

Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation¹

Bruce G. Marcot, J. Douglas Steventon, Glenn D. Sutherland, and Robert K. McCann

Abstract: Bayesian belief networks (BBNs) are useful tools for modeling ecological predictions and aiding resource-management decision-making. We provide practical guidelines for developing, testing, and revising BBNs. Primary steps in this process include creating influence diagrams of the hypothesized “causal web” of key factors affecting a species or ecological outcome of interest; developing a first, alpha-level BBN model from the influence diagram; revising the model after expert review; testing and calibrating the model with case files to create a beta-level model; and updating the model structure and conditional probabilities with new validation data, creating the final-application gamma-level model. We illustrate and discuss these steps with an empirically based BBN model of factors influencing probability of capture of northern flying squirrels (*Glaucomys sabrinus* (Shaw)). Testing and updating BBNs, especially with peer review and calibration, are essential to ensure their credibility and reduce bias. Our guidelines provide modelers with insights that allow them to avoid potentially spurious or unreliable models.

Résumé : Les réseaux de croyance bayésiens (RCB) sont des outils utiles pour faire des prédictions de nature écologique à l'aide de modèles et pour aider dans la prise de décision en aménagement des ressources. Nous fournissons des directives pratiques pour développer, tester et réviser des RCB. Les principales étapes dans ce processus incluent : la création d'un diagramme d'influences du « réseau causal » hypothétique des facteurs-clefs qui affectent une espèce ou un enjeu écologique d'intérêt; le développement d'un premier modèle RCB de niveau alpha à partir du diagramme d'influences; la révision du modèle par des experts; l'essai et la calibration du modèle à partir d'études de cas; la création d'un modèle de niveau bêta; l'actualisation de la structure et des probabilités conditionnelles du modèle à partir de données de validation et la création d'un modèle final d'application, de niveau gamma. Nous illustrons et discutons ces étapes à l'aide d'un modèle RCB empirique des facteurs qui influencent la capture des grands polatouches (*Glaucomys sabrinus* (Shaw)). L'essai et l'actualisation des RCB, surtout à l'aide de révisions par les pairs et de la calibration, sont essentiels pour assurer leur crédibilité et réduire les biais. Nos directives donnent aux modélisateurs des indications pour éviter les modèles potentiellement inexacts ou peu fiables.

[Traduit par la Rédaction]

Introduction

Quantifying the relationships between ecological variables (e.g., species' occurrence, attributes of organisms, communities, or ecosystems, and their functional behavior) and sam-

ple measurements obtained from a set of possibly related observations remains a fundamental problem of inference in ecology. Bayesian belief networks (BBNs; also called probability networks) are statistical tools used in ecology and wildlife management to depict the influence of habitat or environmental predictor variables on ecological-response variables. Examples include BBN models of the response by wolverine (*Gulo gulo* (Linnaeus, 1758); Rowland et al. 2003) and greater sage grouse (*Centrocercus urophasianus* (Bonaparte, 1827); Wisdom et al. 2002a, 2002b) to habitat patterns across landscapes, the viability of salmonid populations (Lee and Rieman 1997) in the interior northwest USA, and prediction of water quality in the southeast USA (Reckhow 1999). In such uses, BBNs predict the probability of ecological responses to varying input assumptions such as habitat and population-demography conditions. Ideally, BBNs serve well as part of a risk-management framework by explicitly displaying the “causal web” of interacting factors and the probabilities of multiple states of predictor and response variables.

Development of the BBNs cited above, however, apparently did not follow a standard process. Some authors devel-

Received 18 October 2005. Accepted 26 May 2006.

Published on the NRC Research Press Web site at
<http://cjfr.nrc.ca> on 14 February 2007.

B.G. Marcot.² USDA Forest Service, 620 S.W. Main Street, Suite 400, Portland, OR 97205, USA.

J.D. Steventon. British Columbia Ministry of Forests and Range, Bag 6000, 3333 Tatlow Road, Smithers, BC V0J 2N0, Canada.

G.D. Sutherland. Cortex Consultants Inc., Suite 2a, 1218 Langley Street, Victoria, BC V8W 1W2, Canada.

R.K. McCann. PO Box 421, Mackenzie, BC V0J 2CO, Canada.

¹This article is one of a selection of papers published in the Special Forum on Bayesian Belief Networks in Natural Resource Management.

²Corresponding author (e-mail: bmarcot@fs.fed.us).

oped BBNs based strictly on expert judgment from individuals or expert panels. Few authors attempted to calibrate and validate their models. None apparently updated their models on the basis of validation outcomes and empirical case data. What is needed is a set of guidelines and insights to help modelers avoid potentially spurious or unreliable models. Methods suggested in this paper provide a standard procedure to accomplish this objective.

In this paper we will presume that the reader has a basic knowledge of BBN structures and how to construct a rudimentary BBN model. Introductory references are available on the statistical structure of BBNs and related Bayesian modeling schema (e.g., Cain et al. 1999; Charniak 1991; Jensen 1996; Zhang and Poole 1996). However, practical guidelines for correctly and efficiently developing BBNs are scattered widely in the literature (e.g., Castillo et al. 1998; Marcot et al. 2001; Reckhow 1999; Varis 1997), and useful sources in the gray literature are not generally available (e.g., Cain 2001; Heckerman et al. 1994). The objective of this paper is to provide a set of practical guidelines for developing, testing, and revising BBNs for wildlife and ecological assessment.

We draw from our extensive experience in BBN-model construction (for examples see other papers in this issue), where we developed and successfully applied these guidelines. We use the BBN-modeling shell Netica® (versions 2.17 or later, Norsys Systems Corp., Vancouver, British Columbia), although other BBN-modeling tools are also available and our suggestions may pertain to those tools as well.

Here we (i) describe the steps for creating, testing, calibrating, and updating BBN models at three levels (alpha, beta, gamma); (ii) discuss BBN models in relation to other statistical models; and (iii) provide recommendations on appropriate construction and interpretation of BBN models for ecological prediction and aid in decision-making. Our guidelines include ways to correctly and efficiently craft a BBN model and appropriate uses of the model. We use as an example a BBN model predicting the probability of capturing northern flying squirrels (*Glaucomys sabrinus* (Shaw)).

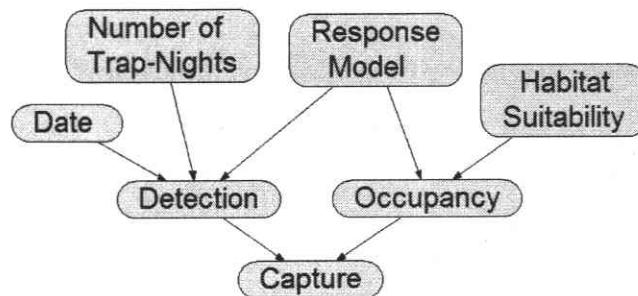
Creating the alpha-level model

Converting the influence diagram to an initial BBN

The first step is to illustrate the "ecological causal web" of the key environmental correlates or other influences affecting the species or outcome of interest. This can be done using influence diagrams (Oliver and Smith 1990; Reckhow 1999; Shachter and Kenley 1989), which, at their simplest, are figures consisting of boxes and arrows showing relevance and influence among variables (Fig. 1). Depending on the intended use of the model, several influence diagrams may be constructed at various spatial scales.

This first step is often done in discussion with species or subject-matter experts and after an initial review of the literature. For example, after identifying the species of interest, experts and literature can be consulted to list all key environmental correlates. These can then be arranged into an influence diagram showing expected causal influences on the species of interest. In our example here (Fig. 1), the outcome of interest is prediction of the probability of capturing a northern flying squirrel at a particular trap site. Based on

Fig. 1. Example of a simple influence diagram showing key factors affecting expected capture of northern flying squirrels (*Glaucomys sabrinus* (Shaw)). Capture is shown to be dependent on detectability, which is influenced by date and trapping effort, and by occupancy, which may or may not be influenced by suitability of local habitat (forest) conditions (response model). Since such diagrams are best used as a basis for further modeling, the correlates shown should also express causal relationships, so that the diagram becomes a "causal web".



field experience, we believe that capture probability is influenced both by detectability of the organism, which is a function of time of year (date) and trapping intensity (number of trap-nights), and by occupancy of the trapping area, which may be a function of the quality of the surrounding forest habitat (habitat suitability) and whether or not habitat suitability is even pertinent (response model).

