criteria in determining appropriate species conservation categories. For example, if there was complete uncertainty about any of the criteria, then the probabilities associated with their alternative states were equally divided, i.e., uniform probabilities denote maximum uncertainty. This could result in nonzero probabilities for more than one conservation category, which the evaluation panelists then considered in their deliberations. Then, these probabilities were combined in the BBN using standard Bayesian learning to calculate the posterior probability of each species conservation category. More traditional look-up tables or a simple set of hierarchical classification or decision rules would not provide this capability.

We built individual BBN submodels for each of the six main evaluation categories (Table 2) to show how the specific evaluation criteria could be applied. The scientific data on each species were used to specify the states of all pertinent input nodes in each BBN submodel (Figs. 2–8), and then the results of each evaluation category were input into the overall model (Fig. 1), and the degree to which the alternative management category(s) pertained was shown.

We created each submodel based on deterministic Boolean representations of the published guidelines. For example, evaluation category 1, geographic range, was evaluated by restating the

Table 2. Guidelines for evaluating species persistence and late successional and old-growth forest (LSOG) association under the Northwest Forest Plan (NWFP), taken from USDA and USDI (2001), with reference to the Bayesian belief network (BBN) decision models. See Fig. 1 for how the overall BBN decision model combined these six evaluation categories. BLM = USDI Bureau of Land Management, and FEMAT = Forest Ecosystem Management Assessment Team.

Evaluation category	Guideline description with references to node numbers in the BBN submodel	BBN model figures
1. Geographic range	The species must occur within the Northwest Forest Plan area (1), or occur close to the NWFP area (2) and have potentially suitable habitat within the NWFP area (3).	Fig. 2
2. LSOG association	<p>A species is considered to be closely associated with late-successional and old-growth forests if it met at least one of the following criteria:</p> <ul style="list-style-type: none"> • The species is significantly more abundant in late-successional and old-growth forest than in young forest, in any part of its range (1). • The species shows association with late-successional and old-growth forest and may reach highest abundance there (2) and the species requires habitat components that are contributed by late successional and old-growth forest (3). • The species is associated with late-successional and old-growth forest, based on field study (2) and is on a federal U.S. Fish and Wildlife Service list (4d/5a) or state threatened or endangered list (4c); the U.S. Fish and Wildlife Service candidate species list (4a); a BLM or Forest Service special status species list in Oregon, Washington, or California (4b); or is listed by the States of Washington, Oregon, or California as a species of special concern or as a sensitive species (4c). • Field data are inadequate to measure strength of association with late-successional and old-growth forest (5c); the species is listed as a federal U.S. Fish and Wildlife Service threatened and endangered species (4d/5a); and the FEMAT suspected, or the panel doing the final placement in Species Review Process suspects, that it is associated with late-successional and old-growth forest (5b). 	Fig. 3
3. Plan provides for persistence	<p>The reserve system and other Standards and Guidelines of the NWFP do not appear to provide for a reasonable assurance of species persistence. Criteria indicating a concern for persistence, i.e., one or more of the following criteria must apply:</p> <ul style="list-style-type: none"> • Low-to-moderate number of likely extant known sites/records in all or part of a species range (1); • Low-to-moderate number of individuals (2); • Low-to-moderate number of individuals at most sites or in most populations (3); • Very-limited to somewhat-limited range (5); • Distribution within habitat is spotty or unpredictable in at least part of its range (6); and 	Fig. 4

-
4. Data sufficiency Information is insufficient to determine whether survey and manage basic criteria are met (1), or to determine what management is needed for a reasonable assurance of species persistence (2). Fig. 5
5. Practicality of survey Surveys are considered “practical” if all of the following criteria apply: Fig. 6
- The taxon appears annually or predictably, producing identifying structures that are visible for a predictable and reasonably long time (1);
 - The taxon is not so minuscule or cryptic as to be barely visible (2);
 - The taxon can authoritatively be identified by more than a few experts, or the number of available experts is not so limited that it would be impossible to accomplish all surveys or identifications for all proposed habitat-disturbing activities in the NWFP area needing identification within the normal planning period for the activity (3);
 - The taxon can be readily distinguished in the field and needs no more than simple laboratory or office examination to confirm its identification (4);
 - Surveys do not require unacceptable safety (5a) or species risks (5b);
 - Surveys can be completed in two field seasons (approximately 7-18 mo) (6); and
 - Credible survey methods for the taxon are known or can be developed within a reasonable time period, i.e., approximately 1 yr (7).
- 6a. Relative rarity The species is relatively rare and all known sites or population areas are likely to be necessary to provide reasonable assurance of species persistence, as indicated by one or more of the following: Fig. 7
- Species poorly distributed within its range or habitat (1,2,3);
 - Limited dispersal capability on federal lands (4);
 - Reproduction or survival not sufficient (5);
 - Low number of likely extant sites/records on federal lands indicates rarity (6).
 - Limited number of individuals per site (7).
 - Declining population trends (8)
 - Low number of sites in reserves (9) or low likelihood of sites or habitat in reserves (10).
 - Highly specialized habitat requirements (narrow ecological amplitude) (11);
 - Declining habitat trend (12);
 - Dispersal capability limited relative to federal habitat;
 - Habitat fragmentation that causes genetic isolation (13);
 - Microsite habitat limited (14); and
-

6b. Relative uncommonness

The species is relatively uncommon rather than rare, and not all known sites or population areas are likely to be necessary for reasonable assurance of persistence, as indicated by one or more of the following: Fig. 8

- A higher number of likely extant sites/records does not indicate rarity of the species (1);
- Low-to-high number of individuals/site (2);
- Less restricted distribution pattern relative to range or potential habitat (3);
- Moderate-to-broad ecological amplitude (4);
- Moderate-to-high likelihood of sites in reserves (5); and
- Populations or habitats are stable (6).

guidelines for this category (Table 2) as a deterministic Boolean formula: geographic range = "in" if the taxon range is within the NWFP area or the species occurs close to the NWFP area, and suitable habitat occurs within the NWFP area. It was relatively simple to then express such statements as deterministically-based BBN models (e.g., Fig. 2).

We also fully documented each node in each BBN by relating back to the specific published guidelines they represented. For example, the input nodes in the BBN submodel pertaining to LSOG association (Fig. 3) were documented according to the specific guideline they each represented for LSOG association (Table 2). Further, as each species was evaluated, the resulting BBN models were saved as part of the public administrative record. This helped provide clear documentation on how each species was assessed.

The BBN models clearly showed the relation of each piece of scientific information to the interpretation of species rarity and persistence status, and thus to alternative conservation management categories for each species. We first tried several other approaches, such as fuzzy logic models, but settled on BBNs because they are intuitive, easy to operate, and clearly display effects of data or the lack of data on potential decision outcomes. We emphasized that the models were not intended to be predictive, or to dictate the outcome decision, but rather to help consolidate information on the species, guide panel discussion and deliberation, and prompt consistency

in how information was considered and how the guidelines were applied as the evaluation panelists each reached their individual conclusions.

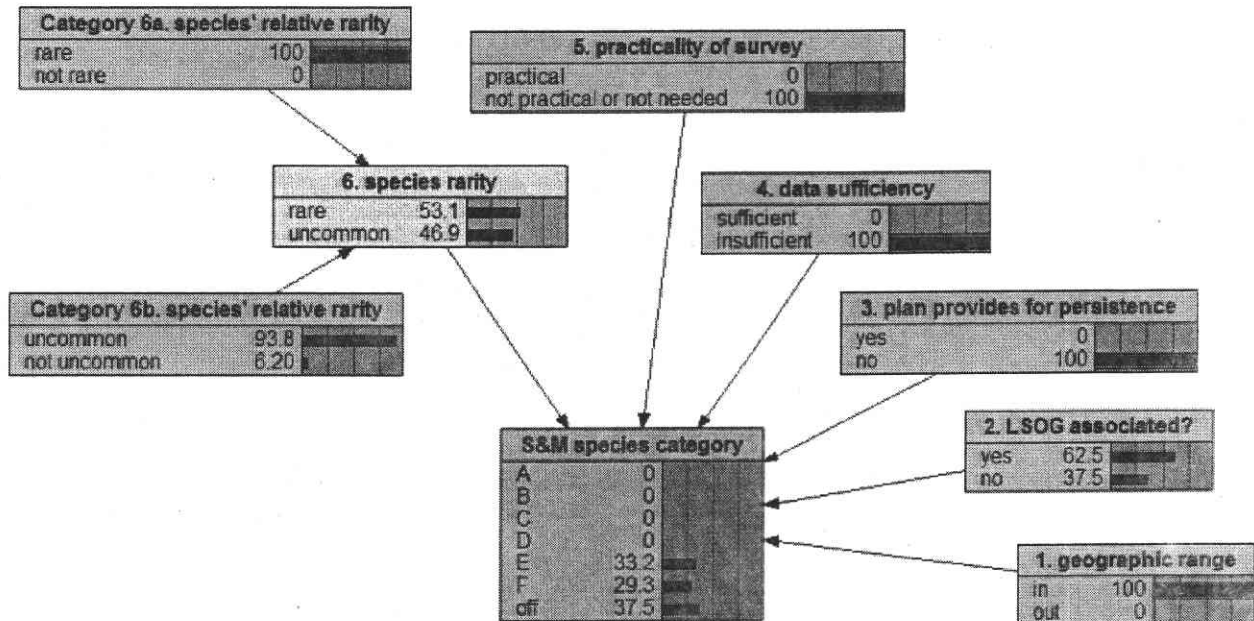
We conducted sensitivity analyses of the BBN decision models to determine the degree to which the outcome nodes were influenced by the input nodes. This helped ensure that the models were constructed in accordance with the evaluation guidelines, and to help the panelists understand the relative influence of each parameter on the outcomes. Formulae used in calculations of sensitivity are presented by Marcot 2006.

