International Congress on Environmental Modelling and Software

5th International Congress on Environmental Modelling and Software - Ottawa, Ontario, Canada - July 2010

Jul 1st, 12:00 AM

Guidelines for Good Practice in Bayesian Network Modelling

Serena H. Chen

C. A. Pollino

Guidelines for Good Practice in Bayesian Network Modelling

Serena H. Chen and Carmel A. Pollino

Abstract

Bayesian networks (BNs) are used increasingly to model environmental systems, for reasons including their ability to: integrate multiple issues and system components; utilise information from different sources; and handle missing data and uncertainty. For a model to be of value in generating and sharing knowledge or providing decision support, it must be built using good modelling practice. This paper provides such guidelines to developing and evaluating Bayesian network models of environmental systems. The guidelines entail clearly defining the model objectives and scope, and using a conceptual model of the system to form the structure of the BN, which should be parsimonious yet capture all key components and processes. After the states and conditional probabilities of all variables are defined, the BN should be evaluated by sensitivity analysis, expert review and testing with cases. All the assumptions, uncertainties, descriptions and reasoning for each node and linkage, data and information sources, and evaluation results must be clearly documented. Following these minimum standards will help ensure the modelling process and the model itself is transparent, credible and robust, within its given limitations.

Introduction

Environmental systems are inherently complex and there is often a high degree of uncertainty of the interactions of system components. While traditional statistical modelling approaches can be used for some models of single components or processes within the system (e.g. rainfall-runoff models, hydraulic models), integrated modelling approaches are often required for whole-of-system models or models incorporating multiple system components. Commonly used integration methods include Bayesian networks (BNs), system dynamics, coupled component models, agent-based models and expert systems. These methods vary in their knowledge and data requirements, technical requirements, treatment of uncertainty, and application suitability (Jakeman et al. 2007).

This paper focuses on Bayesian networks, as it is an approach considered highly suitable for environmental problems due to its ability to integrate multiple issues, interactions and outcomes and investigate tradeoffs. Furthermore, Bayesian networks are apt at utilising data and knowledge from different sources, and handling missing data and uncertainty. Bayesian networks are based on a relatively simple causal graphical structure, which means it can be built without highly technical modelling skills and it can also be understood by non-technical users and stakeholders. This is a very valuable feature of Bayesian networks, particularly in the context of natural resource management which (ideally) involves interdisciplinary and participatory processes.

There is great benefit in the use of modelling as an approach to understanding and supporting decisions on environmental systems. However, for a model to be of value, good practice in its construction, testing and application is essential, as is awareness of the purposes, capabilities and limitations of the modelling approach. Without this, there is a risk of the model user misinterpreting or misusing model outputs, and drawing invalid conclusions (Jakeman et al. 2006). For the model user to be aware of the modelling objectives, assumptions and limitations, the modeller needs clear reporting protocols. Poor modelling practice reduces the credibility of the model and can lead to the model capabilities being 'oversold', potentially causing poor decisions to be made based on models, or where model transparency and testing has not been completed, users mistrusting models and their outputs (Refsgaard and Henriksen 2004). Consequently, guidelines for good modelling practice that create standards to help ensure the development and application of credible and purposeful models are essential.

Several authors have developed modelling guidelines (Refsgaard and Henriksen 2004, Jakeman et al. 2006, Crout et al. 2008), where the key components for good practice include:

- Clearly defining model purpose and the assumptions underlying the model
- Thorough evaluation of the model and its results
- Transparent reporting of the whole modelling process, including its formulation, parameterisation, implementation and evaluation

Good modelling practice will result in better understanding of the development and application of models; this benefits not only the modelling community but also model users who employ the models for improving knowledge of the system or decision making.

The objective of this paper is to introduce guidelines to developing and evaluating Bayesian network models of environmental systems. As with models in general, there is a need for quality assurance standards in developing and applying Bayesian network models. Bayesian network protocols have been published by Cain (2001) and Marcot (2006). Cain (2001) provided guidelines to using BNs for supporting planning and management of natural resources, with a large emphasis on facilitating stakeholder consultation. In the context of natural resources management, stakeholder consultation is seen as essential to ensuring that the management plan is followed through and implemented (Cain 2001). Marcot (2006) developed guidelines for Bayesian networks applied to wildlife and ecological assessment, with the steps to developing and updating the BNs described at three model levels: alpha, beta and gamma. The alpha-level model is the initial functioning BN, suitable only for internal use and review. The BN is considered a beta-level model after formal peer review and revision is conducted. The gamma-level or final application model, is created by further testing, calibrating, validating and updating the beta-level model (Marcot et al. 2006).