The influence diagram can then be developed into an initial BBN where each node (box) is converted to a set of discrete states (such as squirrel capture probability classes in our example). "Parent" nodes feed into "child" nodes. Following through the network, a child node may then become a parent to other nodes. Conditional probabilities for the states of each child node are then specified for all combinations of states of their parent nodes. We distinguish between probabilities, which are frequencies of outcome states given a set of causal conditions, and likelihoods, which are potential causes given an outcome state.

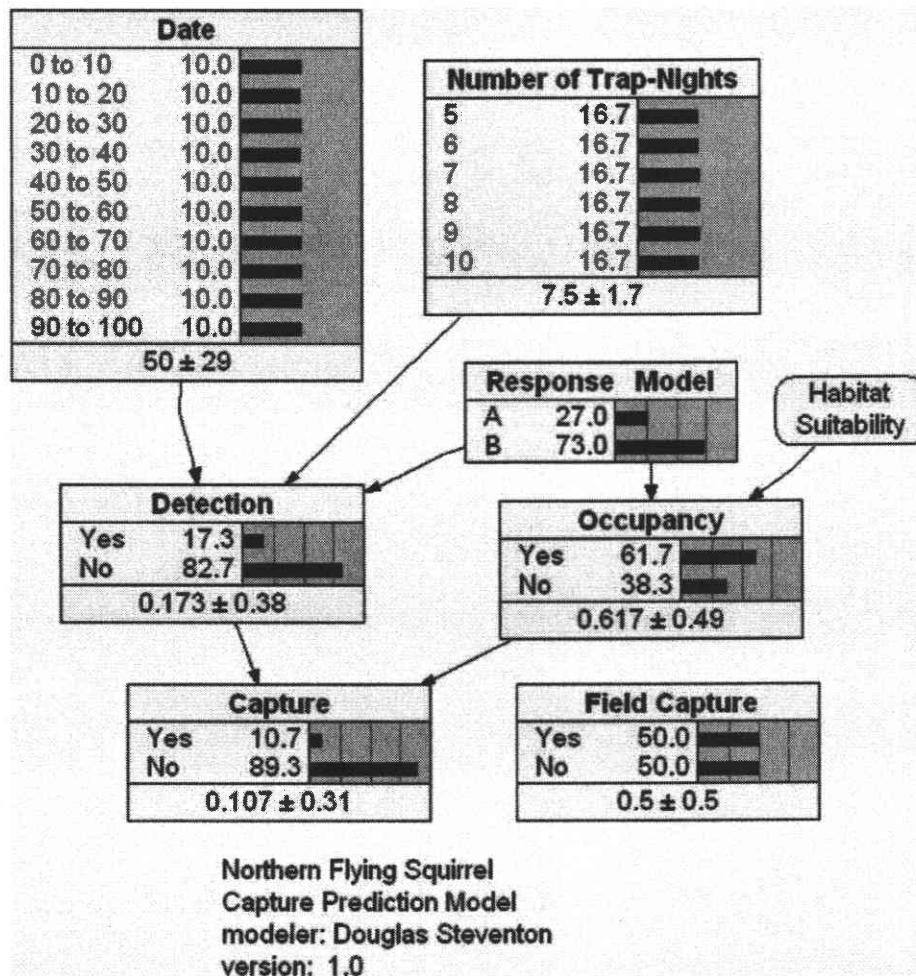
Converting influence diagrams to BBNs can involve sophisticated learning algorithms (Shachter 1988; Smith et al. 1993; Zhang 1998), but we defer presenting these to a later step when we discuss updating BBNs from field data. The current step can also be done simply from consultation with a species or domain expert to specify the states and equations in each BBN node (Fig. 2; Appendix A).

We have found the following guidelines useful for developing the initial structure of the BBN:

- (1) As far as possible, the number of parent nodes to any given node should be kept to three or fewer (e.g., no node in Fig. 2 has greater than three parent nodes) and the number of their states to five or fewer. This keeps the associated conditional-probability table (CPT; see below) small enough to be tractable and understandable if it is to be specified by the modeler and experts rather than by equations or case data. The size of the CPT of the child node is equal to the number of states of the child node times the product of the number of states of

all parent nodes, or with n parent nodes, $S \prod_{i=1}^n P_i$ where S is the number of states of the child node and P_i is the

Fig. 2. Example of an initial parameterized Bayesian belief network model based on the simple influence diagram shown in Fig. 1. Nodes are defined as follows: “capture” is the overall predicted probability of capturing a squirrel at least once at a site; “detection” is the probability of detecting a squirrel if it is present at a site; “occupancy” is the probability that a site is occupied by a squirrel; “response model” denotes the hypotheses that detectability is dependent on trapping date and that occupancy is either constant or varies according to habitat suitability; “habitat suitability” denotes suitability of a habitat for the squirrel in a 16 ha area around each trap site; “date” is the number of days from the start of the trapping session on 22 June; “number of trap-nights” is the number of nights when traps were open at each site; “field capture” denotes field capture data (yes/no) used in assessing model accuracy. The model is shown here running the prior probabilities of input conditions derived from field studies and model parameterization (see Appendix A). The additional unlinked field capture node represents the actual observation at each field site (capture or noncapture). This is used when processing the field-data case file to associate the field result with the predicted capture probability for use in assessing model prediction accuracy.



number of states of the i th parent node. An exception to this guideline occurs when the objective is to represent a complex ecological process where representing that complexity may be more important than predictive accuracy or model parsimony.

- (2) Parentless (input) nodes — typically representing predictor habitat and environmental factors (e.g., “date” and “habitat suitability” nodes in Fig. 2) — should be those items that can be preprocessed or empirically evaluated from existing data, such as geographic information systems (GIS).
- (3) Intermediate nodes (e.g., “detection” and “occupancy” nodes in Fig. 2) should be used to summarize the major themes (the “latent variables”) denoted in the generalized species influence diagrams.

- (4) To the extent possible, all nodes should be observable and quantifiable or testable entities. In some cases, intermediate nodes may not be so, but should still be carefully documented and explained.
- (5) The fewest discrete states necessary within any given node should be used to represent influences, but enough states should be used to ensure the desired precision of estimates and the range of input values in the model (e.g., the date node in Fig. 2 has 10 states to be used in calculating detection probability, but only two states are used to depict two alternative habitat models in the “response model” node). This is a balance between parsimony and precision. Precision is also determined by the overall size of the CPT (see guideline 1). We caution that increased precision is not necessarily equivalent to

increased accuracy. In some models, nodes with fewer states may provide a more accurate prediction, although more states may provide greater precision.

An exception to this guideline occurs when a calculation needs continuous values (e.g., the formula calculating the occupancy node in Fig. 2). In this case, the nodes need to be "discretized" to convert the continuous outcome to discrete states, and this can be done using a large number of discrete states if needed (e.g., as used in the models by Steventon et al. 2006). If sources provide specific cutoff values of some continuous environmental or habitat variable, one should use those values to define the cutoff values of the discrete states.

A corollary to this guideline is to insert states where conditions intermediate between the states shown need to be explicitly represented. That is, spreading probabilities between two states is a measure of uncertainty or of a known or suspected frequency distribution of those states, not a statement of some intermediate condition between those states.

- (6) The depth of the model — the number of layers of nodes — should be kept to four or fewer, if possible. This is desirable for at least three reasons: deep models with many intermediate nodes (latent variables) may contain unnecessary uncertainty propagated from input to output nodes; the sensitivity of the output node to input nodes may be swamped and dampened by intermediate nodes; and output nodes in models with asymmetric structures may be far less sensitive to more distant input nodes with many intervening intermediate nodes than the modeler intended. If it is not possible to keep the model structure relatively shallow, consideration should be given to breaking up the model into two or more networks.