Validating the models

The BBN decision models essentially were representations of the evaluation guidelines, not ecological prediction tools. As such, their validation consisted of ensuring that all evaluation guidelines were fully and correctly represented. We did this in a group setting among ourselves and with peer review by several experts from the NWFP. This differed significantly from validation of species prediction models developed for other segments of the NWFP, which entailed comparing model predictions of species presence with field observations (see Marcot 2006).

However, this also raised the interesting question of what constitutes a valid decision model. We realized that vagueness, ambiguity, and generality of the

Fig. 1. Overall Bayesian belief network (BBN) decision model used by evaluation panels to guide their recommendations on appropriate conservation management categories of Survey and Manage (SM) species under the Northwest Forest Plan of federal agencies in Pacific Northwest, United States. The model displays potential species conservation management categories as a function of six evaluation categories. Figs. 1–8 are parameterized for a rare lichen, *Fuscopannaria saubinetii*, and gray boxes denote data provided by taxa experts.



criteria in the guidelines meant that some latitude needed to be afforded to the evaluation panelists. Thus, the BBN decision models would be more appropriately and validly used as evaluation tools to prompt and guide discussion and interpretation of the data and guidelines, rather than to dictate final decisions.

Using the Bayesian belief network decision models in the paneling process

Running the BBN decision-aiding models entailed using the data from the taxa experts for each species to specify the states of the input nodes. The models thus provided an interpretation of the scientific data in terms of which conservation management categories might pertain to each species, given the

data and application of the evaluation guidelines. Model outcomes and the scientific data were provided to the ASR evaluation panelists for their consideration.

Specifically, the voting process used by the ASR evaluation panels was structured as four stages. Stage 1 consisted of the panelists studying the species data given by the species experts, inspecting the BBN model results, and providing an initial vote on potentially appropriate conservation categories for each species. This initial vote took the form of each panelist spreading 100 voting points across one or more of seven possible outcomes, i.e., the six species categories plus the option of removing or not including the species from the SM mitigation species list. The option of spreading points across more than one alternative outcome allowed

Table 3. Conditional probability table for the outcome node “SM species category” shown in the overall decision model in Fig. 1. See text for explanation of evaluation categories and SM species outcome categories. LSOG = late-successional and old-growth forest. Shown here is only a portion of the very long table.

Input nodes						Outcome node
Evaluation Category 1 – Geographic range	Evaluation Category 2 – LSOG association	Evaluation Category 3 – Plan provides for persistence	Evaluation Category 4 – Data sufficiency	Evaluation Category 5 – Practicality of surveys	Evaluation Category 6 – Species rarity	SM species category
in	yes	yes	sufficient	practical	rare	off
in	yes	yes	insufficient	practical	rare	off
out	yes	yes	sufficient	practical	rare	off
out	yes	yes	insufficient	practical	rare	off
in	yes	no	sufficient	practical	rare	A
in	yes	no	insufficient	practical	rare	E
out	yes	no	sufficient	practical	rare	off
:	:	:	:	:	:	:

panelists to express individual uncertainty about their choices.

In Stage 2, each panelist revealed their votes and their rationale in a structured disclosure process. Tallies of the voting points among all panelists were recorded on a flip-chart for all to see. In Stage 3, a moderated discussion and question-and-answer period allowed the panelists to query the species experts and each other on points of uncertainty.

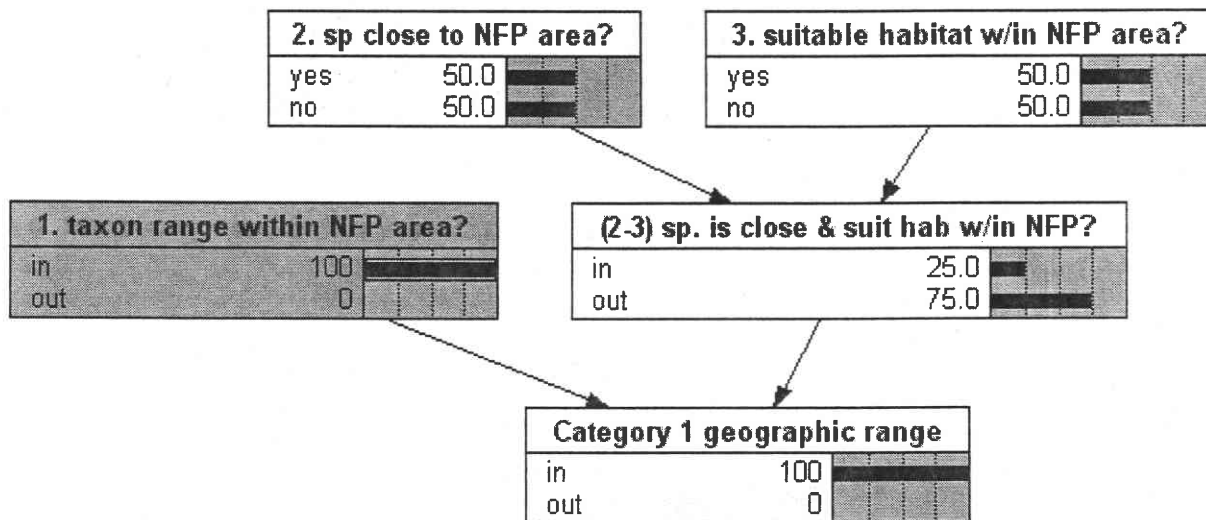
Stage 4 consisted of a subsequent silent vote by each evaluation panel member, but this vote consisted not of spreading points but more simply denoting the single, most appropriate conservation management category for each species because management wanted one single recommendation from each panelist. Voting results again were tallied for all to see, but panelist identities remained anonymous. Finally, in case of tied outcomes, in which equal numbers of panelists voted for different conservation categories with no clear majority outcome, another round of discussion, questioning,

and voting would take place. The distribution of number of votes by conservation category was recorded by a scribe, and this information was later presented to a decision-making body.

The main role of the BBN decision models was in the initial vote and discussions by the panelists in Stage 1, but they could be used by the panelists in any of the four stages such as for exploring possible influence on conservation categories from different interpretations of the scientific data. The models helped ensure that scientific data were treated consistently by the panelists in the context of the evaluation guidelines. The models also helped clarify the most sensitive factors affecting potential species conservation management categories and identified where there was limited or insufficient information and uncertainty. This helped guide some of the questions posed in Stage 3.

The BBN decision models, used to inform the panelists in their initial vote, usually led to multiple possible outcomes, i.e., species conservation

Fig. 2. Bayesian belief network (BBN) submodel for determining evaluation category 1, geographic range conditions for the species in relation to the Northwest Forest Plan (NWFP). See Table 2 for Record of Decision guidelines used to develop this model.



categories, for a given species. The multiple outcomes reflected uncertainty in the input criteria. The panelists were directed to understand the basis for the inputs and their implications for the uncertainty in outcomes.

Comparing Bayesian belief network model outcomes to the species' ratings made by the panelists

Panelists were free to deviate from the outcomes of the BBN decision models, but needed to document why they did so. The guidelines had enough generality and ambiguity, and the data were variable enough among species, to permit such latitude of interpretation by the panelists. This latitude was also consistent with the spirit of using the BBN models as decision-aiding tools to help guide discussion, and not to dictate the outcome.