This paper will explore the development process of Bayesian network models, following the generic guidelines for good modelling practice outlined by Jakeman et al. (2006). These guidelines consist of ten iterative steps (Jakeman et al. 2006):

1. Define model purpose
2. Specify modelling context (scope and resources)
3. Conceptualise the system, specify data and other prior knowledge

4. Select model features and families
5. Decide how to find model structure and parameter values
6. Select estimation performance criteria and technique
7. Identify model structure and parameters
8. Conditional verification and diagnostic testing
9. Quantify uncertainty
10. Model evaluation and testing

The paper is intended to serve a wide readership. It is envisaged that adhering to the proposed guidelines will enhance their quality and value in generating and sharing knowledge on environmental systems and providing advice on their management.

Bayesian networks in Natural Resources Management

In BN models, the studied system is represented as a complex network of interactions from primary cause to final outcome, with all causal assumptions made explicit (Borsuk et al. 2006). Evidence is entered into the model by substituting the a priori beliefs of one or more nodes (variables) with observation or scenario values. Through belief propagation using Bayes' Theorem, the a priori probabilities of the other nodes are updated. This belief propagation enables BNs to be used for diagnostic ('bottom-up' reasoning) or explanatory purposes ('top-down' reasoning) (Castelletti and Soncini-Sessa 2007). So unlike black-box models, such as neural networks, BN users can find out the reasoning behind the model outputs as interactions between variables are clearly displayed. This not only provides clarity to users, but also promotes system learning and increases the transparency of management decisions.

BNs can effectively integrate information from a range of sources and are also able to integrate submodels, even those representing different scales (Borusk et al. 2004). BNs can be used to predict future states/events even when data is partial or uncertain (Park et al. 2005). This is a huge advantage over many other traditional statistical models which often rely on large amounts of empirical data to be built (Marcot et al. 2006). However, as with all modelling approaches, BNs are limited in some respects. BNs are Directed Acyclic Graphs, so they cannot represent feedback loops, which often occur in nature. Also, BNs generally represent static relationships over given temporal scales, although some software packages can handle dynamic models by representing each time slice with a separate network (Kjærulff 1995). BNs can also be useful to decision-makers as the model can be used to investigate tradeoffs. For example, BNs can be used to test and compare the forecasted system response to alternative policy or management options (e.g. Ticehurst et al. 2007), which can help inform managers on which scenario is likely to produce the optimal outcome based on the information given to the network. More details about the advantages and limitations of BNs in environmental modelling can be found in Castelletti and Soncini-Sessa (2007) and Uusitalo (2007).

In the environmental domain, BNs are often used to integrate information about the factors influencing certain aspects of a species, community or system component, to aid management. Examples of such BNs include habitat and population viability models of at-risk fish and wildlife species in the Columbia River basin (Marcot et al. 2001), a dynamic, age-structured population model of brown trout in Swiss Rivers, for assessing the relative influence of different stress factors (e.g. water quality, habitat conditions, stocking practices, flood frequency) to indicate the type of management actions that would be most effective in protecting their populations (Borsuk et al. 2006) and an eutrophication model for the Neuse estuary, North Carolina, developed to quantify the relationship between nitrogen loading and other relevant variables (e.g. shellfish population, size and frequency of algal blooms, fish kill) and assist decision makers who were considering new legislation on total maximum daily load of nitrogen (Borusk et al. 2004).