As with guideline 1, an exception occurs when the objective is to represent complex ecological processes. However, algorithms can be applied (in Netica®, for example) as needed, to simplify overly complex models by "absorbing" nodes that are no longer being updated or have minimal influence while maintaining the inferential relationship among remaining nodes.

- (7) If models are needed for several spatial scales, they should be developed simultaneously. Output of one BBN model can be used as input to another. For example, a broad-scale landscape model identifying general biophysical conditions suitable for a species can become the basis for a fine-scale site model (or vice versa).
- (8) The model, including the rationale for each node and each linkage, should be fully documented. Some modeling shells, such as Netica®, have integrated documentation features that help track authorship, dates and sources of updates, and descriptions of each node.
- (9) Links should be inserted between the input nodes if they are likely to be correlated; a statistical assumption of BBNs and the statistics of Bayesian updating is that prior probabilities associated with unlinked input nodes are uncorrelated.

For example, tree age and height are usually correlated (Kalliovirta and Tokola 2005). A link can be shown from age to height (e.g., height as a partial function of age) if they are used as inputs, and then links can be drawn from age and height to other parts of the model

Table 1. Example of a case data table showing northern flying squirrel (*Glaucomys sabrinus* (Shaw)) captures from multiple trapping nights.

Date	No. of trap-nights	Habitat suitability	Capture
0	5	0.151875	0
0	5	0.634375	0
0	5	0.08125	0
0	5	0.74875	0
0	5	0.553125	0
:	:	:	:
95	7	0.253125	0
95	7	0.421875	0
95	7	0.252500	0
95	7	0.265625	0
95	7	0.150625	1

Note: Each row represents a different trapping-sample outcome. Such data can be used to induce an initial network structure or to refine probabilities of relationships in an existing network. Each column heading refers to a node in the squirrel Bayesian belief network (BBN) model (Fig. 2). Only 10 of the total 135 trapping-sample outcomes, which included 112 no-capture outcomes (capture = 0) and 23 capture outcomes (capture = 1), are shown.

as needed. When processing a case (see below) where values for both age and height are known, then the actual values of both are applied. However, if age is known but height is not, then height is inferred to be conditional on age rather than assumed to be completely unknown. In this way, correlations between the nodes and prior knowledge are used.

Model-induction approaches

One approach to building the alpha-level BBN model is to use case data to induce the model structure, that is, the nodes, their states, and their linkages. In general, this is a form of "data mining" (Hastie et al. 2001) and "rule induction" (Jeffers 1991) whereby specific relationships among variables are calculated from data tables. As in the example given here (Table 1), multiple trapping results for a species may result in a case data table showing conditions (date, number of trap-nights, and habitat suitability) at trap locations. The case data table can then be mined to determine the highest correlations of the prediction variables (trapping conditions and environmental or habitat factors) with the response variable (species capture). However, we do not advocate this approach for initially structuring the BBN model, as it might overfit the data, which are often scant and biased, particularly with rare species or uncommon ecological events. Overfitting data often results in model structures that represent specific and spurious statistical correlations rather than more robust, expert-based knowledge on causation (Clark 2003). In the following section we discuss using case data for updating node probabilities, which is a more appropriate use.

Representing proxy nodes

The ideal BBN model uses input (parentless) nodes that represent true, proximate environmental or habitat attributes with causal influence. In many cases, however, data simply

Table 2. Example of a conditional-probability table for the detection node in Fig. 2.

Date	Parent node states		Outcome states (detection)	
	No. of trap-nights	Response model	Yes	No
0 to 10	5	A	0.0566	99.943
0 to 10	5	B	0.0594	99.941
0 to 10	6	A	0.0680	99.932
0 to 10	6	B	0.0712	99.929
0 to 10	7	A	0.0792	99.921
:	:	:	:	:
90 to 100	8	B	79.229	20.771
90 to 100	9	A	81.618	18.382
90 to 100	9	B	82.760	17.240
90 to 100	10	A	84.621	15.379
90 to 100	10	B	85.699	14.301

Note: Values were calculated according to the method and equation shown in Appendix A. Outcome values sum to 100 across each row and represent frequencies or probabilities of each outcome state, given combinations of input (parent node) states. Not all table entries are shown here.

are not available for such variables, and proxy variables can be used to represent the ideal variable. As an example, data on forest-stand canopy closure may not be available but data on stand tree density could serve to some degree as a proxy. In such cases, for those nodes one can add a parent node that would be the proxy variable to the correlate node it represents. The CPT for the correlate node would then be adjusted to represent the degree to which each state of the proxy node truly represents each state of the correlate node. This essentially denotes the uncertainty with which the proxy node represents the correlate node. Such probabilities can be estimated after consulting with domain or GIS experts or generated from case data.

Creating the CPTs for each node

The heart of any BBN consists of the set of probability tables underlying each node (e.g., Table 2). Parentless nodes have unconditional probability tables that represent prior knowledge on frequencies of each state or, alternatively, uniform probabilities if there is complete uncertainty about those prior conditions. It is generally preferable to at least estimate prior distributions, representing some prior knowledge, rather than use uniform distributions, representing complete uncertainty.

Child nodes have CPTs that represent combinations of all states of its parent nodes. Not every cell in a CPT must have a nonzero entry: some can be 0%. Cells also can be denoted as representing impossible combinations. Rows must total 100% (the sum of probabilities of all possible outcome states for a given set of prior conditions) but column totals do not (the sum of likelihoods of all possible prior conditions for a given state). In fact, column values in a CPT can be interpreted as likelihoods of prior conditions given an outcome state. One can normalize a column and thereby estimate the normalized likelihood function for each outcome state.

There are several practical ways in which CPT values can be established in the initial, alpha-level BBN model. If the

child node has an equation, then probabilities can be calculated explicitly (as used in the detection and occupancy nodes in Fig. 2; see Appendix A). Alternatively, if CPT values are to be at least initially specified by experts, one approach is to set all CPTs to a uniform value, "peg the corners" by setting the extreme cases to 0% or 100%, adjust the middle or most moderate case, and then back-interpolate all other entries. Another method is to represent the states in the node as discrete conditions and then identify the single most likely outcome for each combination of parent-node states, effectively forcing one outcome state for each input combination. Then the outcome states can be considered continuous values, and probabilities for each row can be adjusted to represent a reasonable probability distribution, keeping the most probable outcome to the state identified in the discrete condition. Other, more advanced statistical techniques of populating CPTs from case data are discussed in the section Updating BBN models with case files.

The values entered in a CPT can be cross-checked by scanning down each column and asking whether the entries with the highest (and lowest) values — essentially constituting the likelihood function and the maximum (and minimum) likelihood value(s) for each state — really represent the most (and least) likely causal conditions for that state. If not, the CPT values can be readjusted accordingly. Any defensible empirical information can be used to support the CPT values, and should be fully documented.

Testing and adjusting the BBN

When a CPT is fully specified, the BBN model can be compiled and the behavior of the CPT tested by trying different combinations of input values and observing the resulting probabilities in each intermediate node or the final output node(s). If the model exhibits unrealistic behavior, consideration should be given to readjusting the ill-specified CPTs, combining, splitting, or redefining nodes or their states, or readjusting the overall model structure (connections among nodes and use of summary or intermediate nodes) until it responds reasonably. Also, submodel segments can be tested separately, as can the overall model, using sensitivity analysis to determine the absolute degree and the rank order of influence of parent variables on each outcome variable in the model. CPT values then can be adjusted to make the model's sensitivity conform to known or expected behavior.