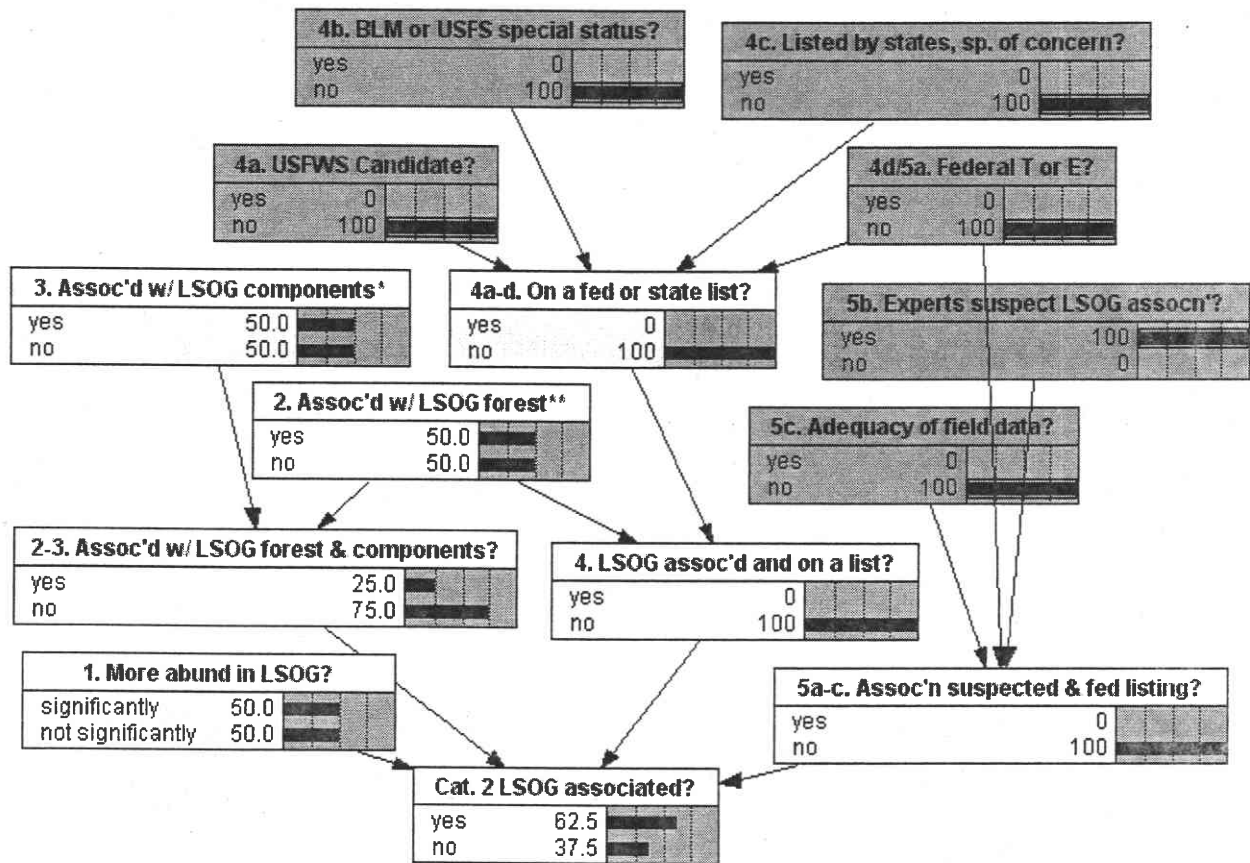
To assess the use and utility of the BBN decision models in the evaluation panels, we tracked and compared model outcomes for each species to each

round of voting by the panelists. We did not expect that the panelists would vote strictly according to the model outcomes. We interpreted the distribution of model and panelist voting outcomes as measures of uncertainty, i.e., spread of model probabilities or of panelist votes among conservation categories, and evaluated the variation in votes among panelists and between panelist and model outcomes by tallying numbers of discrete combinations of the conservation categories.

USING THE BAYESIAN BELIEF NETWORK TO EVALUATE DECISION CRITERIA

Our experience with the Bayesian belief network (BBN) decision models suggested that structured models can help evaluate rule-based guidelines for classifying species conservation categories and management decisions. The models greatly helped to determine if the guidelines were flawed, and provided a way of representing inconsistent guidelines in a consistent structure.

Fig. 3. Bayesian belief network (BBN) submodel for determining evaluation category 2, late-successional and old-growth forest (LSOG) association of the species. See Table 2 for Record of Decision guidelines used to develop this model.

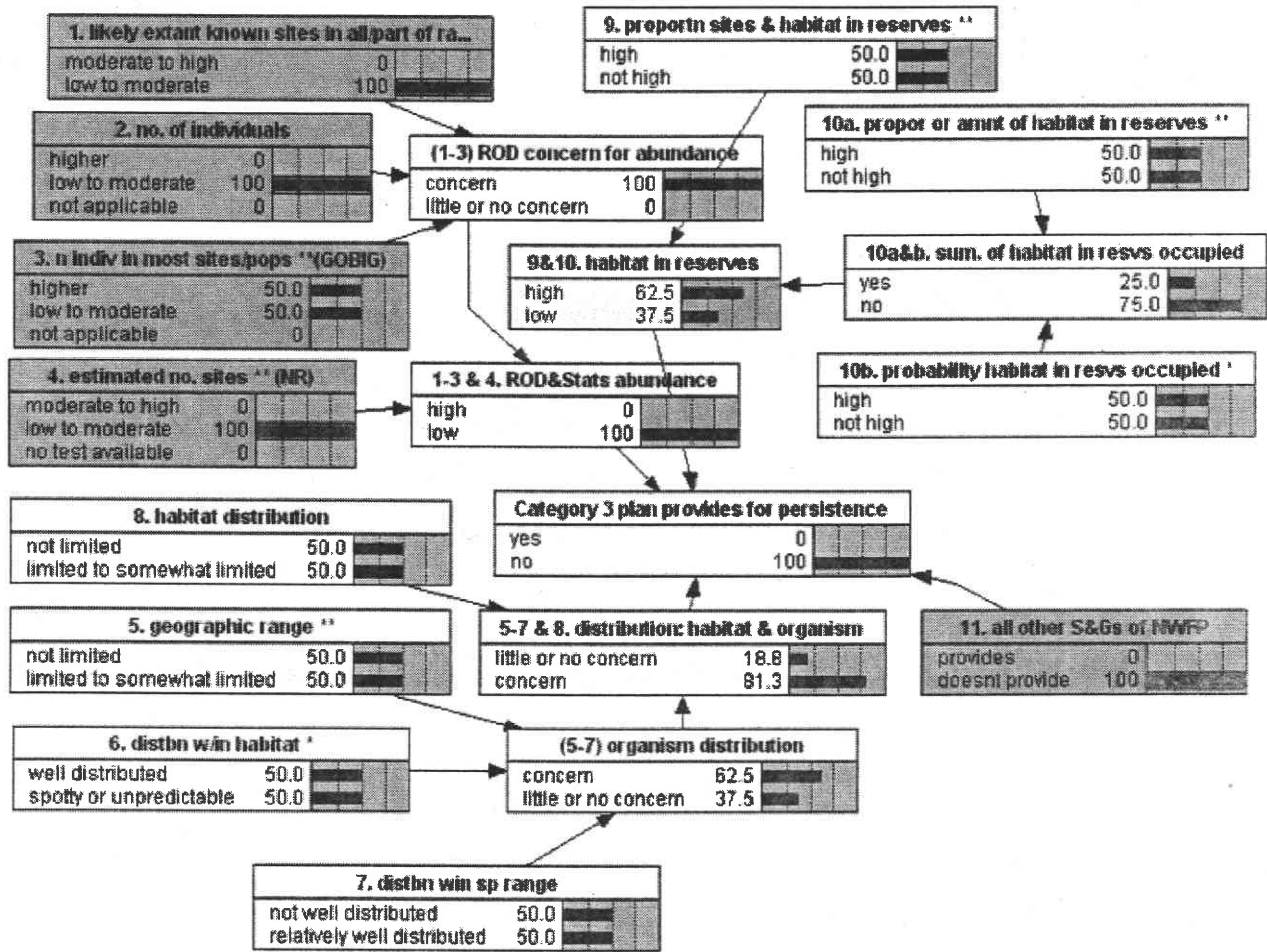


The published guidelines of some of the evaluation categories were very complex (Table 2). This required that we carefully interpreted and represented them as Boolean and then BBN formulations. This was true particularly with the guidelines for determining: species' late-successional and old-growth (LSOG) forest association (Fig. 3); whether the Northwest Forest Plan (NWFP) provides for species persistence (Fig. 4); and species rarity (Figs. 7 and 8). These guidelines were not just complex; some were inconsistent. We dealt with the complexity of these

guidelines in the submodels by using summary nodes to bring order to the extensive criteria.