BNs can effectively integrate physical, social, ecological and economic components of a system into a model. Accordingly they have been applied as integrated models used as decision support tools for testing the impact of various management strategies (pricing, awareness-education, grey water reuse and leak) on domestic water consumption in the Loddon catchment, England (Bromley et al. 2004), assessing the ecological impacts of salinity management scenarios for the Litter River Catchment, Macquarie River basin, NSW (Sadoddin et al. 2005) and exploring the impact of various climate change scenarios on natural resource condition targets (e.g. Red Gum growth rate, bird breeding event, Macquarie Marshes water quality) in the Central West region, NSW (Tighe et al. 2007). BNs can also be applied as risk assessment models, as in Pollino et al. (2008) where a BN was developed to predict the impact of mine-derived heavy metals to the environment and human health in the Ok Tedi and Fly River, Papua New Guinea. In Pollino et al. (2007a), BNs were used as a modelling framework to examine conflicting hypotheses on the main causes of dieback in the Swamp Gum in the Yellingbo Nature Conservation Reserve, Victoria, and make recommendations for future monitoring and research. As these examples demonstrate, there is an enormous scope for the possible applications of BNs in natural resources management.

Guidelines to good practice in Bayesian network modelling

1) Define the model purpose

Clearly defining the model purpose and scope is the first key component of most modelling guidelines (Cain 2001, Jakeman et al. 2006, Crout et al. 2008). It is important that the objectives of the modelling exercise are clear from the beginning, to ensure that the network is built to fulfil the right purpose and captures all the relevant ideas. The model purpose influences many of the choices in the modelling development process, including what variables or information to include, the level of detail required, the complexity of the structure, and the scales considered. Model purpose also determines the role of uncertainty and how the uncertainty should be handled (Brugnach et al. 2008). Furthermore, purpose drives the model evaluation process; without specifying the purpose its success cannot be assessed (Crout et al. 2008).

Motives for developing and applying Bayesian network models can include:

- Improving system understanding
- Social learning
- Knowledge discovery
- Synthesizing or encoding knowledge and data
- Prediction
- Exploratory and scenario analysis
- Tradeoff analysis
- Informing and supporting management and decision making
- Identifying knowledge and data gaps

These are not mutually exclusive. As seen in most of the examples in the previous section, a BN can be built for more than one purpose,

2) *Specify the modelling context*

As well as defining *why* you are modelling, it is essential to clearly state *what* you are modelling. This step involves identifying the scope of the model. This scope should include:

- the specific problems and issues to be addressed
- the geographical area to be modelled, including scale and resolution
- the time period considered
- the anticipated outputs (model endpoints)
- the key drivers

Setting the boundary will help the model to stay focussed on the relevant details. As environmental systems are typically highly complex, if the scope is not clearly defined in the early stages of the study, it is easy to lose focus and waste time on collecting unnecessary information. The latter two points (anticipated outputs and key drivers) also help to clarify what information you hope to gain and what you want to investigate with your model. For instance, if the model purpose is to guide management of a catchment by exploring alternative river restoration options, you should specify what your management objectives are – e.g. improved water quality, improved frog habitat (anticipated outputs), and what management interventions you wish to consider – e.g. riparian revegetation, environmental flows (key drivers).

The level of involvement and collaboration with domain experts or stakeholders depends on the model purpose and scope. If stakeholders are to be affected by decisions based on the BN model, it may be necessary to engage them from this early stage. Model end users (e.g. decision makers, managers) should also be consulted to ensure that the modeller and model users are on the same page regarding the purpose and design of the model. Stakeholders and users can be consulted to help identify values and assets in the system and also relevant issues of concern (Haag and Kaupenjohann 2001). Involving stakeholders in the design and development of the model also enhances the acceptance of the final decision by strengthening their sense of ownership of the decision process (Bromley et al. 2004). If the BN is built to inform science, stakeholder participation is not necessary, but rather expert review should be conducted. As discussed later, peer expert review is an important form of model evaluation when data is limited.

3) Conceptualise system, specify data and other prior knowledge

This step involves finding out what is already known about the system. This step may consist of a review of existing literature on relevant issues on the modelled system, or on similar systems. If there is an adequate amount of data available, data mining may be useful in identifying patterns in the data and potential relationships between variables (Gibert et al. 2008). This step is a process of identifying the key variables of the system and understanding how they (may) interact. Such knowledge can also be elicited from experts in the domain. At this stage, it may also be useful to start identifying relevant experts who may be able to provide peer reviews of your model (at various stages of it construction) or may be able to provide their knowledge to help in parameterising your model particularly if data on your system is scarce.