Standard sensitivity analysis uses calculations of variance reduction (VR) when dealing with continuous variables and entropy reduction with discrete or categorical variables (Appendix B). Sensitivity analysis can help verify correct initial model structure and parameterization. Sensitivity analysis for our squirrel model revealed that date has a far greater influence on detection ($VR(\text{date}) = 0.0227$) than does number of trap-nights ($VR(\text{trap-nights}) = 0.0002$). This confirms that the model is correctly representing the expectations of the expert. It should be clear that, at this point, the goal is to get the model to tell you what you think it should tell you, that is, to represent expert judgment and any initial empirical data (or equations) on how the system works. The rationale for this step should be documented, particularly the specific content and use of expert experience, literature, and initial data.

This completes the initial, alpha-level modeling. The model can be named version 0.10a, where, in modeling tradition, the version number <1 signifies that it is not ready for public release and "a" refers to the alpha level (for internal use and review only). The model files, including all significant versions and changes, should be backed up and the backups tracked with meta-documentation. This may be important for administrative purposes.

Creating the beta-level model

In this step, the alpha-level model undergoes formal peer review from (an)other subject-matter domain expert(s). The modeler can consult with at least one other species expert not involved in creating the model, to review and potentially revise the BBN model structure and CPT values. The documentation created in the alpha-level-modeling step will be useful for informing the reviewer on model structure.

In conjunction with the modeler, the peer reviewer will then review the model structure and CPT values and will either suggest edits or confirm the model's construction. If necessary, the model can be revised and thereafter treated as a competing model for later validation testing.

Following the peer review, the original domain expert(s) is (are) shown the reviewer's comments and, where appropriate, any revised model forms. Reconciliation with the review is conducted by documenting the original expert's accepted changes or other responses. Modelers should conduct the peer review and reconciliation as a formal process — as with writing and publishing papers — because it is essential for ensuring that the model is rigorously developed according to strict standards and is thus credible. The result is what can be termed a beta-level model, which can be initially named version 1.00b. Again, each version of the model should be saved as a track record of changes, along with documentation as to how and why the changes were made. This creates checkpoints and plateaus of model development to save in case they are needed for later reference.

Creating the gamma-level model

Testing, calibrating, validating, and updating BBN models is one of the more interesting and essential steps in the model-building process. Models intended for prediction purposes that are built solely on the basis of expert judgment, even with peer review, without testing represent unconfirmed belief structures or existing theories, with little assurance of reliability and accuracy. Testing and updating models derived from field data may be critical to evaluating and eliminating competing models suggested from the previous peer-review step.

Testing BBN models with case data

Case data can be used to test the accuracy of the beta-level model. One of the simpler but more useful outcomes is a confusion matrix (Kohavi and Provost 1998; Table 3), which compares predicted with actual outcomes. In the example given here, in all 23 cases in which the species was actually captured (Table 1), the model incorrectly predicted lack of capture (type I error) as the more likely outcome, but correctly predicted lack of capture in all 112 cases in which

Table 3. Example of a confusion matrix showing the prediction accuracy of the BBN model in Fig. 2 used with the case data file in Table 1.

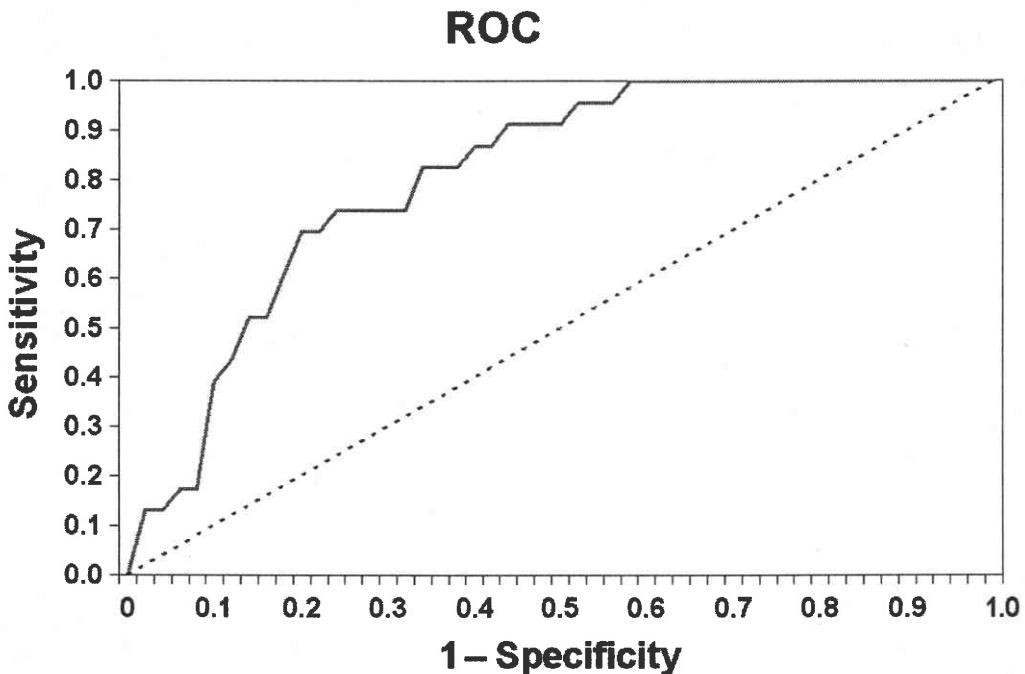
Predicted capture		Actual capture
Yes	No	
0	23	Yes
0	112	No

Note: The confusion matrix shows the number of known cases that were correctly classified (here, whether the species was actually captured or not), given the date, trapping intensity, and habitat-suitability conditions associated with each case.

the species was actually not captured. Thus, the overall model error rate is $23/135 = 17\%$, and all errors were type I (incorrectly rejecting the hypothesis that conditions would lead to a high probability of capture); no type II errors were made in this example. This error rate may seem a bit high, but implications of types of model errors (false positives and false negatives) depend on how the model will be used. Tyre et al. (2003) warned that false negatives can unduly bias prediction of species presence. Researchers may be more tolerant of type I errors (e.g., having greater certainty over conditions controlling the presence of a rare species than its absence) at earlier steps in building models to understand species–environment relations. At later steps, such as when the models are used to design monitoring programs, greater tolerance for type II errors may be permitted (such as allowing for error in predicting presence of a rare species, so that occupied sites are not inadvertently ruled to be absent and unduly disturbed). Another influence on model error rate is the problem of detectability of the species, given survey protocols (MacKenzie et al. 2003). However, indices of species detectability can be included explicitly in BBNs and in population estimates (MacKenzie and Kendall 2002) as shown in our squirrel example.

Actually, the BBN example used here is a great lesson in interpreting rates of error in predicting a relatively rare (or a very common) event, such as capture of an elusive or vagile organism or prediction of the presence of some very scarce or difficult to detect species. Confusion matrices may overstate false-negative outcomes because they rely on predictions based on the most probable outcome, whereas with relatively rare events a positive outcome (e.g., presence of a rare species) may never achieve a probability greater than that for a negative outcome (e.g., absence). Such is the case with the squirrel model, which will, at best, predict capture with <50% probability. For instance, in one case in the squirrel capture case file (Table 1), a squirrel was actually captured, whereas the BBN would predict capture under those conditions with only 39% probability and thus that no capture would be the more probable outcome. However, the 39% probability of capture is far higher than under other case conditions when no squirrel was captured. Squirrel captures (a relatively rare event) would be better compared with these actual probability predictions than just with the most probable predictions. Thus, one should recalculate error rates and confusion-matrix outcomes based on comparing actual values of predicted probabilities instead of just the most probable outcome states.

Fig. 3. A receiver operating characteristic (ROC) curve showing the results of testing the classification success of the model in Fig. 2 with the case-file data in Table 1. A ROC curve plots percent true positives (sensitivity) against percent false positives (1 – specificity).