For example, we used summary nodes in the submodel on species persistence (Fig. 4) to structure and simplify the combination of the 13 evaluation criteria into four summary categories: species statistics and abundance, habitat in reserve land allocations, e.g., late-successional forest reserves, distribution of habitats and the organism, and other guidelines of the NWFP. Summarizing many complex evaluation criteria in this way greatly

Fig. 4. Bayesian belief network (BBN) submodel for determining evaluation category 3, whether the Northwest Forest Plan provides for persistence of the species. See Table 2 for Record of Decision guidelines used to develop this model. GOBIG is a species occurrence database from California.



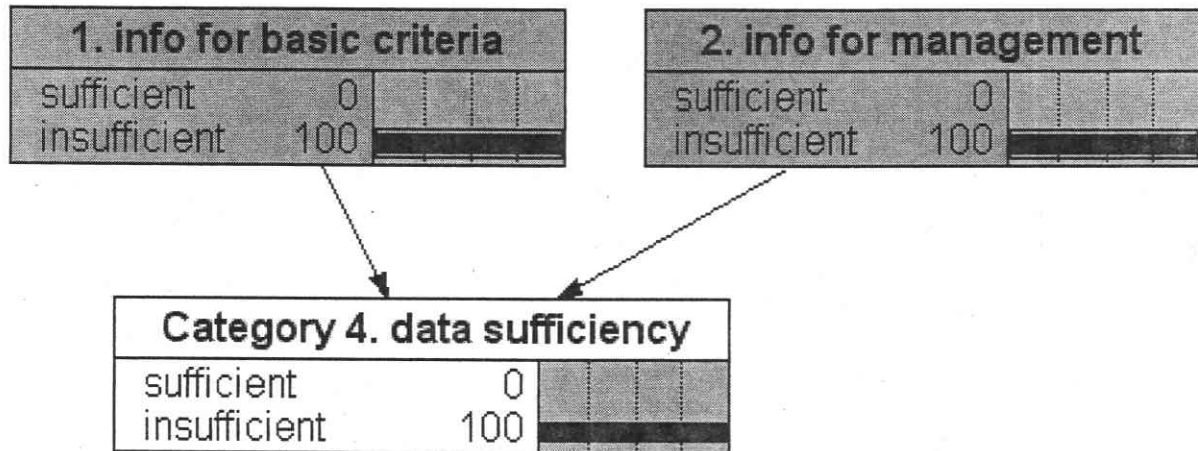
helped both the taxa experts and the evaluation panelists better understand the ecological implications and relationships of the data and the guidelines.

When developing the BBN decision models, we detected several inconsistencies and bias in logic in the guidelines. First, the guidelines presented separate and inconsistent criteria for determining whether a species is rare or uncommon (Table 2). A species in a given geographic setting ought to be either rare or uncommon; logically, it cannot be

both, as these were intended in the guidelines to be mutually exclusive conditions. However, in some cases, the guidelines allowed for a species to qualify for both conditions.

Second, the guidelines also seemed to bias the outcomes toward the "rare" condition by specifying far more criteria, any one of which would lead to a species being "rare," than for being "uncommon." No weights were specified in the guidelines for these criteria. With complete uncertainty about the criteria for "rare" and for "uncommon," the

Fig. 5. Bayesian belief network (BBN) submodel for determining evaluation category 4, data sufficiency. See Table 2 for Record of Decision guidelines used to develop this model.



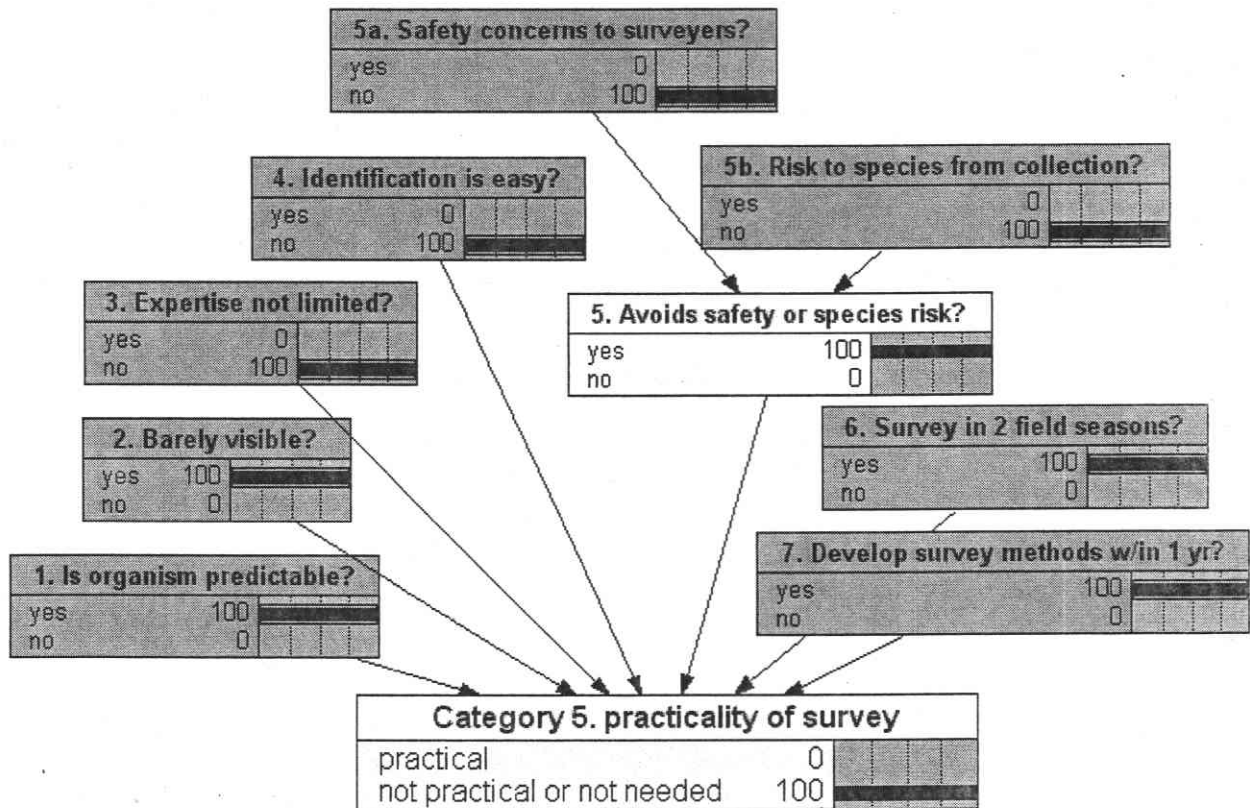
expected outcome was always tilted toward being “rare.” Thus, maximum likelihood model outcomes were never achieved for species conservation categories entailing “uncommon” categories C, D, and F (Table 1), although these categories often appeared as less likely model outcomes.

Third, many criteria determining species persistence and rarity were vague such as low number of sites, poor distribution, and specialization of habitat requirements. Vagueness in the criteria resulted in gray areas that were open to alternative interpretations. To bring consistency to the annual species review (ASR), the scientific data underlying such parameters were fully described by the taxa experts and the vague terms were clearly defined ecologically. Then, the evaluation panelists documented how they interpreted the species data in context of the guidelines. Thus, although the outcomes shown by the BBN decision models were one of many possible interpretations of applying information about species to the evaluation guidelines, the basis for the specific outcomes produced were consistently and well documented.

In general, we did not correct apparent inconsistencies or biases in the guidelines that would have entailed rewriting the guidelines, which we were not permitted to do. For example, we crafted the BBN decision models to explicitly show probabilities of both rare and uncommon status (Figs. 7 and 8), and the panelists were given the flexibility to consider the criteria for rarity in a more unbiased manner.

Overall, then, the BBN decision modeling process helped us to explain and refine selected parts of the guidelines that were inconsistent, incoherent, and biased. If these guidelines are revised, this modeling exercise can be used to clearly determine which guidelines were inconsistent, incoherent, and biased, and which of their specific criteria were ambiguous and vaguely specified. Further, the modeling process could be used to quickly represent a draft set of guidelines to test, and then correct, for faulty logic structures, inconsistency, incoherence, and bias.

Fig. 6. Bayesian belief network (BBN) submodel for determining evaluation category 5, practicality of survey. See Table 2 for Record of Decision guidelines used to develop this model.