The existing knowledge should be synthesized into a conceptual model of the system, in the context of the BN model purpose and scope. More than one conceptual model can be built, for example from different scales, perspectives or levels of detail. The aim of this step is to provide a visual summary of how the drivers (e.g. climate, policy, management intervention) are linked to other variables and the output(s). Therefore it entails identifying the variables that are considered to directly or indirectly impact upon the final output(s) and describing the assumptions about the system processes that link them. Jakeman et al. (2006) advise to always conduct this conceptualisation step even if the model is not being built from scratch, as it can help to expose weaknesses in the underlying assumptions of the model.

After the conceptual model is built, it should be reviewed by a panel of relevant experts and revised if necessary. This feedback may correct any errors in the conceptual model and help identify key variables or processes that were overlooked. The reviews will help ensure that the assumptions and notions described in the model are generally accepted.

In some cases it may be appropriate to build the conceptual models together with stakeholder groups, particularly when the model will be used as a management tool whose outcome will affect the stakeholders. Cain (2001) provides guidelines on facilitating stakeholder consultation in the development process of BNs. In these guidelines, he suggests building separate conceptual models with each of the stakeholder groups to represent each of their perspectives, thereby identifying issues of consensus and conflict between the groups. After joint workshop discussions with the stakeholders, the conceptual models are combined into one (Cain 2001).

4) Select model features and families

The selection of model features and families is dependent on modelling purpose, objectives, and the quality and quantity of prior knowledge and data. Before proceeding with BNs as a modelling approach, you must carefully assess whether it is most appropriate, as other modelling approaches may be more suitable. Modelling features of BNs include their:

- 'white-box' nature, such that relationships between variables are made explicit;
- inability to readily represent feedback loops;
- ability to integrate information from a range of sources;
- ability to integrate different sub-models (e.g. social, ecological and economic);
- ability to be easily updated;
- user-friendliness; and
- limited ability to deal with continuous data.

BNs can be useful for cases:

- where knowledge on the system is poorly structured or involves a high level of uncertainty;
- that have limited/incomplete data on key system variables;
- requiring both qualitative and quantitative information, or data in different forms;
- integrating several system components;
- requiring stakeholder engagement in the modelling process; and/or
- where the relationships between variables are non-linear and complex.

Other modelling techniques may be more suitable in cases:

- involving complex feedbacks, particularly if these feedbacks are important with respect to the model outcome; or
- where there is a lot of data and/or the processes can be effectively described by mathematical equations.

5) *Decide how to find model structure and parameters*

The structure of a BN can be found from domain knowledge and/or data. It is recommended that the structure of BNs is built based on existing theories, knowledge or hypotheses. BNs are capable of structural learning from data using a score-based algorithm, which searches for a structure that maximises the chosen entropy scoring function, or a constraint-based algorithm, which maps out the model structure based on the conditional dependencies found between each pair of variables (Cheng and Greiner 2001, Cansado and Soto 2008). However, structural learning without any prior knowledge or input from the user is considered to be of limited suitability to environmental problems due to the highly stochastic and uncertain nature of environmental processes (Uusitalo 2007). Even if there is high uncertainty of the accuracy of the conceptual models, generally, more reliable BN models are built using prior knowledge, rather than learning solely from data sets (Uusitalo 2007). In cases where different theories or hyotheses about the system exist, separate BNs can be built for each then compared, as demonstrated by Pollino et al (2007a).

All nodes included in the model must affect (or be affected by) the final output; if a node does not, it can be left out. The node should also either be: i) manageable, ii) predictable or iii) observable at the relevant scale of the model (Borusk et al. 2004). The exception to this are aggregate nodes (also referred to as latent or intermediate nodes), mentioned below. The inclusion of insignificant variables can increase the complexity of the network and reduce the sensitivity of the model, not to mention unnecessarily cost extra time and effort, without adding any value to the overall model. Model parsimony is key; keep the network as simple as possible.

BNs cannot contain any cyclic loops. In cases where the system does contain some feedback, you should refer back to Step 1 and consider the relative degree of influence the processes have in the context of the model objectives. Typically you will find that one direction of flow may be of minor importance to the model outcome relative to the other processes, and therefore can be left out. Often no data or good knowledge for representation of feedbacks is available. However if such feedbacks require representation, Dynamic Bayesian Networks can be applied (see Murphy 2002).