To this end, testing the classification success rate can provide useful information on scoring success by using not the most likely state as a prediction for each case but rather the actual belief levels of each state. Some standard scoring rules for evaluating classification success rate include logarithmic loss, quadratic loss, and spherical payoff (Morgan and Henrion 1990). Values of spherical payoff, perhaps the most useful index, vary in the interval [0,1], with 1 being best model performance, and is calculated as

$$\text{MOAC} = \frac{P_c}{\sqrt{\sum_{j=1}^n P_j^2}}$$

where MOAC is the mean probability value of a given state averaged over all cases, P_c is the probability predicted for the correct state, P_j is the probability predicted for state j , and n is the number of states. In our BBN example, spherical payoff is 0.874, indicating a relatively accurate model for predicting species capture and particularly noncapture.

The classification power of a BBN model with binary output (e.g., presence or absence of a species) also can be depicted using a receiver operating characteristic (ROC) curve (Pepe et al. 2006) that assesses prediction accuracy across a continuum of prediction thresholds, rather than an arbitrary cutoff probability (such as 0.5). A ROC curve plots percent true positives ("sensitivity" in ROC parlance) as a function of percent false positives ("1 – specificity", which is the complement of the percentage of true positives; Hand 1997; Cumming 2000; see Fig. 3). Each point on the ROC curve represents the trade-off between making a true positive prediction versus a false positive prediction with increasing prediction threshold. A 45° diagonal line on the ROC curve

denotes the case of a random-guessing model that carries no classifier information whatsoever, and the best-performing model will have a ROC curve that hugs the top left of the diagram. The area under the ROC curve (AUC) varies in the interval [0,1] and is a measure of overall model performance across the full range of possible prediction thresholds. It represents the probability that a positive outcome (a squirrel capture in our example) has a higher predicted probability than a negative outcome (a noncapture) (Cumming 2000). The perfect diagnostic model will have 100% sensitivity and 0% false positives, so the closer to $\text{AUC} = 1$ the better the performance. A ROC curve with $\text{AUC} = 0.5$ represents a random model, and $\text{AUC} < 0.5$ represents a model that consistently gives the wrong answer. The ROC curve from the BBN example (Fig. 3) has $\text{AUC} = 0.81$, suggesting good predictability of a relatively rare event, that is, the trapping of a squirrel. Note, however, that this test used the data that were also used in building the model and would be better repeated with independent data.

The accuracy of different models can be compared on the same ROC diagram (e.g., Cumming 2000; Zheng et al. 2006). As with readjusting confusion-matrix outcomes to better represent model success and error rates with very rare (or very common) events, the ROC curve too is better represented on the basis of comparing actual values for model prediction probability instead of the most probable outcomes. ROC diagrams and calculations can also be extended to n Euclidean dimensions when the number of states in the outcome node is >2 (e.g., with three outcome states the area under the curve becomes a three-dimensional volume under a response surface), but they become exceedingly difficult to depict graphically.

BBN-model testing with case files can handle missing data on some nodes or states. This can be a major advantage

of using a BBN-modeling structure over a more traditional statistical analysis. An alternative to denoting missing data for input nodes is to set their states to uniform or suspected prior-probability distributions.

Updating BBN models with case files

Using case data to update models is one of the hallmarks of Bayesian statistics (Bauer et al. 1995). For example, Mac Nally and Fleishman (2004) applied a Bayesian approach to update their model of indicators of species richness by using both their model-building and their model-validation data sets. There is a rich literature on Bayesian learning and BBN-updating procedures (e.g., Gelman et al. 2003; Mitchell 1997).

Two specific ways in which case data can be used to update BBN models are (1) using test results to calibrate the model states to better align with the data and (2) using case data to automatically update the CPTs. In the first approach, calibration curves can be derived from processing case data to better determine cutoff values of states that better match the data set (Morgan and Henrion 1990). Again, one should be careful not to overfit the model to the data, particularly with small sample sizes. Some BBN-modeling shells such as Netica® provide a means of specifying the weight of individual cases or entire case files when used to update CPTs. To avoid overfitting the model to scant data, low weights can be specified.

Another way to calibrate the model is to identify misclassifications using methods other than the most probable outcome, as mentioned above. For example, the output node in Fig. 2 is a binary variable with only two states, capture or no capture. The confusion matrix and ROC model testing procedures discussed above use the most probable outcome as the model prediction for a given case, that is, when one state has a probability >50%. But each state might have different costs of misclassification. For example, in a management context, false positives (more likely with very common events) may mean that expensive species inventories are conducted or high opportunity costs are incurred by protecting a site when the species is actually absent. False negatives (more likely with very rare events) may mean that site-disturbing activities are allowed that actually cause local extirpation of the species now or in the future. Depending on management objectives, species rarity, and other factors, these two types of error can have very different implications for effects on the organism and social and pecuniary costs. Also, prediction of species presence or absence may be fraught with errors in such factors as detectability of the organism and consistency of surveys (Thompson 2004). Netica® calculates sensitivity (percent correct positives) for several probability cutoff values that it derives from the case data. Using this approach, one can determine more appropriate probability cutoff values in interpreting model predictions.

Updating CPTs to better fit the existing case data can be done using several Bayesian updating methods. One of the more popular is the expectation maximization (EM) learning algorithm (Dempster et al. 1977; Watanabe and Yamaguchi 2003). This approach automatically updates CPT values using case data so that the model better fits known examples and can handle missing data in the case files (Lauritzen

1995). Other updating algorithms include sequential empirical Bayesian analysis, learning gradient analysis, and probability fading. At least after the first pass, the updated model, which may be termed the gamma-level or final-application model, has been tested and revised using validation data sets. It will have known success rates and be calibrated to provide the best interpretation of results.

In general, calibrating and updating Bayesian models — either their overall structure or the CPTs underlying specific nodes — can be an ongoing, iterative process as new case data are gathered. This fits well into an adaptive management process of improving understanding by means of management trials (Nyberg et al. 2006).

Other considerations in BBN-model development

State-modeling objectives and expectations

As with any modeling exercise, the first step should be clear articulation of the objectives and expected use of the model. This should include identifying the geographic area and range of ecological conditions over which the model is expected to be used, the level of acceptable errors (false positives and false negatives), and how the outcome states will be shown.

Attending to coordination needs and watching for shortcuts

Managers and decision-makers who are expected to use the model outcomes should be informed and consulted at key points in the model-development process to ensure that objectives remain clear and in focus. Also, the modeler can watch for shortcuts in the BBN-modeling process, particularly if some species can be grouped into sets of BBN models with common structures. Such structures can be saved as submodels into libraries of BBN nodes and CPTs, such as libraries of GIS proxy variables representing environmental or habitat factors useful for more than one model. Such libraries can be quickly copied and pasted into new models.

Representing uncertainty and error

When developing and updating BBNs, the modeler should consider how they wish to represent uncertainty, variation, and propagation of error within and among the model variables. In one sense the spread of prior probabilities assigned to states in the input nodes, and of conditional probabilities assigned to CPTs in other nodes, represents variability or uncertainty (Toivonen et al. 2001). Continuous variables in a BBN can display weighted means and Gaussian standard deviations of their values, and BBN nodes can also represent uncertainty of continuous variables with statistical distribution functions. As BBNs are run, they execute Bayesian updating of posterior probabilities, which is one way to show propagation of uncertainty throughout a network, although strictly speaking it is not the same as analysis of propagation of error, which, in traditional statistics, entails the thorny step of estimating covariance among predictor variables (Kotz et al. 1982).

BBNs and other statistical modeling approaches

BBNs are but one tool in the modeling toolbox for researchers, resource managers, and decision-makers. Because a belief network expresses functional, probabilistic relationships between predictor variables and ecological-response variables of interest, it is akin to a widening variety of statistical methods for hierarchical modeling (where multiple parameters are structurally related and can be represented by joint probability models; see Gelman et al. 2003), nonlinear association and classification (for examples see Congdon 2001; Denison et al. 2002), and a large variety of general linear models and clustering techniques. BBNs can also be used to represent the primary relationships between variables in predictive simulation models, and can be used as a “post-processing” step to summarize results from model projections, allowing nonmodelers to explore and interpret results (for examples of this usage see McNay et al. 2006; Steventon et al. 2006).