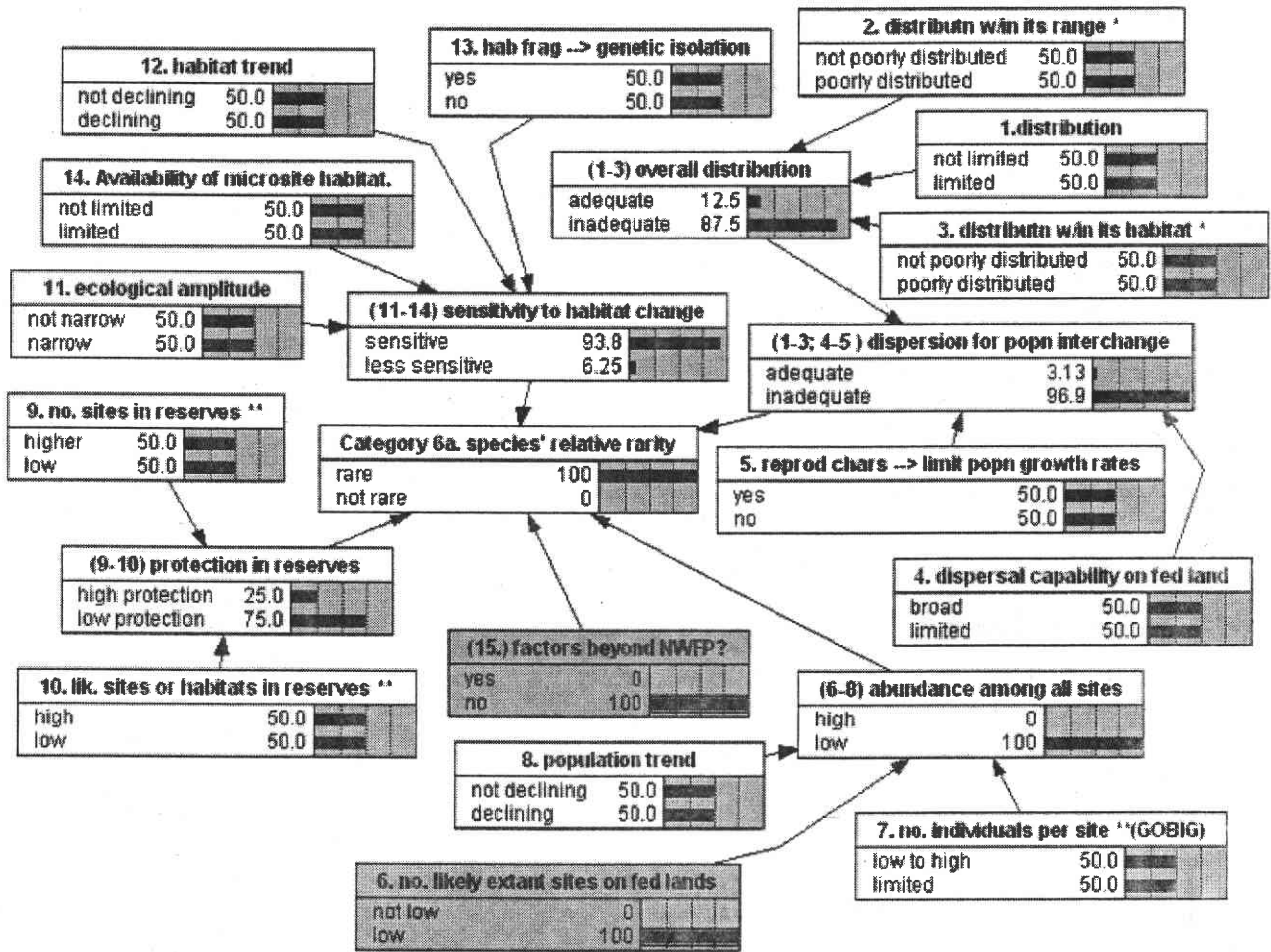
RESULTS

Bayesian belief network decision models: construction, use, and an example

The Bayesian belief network (BBN) decision model framework consisted of the one main model (Fig. 1) and seven submodels (Figs. 2–8). Structuring the BBN models as deterministic greatly simplified the model-building task, as we did not need to specify exact probability values, other than 0 and 1.0, in the conditional probability tables. A deterministic structure was consistent with the declarative form of the evaluation guidelines (Tables 1 and 2).

The models are shown in Figs. 1–8 as examples parameterized for one of the SM species evaluated by the plant annual species review (ASR) panel in 2003, a rare shingled lichen, *Fuscopannaria saubinetii* (Mont.) P. M. Jørg. This lichen is known from just a few locations in the Northwest Forest Plan (NWFP) area. It is found on bark, wood, or rock in moist or wet forest conditions (McCune and Geiser 1997). It was recently split from the genus *Pannaria*, so most of the known locations of this species were reexamined by the taxa experts who found them to be *Fuscopannaria pacifica*, a common species; fewer than five sites were *F. saubinetii*.

Fig. 7. Bayesian belief network (BBN) submodel for determining part of evaluation category 6, species relative rarity. See Table 2 for Record of Decision guidelines used to develop this model.

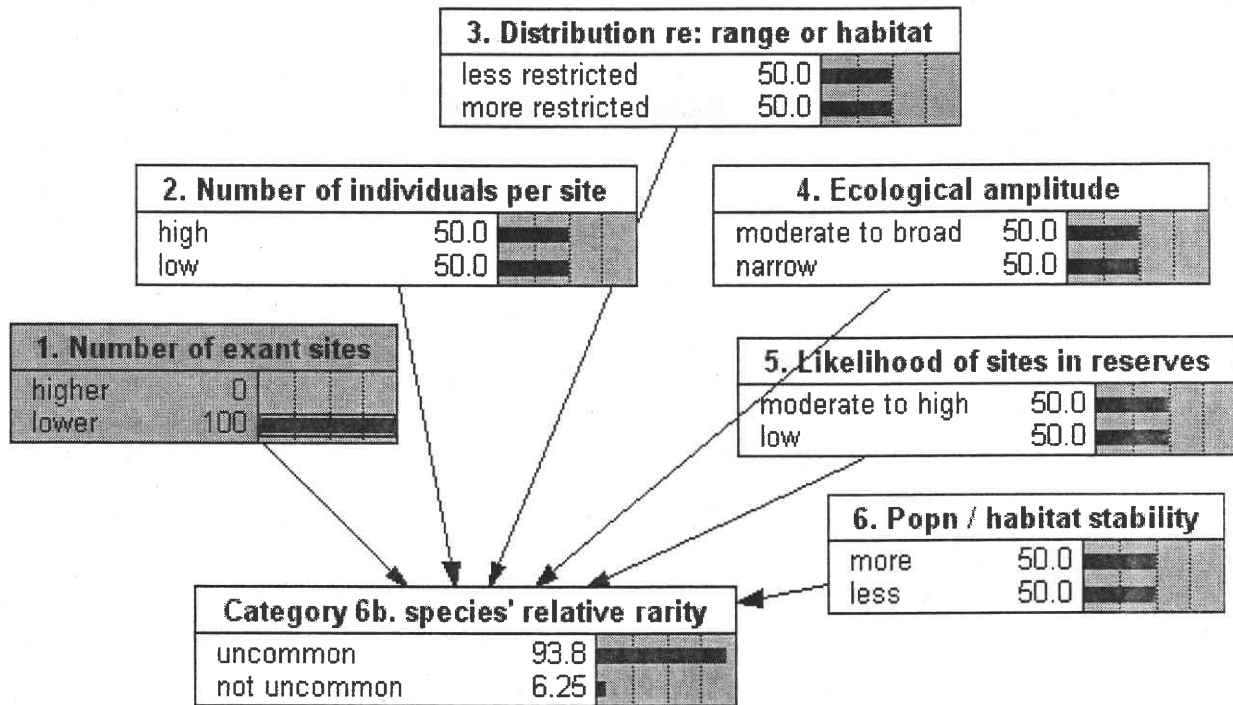


Applying the first three evaluation categories in the BBN decision model (Fig. 1) determined that *F. saubinetii* qualified for consideration under Survey and Manage (SM) mitigation. That is, it occurs within the geographic range of the NWFP (Fig. 2); it may be associated with late-successional and old-growth (LSOG) forests, even given uncertainty about its specific association with LSOG components and abundance in LSOG (Fig. 3); and its persistence likely is not otherwise provided by the NWFP, even given uncertainty about its distribution and occurrence in reserves (Fig. 4).