The strength of the relationships between nodes are quantified in the Conditional Probability Tables (CPTs) attached to each node. Parentless nodes are described by marginal probability distributions, rather than conditional probabilities. For each child node, conditional probabilities are allocated for each combination of states in their parent nodes, so the size of each of these CPTs depends on the number of parent nodes and the number of their states. This CPT size can increase exponentially, which can make the process of filling the CPTs intractable especially if this is done through expert elicitation. It is generally recommended that each node should have no more than three parent nodes (Marcot et al. 2006).

The approach to obtaining the conditional probabilities will depend on the type and amount of data you have access to. These can include:

- Datasets, from field monitoring or laboratory studies
- Process equations, derived from peer-reviewed studies or models
- Datasets, derived from models
- Information elicited from experts or stakeholders
- A combination of the sources above

Each of these data sources has their advantages and limitations, and it is crucial that the modeller is aware of these and conveys them to the user.

When using datasets collected by direct measurements, the data must represent how the node changes according to changes to the states of the parent nodes, in other words the observed state given the combination of parent nodes. Each of these data samples is referred to as a case. The inherent stochasticity and uncertainty of environmental systems is likely to be reflected in the dataset, with cases varying even with the same combination of parent nodes. Accordingly, the accuracy of the conditional probabilities will increase with a larger number of cases. It is recommended there be at least 20 cases for each combination of states of the parent nodes (Cain 2001); this helps to avoid overfitting the model. The conditional probabilities can be learnt from data using algorithms such as Laurizen Spiegelhalter algorithm (a basic representation of Bayes' theorem), Expectation Maximization (EM) or Gradient Descent (GD), which are built into most BN software. These algorithms estimate the conditional probabilities based on the network structure and dataset. Both EM and GD algorithms can approximate probability distributions for datasets containing missing values.

If available, another alternative is to use datasets generated from models. Such models may have been calibrated by measured data. Methods such as Monte Carlo sampling are suitable for generating cases from models. Monte Carlo sampling performs repeated runs of the model with different sets of inputs and model parameters, which are randomly varied within defined limits (Cain 2001). A case is produced with each run and conditional probabilities can be learnt from these cases using the learning algorithms.

If no appropriate datasets or models are available to parameterise the BN model, expert judgement based on past observation and experience, can be used to estimate conditional probabilities. The ability of BN models to be parameterised using expert opinion is an advantage particularly for environmental systems that simply do not have the required quantitative data which would be necessary for statistical modelling approaches (see Smith et al. 2007). However the problems with the inherent subjectiveness of expert opinion must be considered. Numerous factors can limit a person's judgement and estimation of quantities (regardless of their expertise) including heuristics, biases, values, attitudes and motivations. See Burgman et al. (2006) for a comprehensive review of techniques for eliciting expert judgment.

It is also possible to combine expert elicitation with data or model outputs, to specify the CPTs (Pollino et al. 2007b). Expert judgement can be used to provide an initial estimate of the probabilities (i.e. prior probabilities). Uncertainty values attached to these prior probabilities can be expressed as the number of observed cases your degree of knowledge is equivalent to (Cain 2001). The prior probabilities can then be updated using the available observed data.

6) Select estimation performance criteria and technique

The performance criteria must be contingent on the model purpose and objectives. The desired outcome of the modelling exercise can include an acceptable prediction performance or the model exhibiting realistic/plausible behaviour. Ideally, the accuracy of the model should be tested with empirical data, however in some cases this data is not available. Data independent from that used to parameterise the model should be used for testing. You may need to randomly split your data set into two parts, one for training and the other for testing (e.g. 80%/20%). Most BN software have a function that allow a set of data to be tested against model predictions or diagnosis. The software updates the probability values of all the samples within the case, except the 'unobserved nodes' (i.e. the nodes you have selected to be predicted/diagnosed), and then generates beliefs for each unobserved node. This generated value is then compared with its true value, and this is repeated for all given cases. One of the outputs for the test is a confusion matrix, which compares the predicted with actual outcomes. The most likely state is chosen as the model's prediction for that case. The columns in the matrix represent the instances in a predicted state and the rows represent the instances in the actual state and the number of cases is tallied up accordingly (Norsys 2003). In this case the performance criteria can be a certain error rate (e.g. <5%).