The resource manager may wish to compare predictions derived from other modeling constructs to better understand whether results reflect model structure more than actual ecological conditions. We advocate starting with influence diagrams, not just because they are analogous to a conceptual diagram of a model system but also because they can be used as the basis for many other model forms (Marcot 2006). For instance, influence diagrams can become the basis for BBNs, graph theory or loop analyses of habitat connectivity, a path regression model, a process-simulation model, a fuzzy logic model, and other modeling approaches (e.g., Jensen et al. 1994).

BBNs belong to the school of Bayesian statistics, which has a history of tension with respect to the traditional “frequentist” school, although there are advantages to be gained by dining at both tables (Anderson et al. 2001). Some authors have helped quell some of the tension and merge aspects of the two schools by interpreting BBN posterior probabilities as probabilities of outcomes and comparing these probabilities with confidence levels in traditional statistics, such as those used in hypothesis-testing (Reckhow 1990). BBNs can be used, for example, to calculate a ratio of a likelihood function of conditions given a null hypothesis to a likelihood function of conditions given an alternative hypothesis; this odds ratio is analogous to a *P* value in classical statistics.

Conclusions and recommendations

Scientists, resource managers, decision-makers, and stakeholders all share the challenge of how to draw appropriate conclusions from scientific data, prior beliefs, and simulation modeling. As one type of tool for structuring ecological data, formulating and testing hypotheses, and exploring consequences of decisions, BBNs are proving tractable and flexible. Examples are found in other papers of this series, such as those on mapping of ecosystems (Walton and Meidinger 2006) and management of woodland caribou (*Rangifer tarandus caribou* (Gmelin, 1788); McNay et al. 2006) and marbled murrelets (*Brachyramphus marmoratus* (J.F. Gmelin, 1789); Steventon et al. 2006).

BBNs are easy to build and are thereby double-edged, combining ease of creation, explanation, and use with the potential for misuse. We suggest that the BBN modeler follow some version of the modeling guidelines presented here to avoid creating spurious or unreliable models. This entails ensuring that the relationships represented in the model structure are real and, at best, causal, and that all nodes in a BBN are empirically observable and quantifiable or at least credible and defensible through peer review. Following rigorous methods for testing and updating BBNs can greatly help in their effective application in adaptive management, where results of well-designed management experiments can be used to test management guidelines and refine BBN predictions (Nyberg et al. 2006).

When developing and updating a BBN model, the modeler should consider how it will be explained to the audience of intended users, particularly decision-makers. We have found that although BBN models are fairly intuitive, starting with the simpler influence diagram (Fig. 1) helps to explain their structure to nonmodelers. Documenting the source material and rationale for how the model was structured, including choice of variables and their states, how nodes were connected, and what was left out, as well as using experts and data sets in constructing, reviewing, and updating the model, can go far in avoiding later “battles of the experts” when the model is applied to resource decision-making.

But it is our experience that the clearest way to explain a BBN model is simply to show how it works by changing input states and demonstrating their effect on probabilities generated throughout the model. In an important sense, BBNs are holistic representations of multivariate outcomes. Showing how probabilities of these outcomes are affected by input conditions is the best way to explain their operation and use. It is also important to be very clear about the scope of inference or application to be made from such a model, such as the geographic area, ecological conditions, species, and decision context in which the model was designed, and how to scientifically interpret the model results (probability outcomes) as representations of uncertainty.

The oft-occurring “battle of the models” should be minimized by including peer review and developing, testing, and eliminating competing model structures during the model-building phases. Then, the remaining models should be used only as decision-aiding tools, not to dictate decisions. The modeler should clearly differentiate ecological prediction models from decision models; the latter, in BBNs, can incorporate decision and utility nodes that we have not discussed here (see Nyberg et al. 2006).

BBNs can complement or integrate with other kinds of models, especially for determining effects of environmental conditions on organisms. Testing, calibration, and validation updating are paramount in ensuring correct model structure, especially for prior and conditional probabilities.

Acknowledgments

Many thanks are extended to Tim Haas, who favorably reviewed an early version of our guidelines, and to Marianne Turley for reviewing the manuscript. We thank two anonymous reviewers and the Guest Associate Editor for helpful reviews. The example of the BBN model was developed

with the support of the Forest Science Program of the British Columbia Ministry of Forests and Range, and the Bulkley Valley Centre for Natural Resources Research and Management.