In the next set of evaluation categories, data seemed insufficient by which to evaluate the species' conservation and management requirements (Fig. 5); surveys did not seem practical particularly given the lack of identification keys to distinguish this species (Fig. 6); and the species seemed rare (Fig. 7) although there was much uncertainty over its distribution and abundance, which also qualified the species as potentially being uncommon (Fig. 8).

The combination of these evaluation categories in the overall BBN model suggested three possible conservation management categories for this

Fig. 8. Bayesian belief network (BBN) submodel for determining part of evaluation category 6, species relative uncommonness. See Table 2 for Record of Decision guidelines used to develop this model.



species consistent with the evaluation guidelines: E, F, and to be taken off the SM species mitigation list (Fig. 1). The “off” category appeared because of the uncertainty over LSOG forest association. The likelihoods of these three outcomes illustrated in the BBN model were approximately equal.

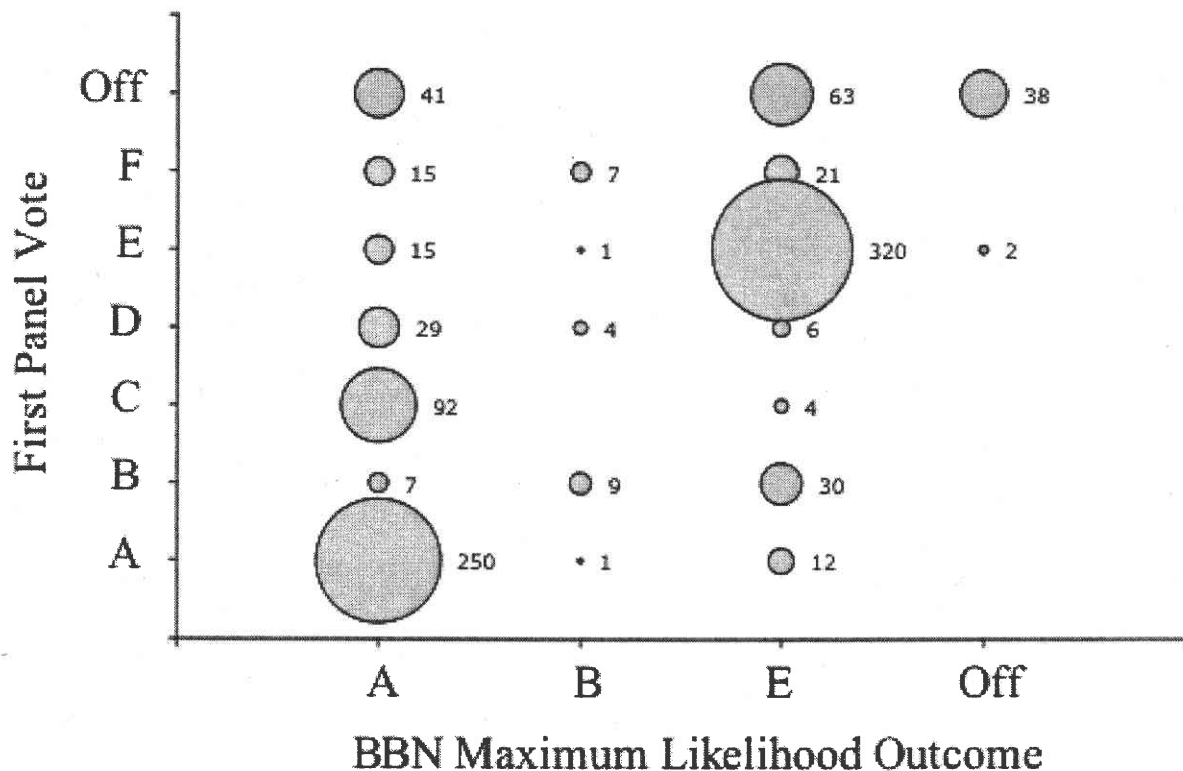
Taking into account the data on this species and the outcomes of the BBN models in their Stage 1 vote, all but one of the panelists voted fully for outcome E, and one voted for outcome A. After Delphi roundtable disclosures, questions, and discussions in panel Stages 2 and 3, the Stage 4 final vote resulted in unanimous agreement on outcome E for this species. In this case, the uncertainty over, and thus possibility of LSOG forest association was, in the panelists’ opinions, insufficient to warrant voting for taking this apparently very rare species off the

list. The conservation category E in turn meant that only strategic surveys would be conducted and that all known sites would be protected (Table 1).

Overall use of the Bayesian network model decision models

The 2002 and 2003 ASR panels used the BBN decision-aiding models to evaluate 119 species. Each year, separate flora and fauna evaluation panels each convened 4 to 5 d, and each evaluated on average 30 species. The BBN models contributed tangibly to the ASR species evaluations. Of the 119 species reviewed, about half were suggested by the ASR evaluation panels to change conservation management.

Fig. 9. Distribution of the first round of votes across all species from the 2002 and 2003 evaluation panels of the Annual Species Reviews under the Northwest Forest Plan, by species conservation management category (A-F and Off the list; see Table 1 for category descriptions), for each maximum likelihood outcome of the overall Bayesian belief network (BBN) model (Fig. 1). The area of each bubble corresponds to the number of votes, also shown next to each bubble.

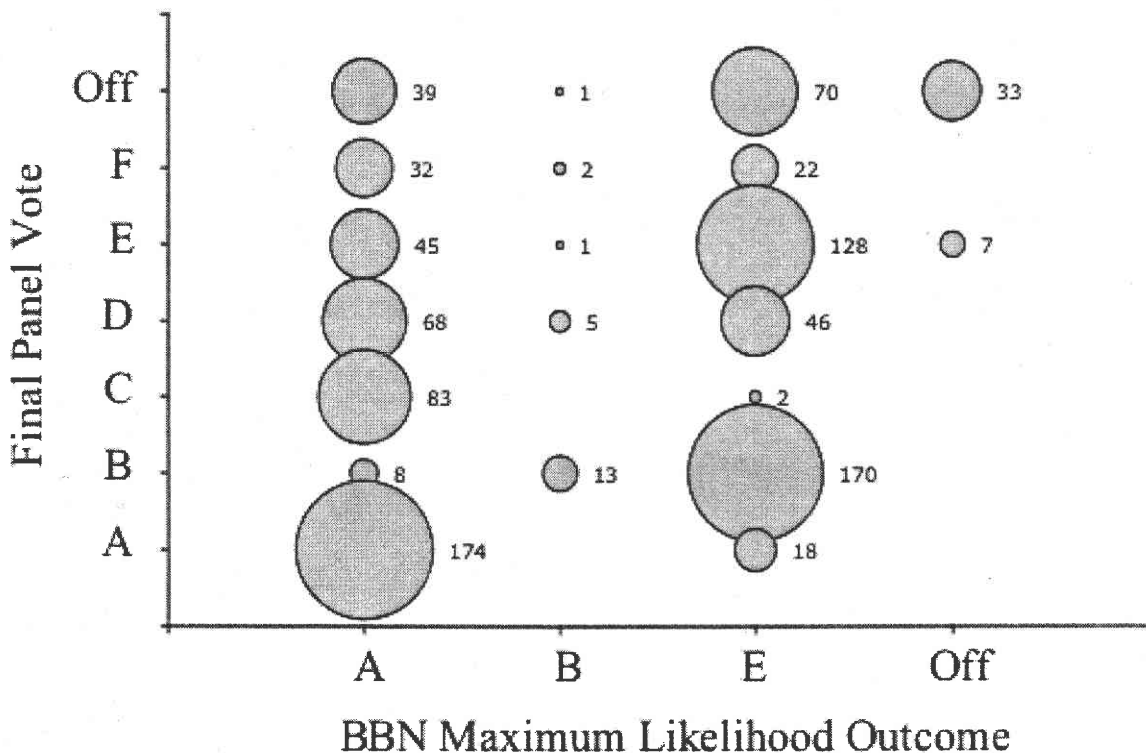


One of the advantages of the BBN modeling approach was being able to determine alternative outcomes of species conservation management categories when data were lacking or uncertain on the input parameters. For example, because of the recent taxonomic split of the lichen *F. saubinetii*, data are lacking for this species on LSOG forest association. Thus, in the BBN submodel pertaining to LSOG association (Fig. 3), the input parameters on abundance in LSOG forest (node 1) and association with LSOG forest (node 2) and LSOG components (node 3) are set to uniform probabilities, denoting maximum uncertainty. This uncertainty then propagates in the model as

alternative outcomes (Fig. 3) and ultimately as alternative conservation management categories for the species (Fig. 1).