Another form of quantitative evaluation is a sensitivity analysis, which can be conducted to identify sensitive parameters. Sensitivity analyses typically apply variance reduction calculations to continuous variables, and entropy reduction calculations to discrete or categorical variables. These analyses essentially rank the variables in order of importance relative to the variable of interest, which is generally the final output. Sensitivity analysis can be used to verify whether the model's response is correctly conforming to expectations. Sensitivity analysis can identify which variables have the most influence on the final outcome, and subsequently these variables indicate priority risks or key knowledge gaps (Pollino et al. 2007b).

If data on the system, in particular for the model output node is limited or unavailable, qualitative forms of model evaluation, such as peer review, are valuable. By applying different combinations of inputs, and examining the resulting probabilities throughout the network, reviewers can test whether the behaviour of the model is consistent with current understanding about the system.

7) *Identify model structure and parameters*

The model structure and parameters should be identified from domain knowledge and data, as described in step 5. The final structure of the BN should be parsimonious, by having as few nodes as necessary and keeping the CPTs to a manageable size. The size of a CPT is determined by the numbers of parent nodes and their states. One approach to reducing the number of parent nodes, and thereby simplifying a BN is by 'divorcing' nodes (Fig. 1) (Henderson et al. 2009). This process involves aggregating a few of the nodes by adding a new node (A) that summaries themes. This can only be done if the aggregations are logical and no interactions are lost in the procedure. Although this process adds nodes to the network, it actually reduces the combined size of CPTs in the network (Cain 2001). It should be noted that divorcing may, to some degree 'dilute' the sensitivity of the model, as it increases the number of nodes between the input nodes and the final node(s). Increasing the number of layers of nodes also increases the uncertainty propagated through the network (Cain 2001).

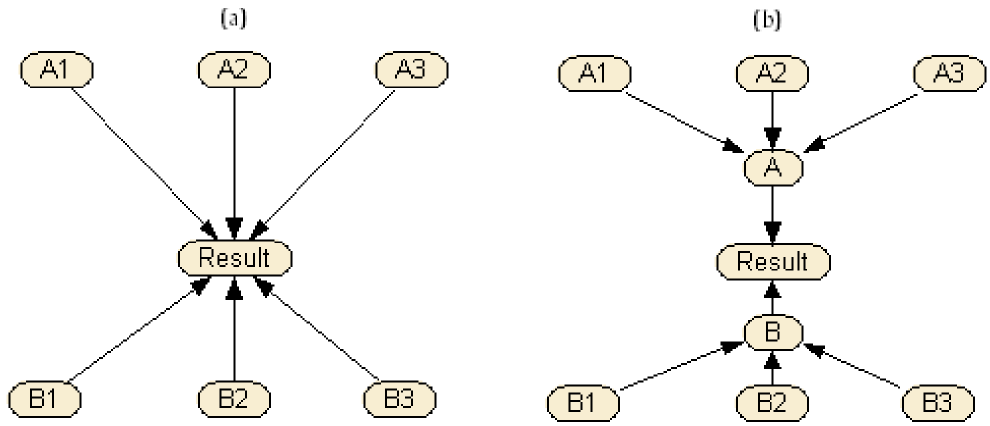

Fig. 1: Divorcing, where nodes ‘A1’, ‘A2’ and ‘A3’ and ‘B1’, ‘B2’ and ‘B3’ in (a) are divorced from the ‘Result’ node by including intermediate nodes ‘A’ and ‘B’ (b). (Source: Henderson et al. 2009)

Marcot (2006) suggests that the BN network should have less than five layers of nodes if possible, to avoid ‘diluting’ pathways and increasing the uncertainty propagated to the final output. This of course is not always possible, and depends on the complexity and scale of the modelled system. Furthermore, if the network structure is asymmetric, such that some input nodes are closer (in terms of the number of layers between them) to the final node than others, this creates uneven sensitivity of the various input nodes to the outcome (Marcot et al. 2006).

In discrete BNs, each node has a set of mutually exclusive states, which can be categorical, Boolean, continuous or discrete. Continuous variables, however, must be discretised into a finite set of states (e.g. <5, 5-10, >10) in most BN software. As the size of the CPTs depends on both the number of parent nodes and the combined number of their states, it is better to have the fewest number of states necessary in each node. It is recommended to select states that describe: 1) the current state the variable is in, 2) the state it may shift to under the system change scenarios, and 3) if necessary, any intermediate states (Cain 2001). Unless the modelling exercise is for purposes such as risk assessment or disaster management, a state should not be included if it is unlikely to be reached or is not relevant to the model objectives. The states must also match the logic of the network. To ensure this, check that the states of the parent and child nodes directly affect or are directly affected by the states in each node.