References

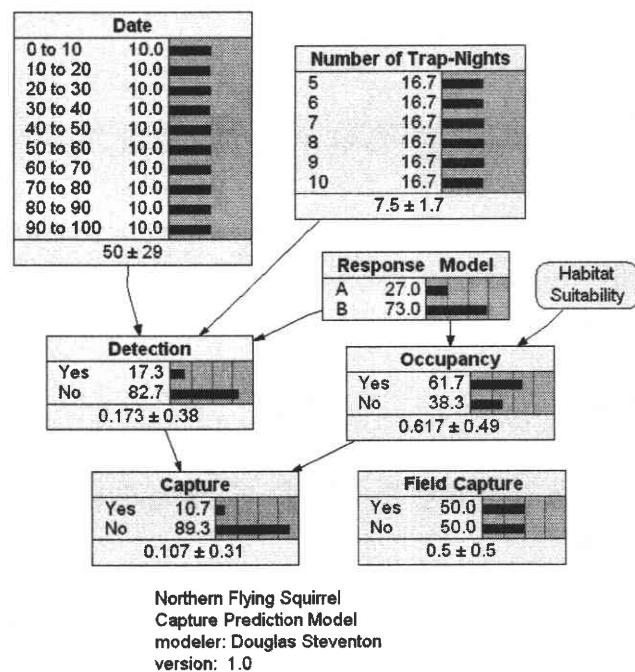
- Anderson, D.R., Link, W.A., Johnson, D.H., and Burnham, K.P. 2001. Suggestions for presenting the results of data analyses. *J. Wildl. Manage.* **65**(3): 373–378.
- Bauer, E., Koller, D., and Singer, Y. 1995. Update rules for parameter estimation in Bayesian networks. In *Uncertainty in Artificial Intelligence: Proceedings of the Eleventh Conference*, McGill University, Montréal, Quebec, 18–20 August 1995. Edited by P. Besnard and S. Hanks. Morgan Kaufmann Publishers, San Francisco, Calif. pp. 3–13.
- Cain, J. 2001. Planning improvements in natural resources management: guidelines for using Bayesian networks to support the planning and management of development programmes in the water sector and beyond. Centre for Ecology and Hydrology, Crowmarsh Gifford, Wallingford, UK.
- Cain, J.D., Batchelor, C.H., and Waughray, D.K.N. 1999. Belief networks: a framework for the participatory development of natural resource management strategies. *Environ. Dev. Sustain.* **1**: 123–133.
- Castillo, E., Gutierrez, J.M., and Hadi, A.S. 1998. Modeling probabilistic networks of discrete and continuous variables. *Int. J. Hum. Comput. Stud.* **42**(6): 48–65.
- Charniak, E. 1991. Bayesian networks without tears. *AI Magazine*, **12**(4): 50–63.
- Clark, J.S. 2003. Uncertainty and variability in demography and population growth: a hierarchical approach. *Ecology*, **84**(6): 1370–1381.
- Congdon, P. 2001. *Bayesian statistical modelling*. John Wiley & Sons, Ltd., Chichester, UK.
- Cumming, G.S. 2000. Using between-model comparisons to fine-tune linear models of species ranges. *J. Biogeogr.* **27**: 441–455.
- Dempster, A., Laird, N., and Rubin, D. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B*, **39**: 1–38.
- Denison, D.G.T., Holmes, C.C., Mallick, B.K., and Smith, A.F.M. 2002. *Bayesian methods for nonlinear classification and regression*. John Wiley & Sons, Ltd., Chichester, UK.
- Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (Editors). 2003. *Bayesian data analysis*. 2nd ed. Chapman & Hall/CRC, New York.
- Hand, D. 1997. *Construction and assessment of classification rules*. John Wiley & Sons, New York.
- Hastie, T., Tibshirani, R., and Friedman, J. 2001. *The elements of statistical learning: data mining, inference, and prediction*. Springer-Verlag, New York.
- Heckerman, D., Geiger, D., and Chickering, D.M. 1994. Learning Bayesian networks: the combination of knowledge and statistical data. In *Uncertainty in Artificial Intelligence: Proceedings of the Tenth Conference*, Seattle, Washington, 29–31 July 1994. Edited by R. Lopez de Mantaras and D. Poole. Morgan Kaufmann Publishers, San Francisco, Calif. pp. 293–301.
- Jeffers, J.N.R. 1991. Rule induction methods in forestry research. *AI Appl.* **5**: 37–44.
- Jensen, F.V. 1996. *An introduction to Bayesian networks*. Springer-Verlag, New York.
- Jensen, F., Jensen, F.V., and Dittmer, S.L. 1994. From influence diagrams to junction trees. In *Uncertainty in Artificial Intelligence: Proceedings of the Tenth Conference*, Seattle, Washington, 29–31 July 1994. Edited by R. Lopez de Mantaras and D. Poole. Morgan Kaufmann Publishers, San Francisco, Calif. pp. 367–373.
- Kalliovirta, J., and Tokola, T. 2005. Functions for estimating stem diameter and tree age using tree height, crown width and existing stand database information. *Silva Fenn.* **39**(2): 227–248.
- Kohavi, R., and Provost, F. 1998. Glossary of terms. *J. Mach. Learn.* **30**(2–3): 271–274.
- Kotz, S., Johnson, N.L., and Read, C.B. (Editors). 1982. *Encyclopedia of statistical sciences*. Vol. 2. Classification — eye estimate. John Wiley & Sons, New York.
- Lauritzen, S.L. 1995. The EM algorithm for graphical association models with missing data. *Comput. Stat. Data Anal.* **19**(2): 191–201.
- Lee, D.C., and Rieman, B.E. 1997. Population viability assessment of salmonids by using probabilistic networks. *N. Am. J. Fish. Manage.* **17**: 1144–1157.
- MacKenzie, D.I., and Kendall, W.L. 2002. How should detection probability be incorporated into estimates of relative abundance? *Ecology*, **83**(9): 2387–2393.
- MacKenzie, D.I., Nichols, J.D., Hines, J.E., Knutson, M.G., and Franklin, A.B. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, **84**(8): 2200–2207.
- Mac Nally, R., and Fleishman, E. 2004. A successful predictive model of species richness based on indicator species. *Conserv. Biol.* **18**(3): 646–654.
- Marcot, B.G. 2006. Habitat modeling for biodiversity conservation. *Northwest. Nat.* **87**(1): 56–65.
- Marcot, B.G., Holthausen, R.S., Raphael, M.G., Rowland, M.M., and Wisdom, M.J. 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. *For. Ecol. Manage.* **153**(1–3): 29–42.
- McNay, R.S., Sulyma, S.R., and Marcot, B.G. 2006. A Bayesian approach to evaluating habitat for woodland caribou in north-central British Columbia. *Can. J. For. Res.* **36**. This issue.
- Mitchell, T.W. 1997. *Machine learning*. McGraw-Hill, Boston, Mass.
- Morgan, M.G., and Henrion, M. 1990. *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge Press, New York.
- Nyberg, J.B., Marcot, B.G., and Sulyma, R. 2006. Using Bayesian belief networks in adaptive management. *Can. J. For. Res.* **36**. This issue.
- Oliver, R.M., and Smith, J.Q. (Editors). 1990. *Influence diagrams, belief nets and decision analysis*. John Wiley & Sons, Chichester.
- Pepe, M.S., Cai, T., and Longton, G. 2006. Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics*, **62**(1): 221–229.
- Ransome, D.B., Lindgren, P., Sullivan, D.S., and Sullivan, T.P. 2004. Long-term responses of ecosystem components to stand thinning in young lodgepole pine forest. I. Population dynamics of northern flying squirrels and red squirrels. *For. Ecol. Manage.* **202**: 355–367.
- Reckhow, K.H. 1990. Bayesian inference in non-replicated ecological studies. *Ecology*, **71**: 2053–2059.
- Reckhow, K.H. 1999. Water quality prediction and probability network models. *Can. J. Fish. Aquat. Sci.* **56**: 1150–1158.
- Rowland, M.M., Wisdom, M.J., Johnson, D.H., Wales, B.C., Cope-land, J.P., and Edelmann, F.B. 2003. Evaluation of landscape models for wolverines in the interior northwest, United States of America. *J. Mammal.* **84**(1): 92–105.

- Shachter, R.D. 1988. Probabilistic inference and influence diagrams. *Oper. Res.* 34(4): 589–604.
- Shachter, R., and Kenley, C. 1989. Gaussian influence diagrams. *Manage. Sci.* 35(5): 527–550.
- Smith, J.E., Holtzman, S., and Matheson, J.E. 1993. Structuring conditional relationships in influence diagrams. *Oper. Res.* 41(2): 280–297.
- Steventon, J.D., Bergerud, W.A., and Ott, P.K. 2005. Analysis of presence/absence data when absence is uncertain (false zeroes): an example for the northern flying squirrel using SAS [online]. Ex. Note 74, Research Branch, British Columbia Ministry of Forests, Victoria, B.C. Available from <http://www.for.gov.bc.ca/hfd/pubs/Docs/En/EnXX.htm> [cited 21 June 2005].
- Steventon, J.D., Sutherland, G.D., and Arcese, P. 2006. A population-viability risk assessment of marbled murrelet nesting habitat policy in British Columbia. *Can. J. For. Res.* 36. This issue.
- Thompson, W.L. (Editor). 2004. Sampling rare or elusive species: concepts, designs, and techniques for estimating population parameters. Island Press, Washington, D.C.
- Toivonen, H.T.T., Mannila, H., Korhala, A., and Olander, H. 2001. Applying Bayesian statistics to organism-based environmental reconstruction. *Ecol. Appl.* 11(2): 618–630.
- Tyre, A.J., Tenhumberg, B., Field, S.A., Niejalke, D., Parris, K., and Possingham, H.P. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecol. Appl.* 13(6): 1790–1801.
- Varis, O. 1997. Bayesian decision analysis for environmental and resource management. *Environ. Modell. Software*, 12: 177–185.
- Watanabe, M., and Yamaguchi, K. (Editors). 2003. The EM algorithm and related statistical models. Marcel Dekker, New York.
- Wisdom, M.J., Rowland, M.M., Wales, B.C., Hemstrom, M.A., Hann, W.J., Raphael, M.G., Holthausen, R.S., Gravenmier, R.A., and Rich, T.D. 2002a. Modeled effects of sagebrush-steppe restoration on greater sage-grouse in the interior Columbia Basin, U.S.A. *Conserv. Biol.* 16(5): 1223–1231.
- Wisdom, M.J., Wales, B.C., Rowland, M.M., Raphael, M.G., Holthausen, R.S., Rich, T.D., and Saab, V.A. 2002b. Performance of greater sage-grouse models for conservation assessment in the interior Columbia Basin, U.S.A. *Conserv. Biol.* 16(5): 1232–1242.
- Zhang, N.L. 1998. Probabilistic inference in influence diagrams. *Comput. Intell.* 14(4): 475–497.
- Zhang, L., and Poole, D. 1996. Exploiting causal independence in Bayesian network inference. *J. Artif. Intell. Res.* 5: 301–328.
- Zheng, Y., Cai, T., and Feng, Z. 2006. Application of the time-dependent ROC curves for prognostic accuracy with multiple biomarkers. *Biometrics*, 62(1): 279–287.