In this way, the BBN decision models clearly showed the availability and uncertainty of the scientific data for each input parameter, and the influence on conservation management categories, in an unbiased and consistent manner. For some species and some evaluation categories, very little data were available. An example again with *F. saubinetii* is shown with determining its relative rarity or uncommonness (Figs. 7 and 8), for which only a few of the input parameters were able to be

Fig. 10. Distribution of the final votes across all species from the 2002 and 2003 evaluation panels of the Annual Species Reviews under the Northwest Forest Plan, by species conservation management category (A-F and Off the list; see Table 1 for category descriptions), for each maximum likelihood outcome of the overall Bayesian belief network (BBN) model (Fig. 1). The area of each bubble corresponds to the number of votes, also shown next to each bubble. The BBN decision model did not result in species conservation management categories C, D, or F being the most likely outcome for any species, although they were sometimes shown as being less likely outcomes.



specified. For only one-third of the 119 species assessed did the BBN models provide a single conservation management category outcome; for the rest of the species, nonzero probabilities for two or more alternative categories were propagated from uncertainty in the input parameters.

Use of the models to guide the paneling process

Overall, the BBN decision models helped guide the ASR evaluation panelists' discussions and voting process by helping to focus discussion and deliberation on areas of greater uncertainty and on those guideline parameters that seemed to most affect conservation management categories. Because of uncertainty in the input parameters and

vagueness in the evaluation guidelines, it was not expected that each panelist would vote strictly according to the model outcomes. In fact, 60% of the time, panelists deviated in their initial votes from the maximum likelihood outcome of species conservation management categories denoted in the overall BBN model (Fig. 1) for each species, although their deviation from any of the model's nonzero outcomes was infrequent.

For example, when the overall BBN model suggested conservation management category A as the most likely outcome for a species, panelists put most of their 100 points of their initial vote into outcome A only 56% of the time (Fig. 9); they were more likely to initially vote according to model results for off, 95% of the time, and outcome E (70%). In their final votes, the panelists concurred with the maximum likelihood category shown in the overall BBN model only 36% of the time, ranging from 28% concurrence with model outcome E to 83% with outcome off (Fig. 10).

Again, the deviation between model outcomes and panelist votes was expected and did not invalidate the structure or use of the models. In fact, only 51% of the time did the panelists' final vote match their own, most likely initial vote. This was because the paneling process of modeling, initial voting, vote disclosure, questioning, discussing, and final voting served to clarify points of uncertainty, reduce error and bias, and allowed the panelists to change their votes accordingly.

Results also varied by taxonomic group. The panelists tended to cast their final votes most often for more conservative management categories than the BBN models suggested for fungi; for equally conservative categories for lichens, vascular plants, and mollusks; and for less conservative categories for the other taxa. These differences typically were because of number of detections and known sites of species, e.g., the relative rarity of fungi, which the panelists weighed more heavily than did the models as evidenced by the panelists' written explanations.

An example of comparing the BBN model outcomes with panelist votes for the lichen *F. saubinetii* is shown in Table 4. The BBN model (Fig. 1) suggested three potential outcomes with probabilities ranging 30–37%. The evaluation panel's initial votes were mostly for category E (84% of all vote points) with some votes for A (16%); none of the panelists initially voted for the "off" category because, despite uncertainty over LSOG association

for this species, the panelists felt that the species' apparent rarity and other attributes led to retaining this species on the SM species list. After discussions and questions to taxa experts, which focused on clarifying those parameters that would lead to outcomes A or "off," their final votes were unanimous for category E. Thus, in this example, the panelists' initial votes only partially overlapped, and included a slightly wider set of outcome categories, than from the BBN model, and their final votes were more narrow and were fully a subset of the model outcome. This pattern varied considerably among species and taxa for reasons noted above.

DISCUSSION

Utility of the Bayesian belief network decision models in the annual species review panels

The Bayesian belief network (BBN) decision models explicitly laid out the parameters and the species evaluation process as a risk analysis, that is, showing alternative management conservation categories that would be consistent with the scientific data on each species, and the degree, i.e., probability, to which each category might pertain in each case. Showing the outcomes as probabilities fit well within a risk management framework, which was requested by the decision-making body to which the evaluation panels presented their results.

We found that the BBN decision models helped the most in prompting the members of the evaluation panels to consistently consider all guideline parameters, including those for which data were uncertain or unavailable. It was human nature to otherwise focus on parameters which were most clearly explained in the guidelines, for which the scientific data were most complete, for which a panelist had the most experience, or for which a panelist put the most weight of judgment as happened with the raw counts of number of known locations of a species as a proxy to rarity. The models helped avoid such motivational bias by always displaying all parameters, whether known and clearly explained, or not, and their influence on outcomes of conservation management categories for each species.

As a discussion aid, the panelists also ran the BBN models during panel deliberations to determine the influence of parameters for which conservation

Table 4. An example of an evaluation of a species of rare lichen, *Fuscopanaria saubinetii*, from the 2003 annual species review, showing outcomes of the Bayesian belief network (BBN) decision models and the voting results from the evaluation panel. Panelist 7 was absent during this particular panel. The prior (2002) conservation management category for this species was category F (see Table 1).

Model or Panel	Species conservation management category						Off
	A	B	C	D	E	F	
BBN species model (probability)							
	0	0	0	0	33	30	37
Panel – 1st votes (percent)							
Panelist 1	0	0	0	0	100	0	0
Panelist 2	100	0	0	0	0	0	0
Panelist 3	0	0	0	0	100	0	0
Panelist 4	0	0	0	0	100	0	0
Panelist 5	10	0	0	0	90	0	0
Panelist 6	0	0	0	0	100	0	0
Panelist 7							
Panelist 8	0	0	0	0	100	0	0
TOTAL voting points (%)	110 (16%)	0 (0%)	0 (0%)	0 (0%)	590 (84%)	0 (0%)	0 (0%)
Panel – 2nd votes (single vote)							
Panelist 1					X		
Panelist 2					X		
Panelist 3					X		
Panelist 4					X		
Panelist 5					X		
Panelist 6					X		
Panelist 7							
Panelist 8					X		
Total number of votes	0	0	0	0	7	0	0

outcomes were most sensitive for a given species. Running the models entailed posing "what if" type questions and altering the model inputs for a given species, such as, what would be the conservation category outcomes if a given species had a narrow ecological amplitude, or a stable population, or a more limited dispersal capability? In a sense, this was an informal sensitivity analysis that helped focus the topics of discussion and questions posed by the panelists to the taxa experts.

A more formal sensitivity analysis was run on species for which panelists thought it useful, to determine which input parameters most influenced the conservation category outcomes. This was useful when the BBN decision model projected multiple outcomes with similar probabilities, or when the panelists diverged widely in their vote categories. For instance, with the example lichen species presented here, the species conservation categories in the overall model (Fig. 1) were most sensitive to evaluation categories 2, i.e., late-successional and old-growth (LSOG) association, and 6b, the "uncommon" part of relative rarity, so the panelists would desire greater clarification on scientific knowledge and certainty of the factors influencing those parameters.

Building and using Bayesian belief network decision models

These were simple BBN decision models despite their size and apparent complexity because they represented the evaluation guidelines in a deterministic way. Because the models were based on simple, deterministic linkages, such tools should be used at most to rank-order possible alternative outcomes, not to suggest specific probabilities.

Such models also should be used most appropriately to interpret implications of uncertainty. It is vital to understand how such models represent expected values of outcomes based on the Boolean linkage of the evaluation guideline criteria. For instance, if 5 criteria are to be used to evaluate population rarity and if any one criterion can trigger a "rare" condition, then the "expected value" in the face of total uncertainty of the input criteria will weigh the outcome strongly toward rare. In fact, this was the case with the guidelines and how they were represented in the models.

Our crafting the models also helped identify imperfections in the evaluation guidelines, particularly that some were vague or ambiguous, and all were qualitative in their descriptions and thus subject to multiple interpretations. Their qualitative nature provided latitude and flexibility but also some unavoidable inconsistency in their application.

CONCLUSIONS AND CONSIDERATIONS FOR FUTURE APPLICATION

Our experience suggests that it is vital to craft clear, unambiguous, and internally consistent evaluation guidelines to help reduce uncertainty in recommended management actions. Using structured decision-aiding tools, e.g., Bayesian belief network (BBN) decision models, and assessment procedures, e.g., the Delphi paneling process, can greatly help to reduce inconsistencies and bias in the interpretation of evaluation guidelines, but such tools and procedures alone cannot eliminate semantic uncertainty inherent in a set of guidelines.