Appendix A. Northern flying squirrel (*Glaucomys sabrinus* (Shaw)) Bayesian belief network model

The example of Bayesian belief network (BBN) is based on pilot-study data examining whether the occupancy of otherwise suitable habitat by northern flying squirrels was affected by the surrounding forest context. In addition, for designing future field efforts we wanted to know the influence of trapping date on capture efficiency; Ransome et al. (2004) suggested that capture rates are much lower in the summer than in the fall. Two live traps were placed for five or seven nights at each of 135 mature-forest locations representing a range of habitat quality in the 16 ha square sur-

Fig. A1. Example of an initial parameterized Bayesian belief network model based on the simple influence diagram (Fig. A2). Nodes are defined as follows: “capture” is the overall predicted probability of capturing a squirrel at least once at a site; “detection” is the probability of detecting a squirrel if it is present at a site; “occupancy” is the probability that a site is occupied by a squirrel; “response model” denotes the hypotheses that detectability is dependent on trapping date and that occupancy is either constant or varies according to habitat suitability; “habitat suitability” denotes suitability of a habitat for the squirrel in a 16 ha area around each trap site; “date” is the number of days from the start of the trapping session on 22 June; “number of trap-nights” is the number of nights when traps were open at each site; “field capture” denotes field capture data (yes/no) used in assessing model accuracy. The model is shown here running the prior probabilities of input conditions derived from field studies and model parameterization. The additional unlinked field capture node represents the actual observation at each field site (capture or noncapture). This is used when processing the field-data case file to associate the field result with the predicted capture probability for use in assessing model prediction accuracy.



rounding each site (the “habitat suitability” input node in the BBN model shown in Fig. A1).

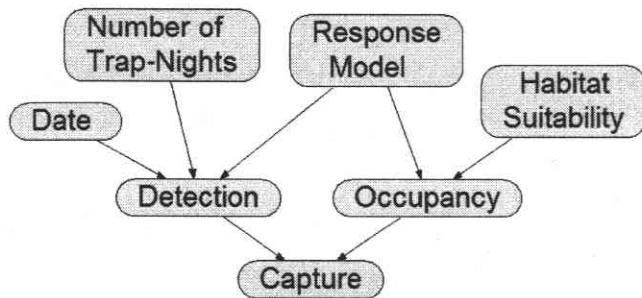
The following are node definitions and the equations applied in the BBN to populate the CPTs. Equations are based on the zero-inflated binomial MLE method of Tyre et al. (2003) applied in SAS® version 8 (Steventon et al. 2005).

Capture: The predicted probability [0,1] of capturing a squirrel at least once, conditional on detection probability (“detection” node) and occupancy probability (“occupancy” node), which both vary in the interval [0,1]:

$$P_{\text{(capture)}} = P_{\text{(detection)}} P_{\text{(occupancy)}}$$

Detection: The probability [0,1] of catching a squirrel at least once if it is present, conditional on “date”, “number of trap-nights”, and “response model” node states.

Fig. A2. Example of a simple influence diagram showing key factors affecting expected capture of northern flying squirrels (*Glaucomys sabrinus* (Shaw)). Capture is shown to be dependent on detectability, which is influenced by date and trapping effort, and by occupancy, which may or may not be influenced by suitability of local habitat (forest) conditions (response model). Since such diagrams are best used as a basis for further modeling, the correlates shown should also express causal relationships, so that the diagram becomes a “causal web”.



If the response model node assumes state A, then

$$P_{(\text{detection})} = 1 - [1 - e^{-9.535+0.084\text{date value}} / (1 + e^{-9.535+0.084\text{date value}})]^m$$

If the response model node assumes state B, then

$$P_{(\text{detection})} = 1 - [1 - e^{-9.489+0.084\text{date value}} / (1 + e^{-9.489+0.084\text{date value}})]^m$$

“Date value” is the date as a sequential number from 22 June and m is the number of trap-nights. The expression within the braces is the probability of *not* detecting a squirrel on any single trap-night if the site is indeed occupied: 1 minus the log-odds of detection (Tyre et al. 2003, eq. 4) converted to the probability scale. The probability of *not* detecting a squirrel at least once (if it is present) over m nights is the product of the individual trap-night probabilities (the single-night probability to the power m assuming constant probability). Finally, 1 minus that probability is the probability of detecting a squirrel on m trap-nights (if it is present).

Occupancy: The probability [0,1] of site occupancy conditional on the habitat suitability (x) and response model nodes. For response model node state A, occupancy probability is

$$P_{(\text{occupancy})} = e^{0.396-0.295x} / (1 - e^{0.396-0.295x})$$

For response model node state B, occupancy probability was estimated to be constant at 0.64. The equation is based on transforming the log-odds of occupancy versus non-occupancy (Tyre et al. 2003, eq. 4) to the probability scale.

Response model: Represents two a priori assumptions: state A — that detection probability is dependent on date and occupancy probability is constant; state B — that detection probability is dependent on date and occupancy probability on habitat suitability. The weighting is based on the small-sample Akaike’s Information Criterion (AIC) weight for each response model (Steventon et al. 2005).

Habitat suitability: An area-weighted habitat-suitability

score (D. Steventon, unpublished data) from GIS for a 16 ha square surrounding a trap site. Each 1 ha raster cell in the 16 ha square is assigned a value (0–1) based on the forest-inventory attributes (British Columbia Ministry of Forests and Range) stand age, site productivity, and biogeoclimatic classification to the variant level. The values for all 16 raster cells are then averaged to give the habitat-suitability score (0–1) applied in the BBN.

Date: A consecutive number from 22 June, representing the date of the start of each trapping session.

Number of trap-nights: The number of nights when traps were open at each site (five or seven).

Field capture: The additional unlinked node “field capture” represents the actual observation at each field site (capture or noncapture). This is used when processing the field-data case file to associate the field result with the predicted capture probability for use in assessing model prediction accuracy.

References

- Ransome, D.B., Lindgren, P., Sullivan, D.S., and Sullivan, T.P. 2004. Long-term responses of ecosystem components to stand thinning in young lodgepole pine forest. I. Population dynamics of northern flying squirrels and red squirrels. *For. Ecol. Manage.* **202**: 355–367.
- Steventon, J.D., Bergerud, W.A., and Ott, P.K. 2005. Analysis of presence/absence data when absence is uncertain (false zeroes): an example for the northern flying squirrel using SAS [online]. Ex. Note 74, Research Branch, British Columbia Ministry of Forests, Victoria, B.C. Available from <http://www.for.gov.bc.ca/hfd/pubs/Docs/En/EnXX.htm> [cited 21 June 2005].
- Tyre, A.J., Tenhumberg, B., Field, S.A., Niejalke, D., Parris, K., and Possingham, H.P. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecol. Appl.* **13**(6): 1790–1801.

Appendix B. Calculations of model sensitivity

The following formulae are used to calculate model sensitivity in the modeling shell Netica® (B. Boerlage, personal communication). Variance reduction (VR) is the expected reduction in the variation, $V(Q)$, of the expected real value of the output variable Q having q states due to the value of an input variable F having f states, and is calculated as $VR = V(Q) - V(Q|F)$, where $V(Q) = \sum_q P(q)[Xq - E(Q)]^2$, $V(Q|F) = \sum_q P(q|f)[Xq - E(Q|f)]^2$, $E(Q) = \sum_q P(q)Xq$, where Xq is the numeric real value corresponding to state q , $E(Q)$ is the expected real value of Q before any new findings, $E(Q|f)$ is the expected real value of Q after new findings f for node F , and $V(Q)$ is the variance of the real value of Q before any new findings. Entropy reduction, I , is the expected reduction in mutual information of Q (measured in information bits) due to a finding at F , and is calculated as

$$I = H(Q) - H(Q|F) = \sum_q \sum_f \frac{P(q, f) \log_2[P(q, f)]}{P(q)P(f)}$$

where $H(Q)$ is the entropy of Q before any new findings and $H(Q|F)$ is the entropy of Q after new findings from node F .