Also, it may be useful to test a draft set of evaluation guidelines by formalizing them in a decision-aiding tool such as a BBN model, even if such tools are not explicitly used in the subsequent application of the guidelines. This could help identify incompleteness, ambiguities, and inconsistencies within the guidelines. It may even be desirable to hold mock panel evaluations, with panels of managers or resource experts using the decision-aiding models, to test the interpretation and identify weaknesses in the guidelines.

Pros and cons of the Bayesian belief network decision models

Like any such tools, the BBN decision models had both strong and weak points. Among the strong points, using BBN decision models:

- formalized and structured the evaluation guideline criteria, and made them explicit and repeatable;
- organized thinking;
- allowed organizing complex combinations of evaluation criteria into summary nodes;

- prompted new considerations;
- provided quick sensitivity testing of the influence and value of new knowledge, change in interpretations, and effects of uncertainty;
- identified where there was limited or insufficient information and uncertainty;
- provided a parity across all species and panelists, greatly reducing bias in thinking, i. e., considering only favorite or better-known ecological factors;
- provided a rigorous basis for expert paneling whereby the panelists had to explicitly explain why their votes deviated from the model outcomes; and
- provided a means of efficiently representing and storing some of the administrative records, explaining how the models were parameterized from the scientific data for each species.

However, some cons of using BBN decision models, at least with this set of species management guidelines, were that:

- these particular models did not incorporate any estimates of utility values, e.g., social costs, biological benefits of ensuring species viability, etc.;
- the models were perhaps too rigidly structured for use with fuzzy or ambiguous decision criteria, evidenced by how often the panelists' final votes deviated from model outcomes;
- the models were at times confusing in how they mixed management criteria, i.e., whether the Northwest Forest Plan (NWFP) provides for species persistence, whether data are sufficient for management decisions, and the practicality of conducting species surveys; see evaluation categories 3, 4, and 5 (Table 2) and ecological criteria such as geographic range and late-successional and old-growth (LSOG) association (see evaluation categories 1, 2, 6a, and 6b in Table 2), although such was the nature of the published guidelines for

this exercise;

- the models at times constrained thinking on additional ecological factors, which could greatly affect outcomes, that were not specified in the evaluation guidelines, e.g., effects of climate change, species competition, the legacy of past management actions, catastrophic disturbance events, etc.; and
- at times the process of using the models led to fully equivocal outcomes because of ambiguities in the evaluation guidelines, such as a species being both uncommon and rare, with great difficulty to identify any one factor most responsible for the ambiguities.

The rigidity of the BBN models resulted directly from that of the published guidelines. It was clear, however, that the discrepancies between model results and panelist judgments and voting patterns (Figs. 9 and 10) meant that the specific guideline criteria (Table 2) per se did not represent the degree of flexibility in interpretation that was also afforded to the panelists when evaluating those criteria. That is, the written guidelines did not explicitly incorporate such flexibility of interpretation. This outcome, though, does not fully exclude the possibility that the panelists' judgments and voting categories were faulty, although the rigor of the Delphi paneling process was established to minimize such possible faults.

Mixing management criteria with ecological criteria meant that the resulting BBN decision model was not to be interpreted as an ecological prediction model, but rather as a representation of the evaluation guidelines. This was not a problem as long as the model was used in its proper decision-aiding context and not to predict species viability.

Considerations for use of decision-aiding models for evaluating multiple stressors

We suggest, overall, that such BBN decision models served well to organize thinking, structured the use of the evaluation guidelines, and provided an intuitive means of exploring implications of data and uncertainty in an effective adaptive management process. The models generally worked well in the species evaluation process.

We also suggest that such models be developed, as we did, using simple deterministic relations as far as possible unless a set of evaluation guidelines clearly specify probabilistic or other quantitative combinations of criteria. This does not mean that input parameters must be solely qualitative or categorical in nature, as was the case with the guidelines and models in this effort. In fact, quantifying input parameters with clearly measurable units and ranges of values for their states often would have been preferred, to help reduce semantic uncertainty and vagueness.

Use of decision-aiding models does not obviate the need to fully explain the evaluation or decision guidelines in simple, written terms. Such guidelines often contain more subtle distinctions, inconsistencies, and contexts that cannot be fully captured in decision-aiding models.

Responses to this article can be read online at:
<http://www.ecologyandsociety.org/vol11/iss2/art12/responses/>

Acknowledgments:

Our thanks to two anonymous reviewers and the editor for helpful comments on the manuscript. We thank Luka Jordan and Charrise Chesbrough for their help in data entry of the panel votes. We also thank all 15 managers and specialists involved in the evaluation panels, and the many species experts involved in compiling the basic ecological data used in the panels. We developed the example decision models shown on *Fuscopannaria saubinetii* from ecological data provided by lichenologists Marty Stein and Chiska Derr. Thanks also to Terry Brumley for administrative support. This work was conducted under the Survey and Manage Species Program, Northwest Forest Plan, of USDA Forest Service and USDI Bureau of Land Management.

LITERATURE CITED

Ayyub, B. M. 2001. *Elicitation of expert opinions for uncertainty and risks*. CRC Press, Boca Raton, Florida, USA.

Cleaves, D. A. 1995. Assessing and communicating uncertainty in decision support systems: lessons

from an ecosystem policy analysis. *AI Applications* 9(3):87-102.

Edwards, T. C., D. R. Cutler, L. Geiser, J. Alegria, and D. McKenzie. 2004. Assessing rarity of species with low detectability: lichens in Pacific Northwest forests. *Ecological Applications* 14(2):414-424.

Iliadis, L. S. 2005. A decision support system applying an integrated fuzzy model for long-term forest fire risk estimation. *Environmental Modelling and Software* 20(5):613-621.

Leshner, R. D. 2005. An environmental gradient model predicts the spatial distribution of potential habitat for *Hypogymnia duplicata* in the Cascade Mountains of northwestern Washington. Dissertation. University of Washington, Washington, USA.

Marcot, B. G. 2006. Characterizing species at risk I: modeling rare species under the Northwest Forest Plan. *Ecology and Society* 11(2):10. Online URL: www.ecologyandsociety.org/vol11/iss2/art10/.

Marcot, B. G., J. D. Steventon, G. D. Sutherland, and R. K. McCann. Guidelines for developing and updating Bayesian belief networks for ecological modeling. *Canadian Journal of Forest Research*, in press.

McCune, B., and L. Geiser. 1997. Macrolichens of the Pacific Northwest. Oregon State University Press, Corvallis, Oregon, USA.

Molina, R., B. G. Marcot, and R. Leshner. 2006. Protecting rare, old-growth forest associated species under the survey and manage guidelines of the Northwest Forest Plan. *Conservation Biology* 20(2):306-318.

Molina, R., D. McKenzie, R. Leshner, J. Ford, J. Alegria, and R. Cutler. 2003. *Strategic survey framework for the Northwest Forest Plan Survey and Manage Program*. USDA Forest Service General Technical Report PNW-GTR-573. Pacific Northwest Research Station, Portland Oregon, USA.

Pieterse, N. M., A. W. Verkroost, M. J. Wassen, V. H. Olde Venterink, and C. Kwakernaak. 2002. A decision support system for restoration planning of stream valley ecosystems. *Landscape Ecology* 17 (1 Supp.):69-81.

Reynolds, K. M., and P. F. Hessburg. 2005.

Decision support for integrated landscape evaluation and restoration planning. *Forest Ecology and Management* 207(1-2):263-278.

Rojo, J. M., and S. S. Orois. 2005. A decision support system for optimizing the conversion of rotation forest stands to continuous cover forest stands. *Forest Ecology and Management* 207 (1-2):109-120.

U.S. Department of Agriculture Forest Service, and U.S. Department of the Interior Bureau of Land Management (USDA and USDI). 1994. *Record of decision for amendments to Forest Service and Bureau of Land Management planning documents within the range of the northern spotted owl and the standards and guidelines for management of habitat for late-successional and old-growth related species within the range of the northern spotted owl.* USDA Forest Service and USDI Bureau of Land Management, Portland, Oregon, USA.

U.S. Department of Agriculture Forest Service, and U.S. Department of the Interior Bureau of Land Management (USDA and USDI). 2001. *Record of decision and standards and guidelines for amendments to the survey and manage, protection buffer, and other mitigation measures standards and guidelines in Forest Service and Bureau of Land Management planning documents within the range of the northern spotted owl.* USDA Forest Service and USDI Bureau of Land Management, Portland, Oregon, USA.

Zhu, X., and A. P. Dale. 2000. Identifying opportunities for decision support systems in support of regional resource use planning: an approach through soft systems methodology. *Environmental Management* 26:371-384.