One of the drawbacks of BNs is the potential loss of statistical accuracy through discretisation of continuous variables. This loss of information is particularly the case when dealing with linear relationships. On the other hand, such discretisation can easily capture non-linear relationships and complex variable distributions, such as bi- or multimodal distributions, with relatively little computational power (Henderson et al. 2009). Choosing the number of intervals requires a compromise between model simplicity and accuracy. Consider what level of detail of the node states may actually affect the final output. For example if the node is representing rainfall and we know that annual rainfall below 500mm is too low for the plant species we are considering, above 1000mm is too high, and levels in between are fine, rather than having multiple states which lead to the same result (e.g. <250, 250-500, 500-750, 750-1000, etc) we should just include those that are relevant to the outcome (i.e. <500, 500-1000, >1000). Ideally, the continuous values are divided at points where there are breakpoints or thresholds relevant to the child nodes or model objectives. Another consideration, particularly if the CPTs will be learnt from data, may be whether the intervals contain a reasonable number of observations.

8) Conditional verification and diagnostic testing

This step involves conditionally verifying and testing the model, to ensure that the model's interactions and outcomes are feasible and defensible. The behaviour of the parts of or the whole model can initially be tested by applying different scenarios (i.e. combinations of inputs) and examining whether the resulting probabilities are reasonable and logical. Domain experts should also be reviewed by a panel of relevant experts, to ensure that the model contains all key components and processes, is structurally sound and is consistent with current understanding about the system.

This step should also involve measuring the other performance criteria identified in step 6. If the results are not plausible or if the model does not behave in a manner that is feasible or defensible, it may be necessary to reassess the model structure and assumptions. This may involve readjusting the network structure, fine-tuning the questionable CPTs, or combining, separating or redefining nodes or states (Marcot et al. 2006).

9) Managing Uncertainty

Due to the complexities of environmental systems, all models contain uncertainties. Uncertainties can stem from a range of sources, and manifest themselves in the parameter values, data, structure and framing of the model. It is crucial that these model uncertainties are managed appropriately in the model development process and are conveyed to model users and in association with any model outcomes. Uncertainties must be expressed to users, particularly if the model is to be used to inform decisions in management or policy. This allows users to consider the desirability of a certain outcome against the risk of it not being achieved when making their decision (Cain 2001).

Uncertainties in a BN model can originate from: an incomplete understanding of the system processes; the stochastic nature of environmental systems; incomplete, finite or imprecise data; and the subjective biases of expert elicitation of conditional probabilities. These probabilities are expressed through the distribution of probabilities assigned to the states of nodes. The probabilities in the network are propagated through to the final model endpoints. However, no distinction is made between uncertainty, variation and propagation of error.

One type of uncertainty that is not accounted for in BN models is concerning the causal structure of the network. This structural uncertainty can result in an underestimation of the uncertainty of the BN model outputs (Mead et al. 2006). Expert review of the conceptual model and BN model can help in reducing this structural uncertainty and sensitivity analyses can identify where linkages in models need to be inserted or removed.

10) Model evaluation and testing

As noted by Robson et al. (2008), there is some overlap with this step and step 8. The aim of this step is to evaluate the model with respect to its objectives. Testing model performance against independent data is often not possible, for example, for large integrated models. Therefore Jakeman et al. (2006) stress that model evaluation should go beyond the traditional attitude of 'validation' based only on model accuracy, to also include subjective criteria such as fitness for purpose and transparency of the modelling process. Model evaluation may include: sensitivity of model to plausible parameter changes; critique of assumptions; documentation; critique of the model development process; and ability to perform under a range of conditions including unexpected scenarios (Jakeman et al. 2006).

There must be a thorough analysis of how well the model achieves the modelling purpose and objectives specified in the first two steps. It is important that the entire modelling process is well documented; this includes specifying the rationale for modelling approach, the definition and rationale of all nodes and states in the network, citations of information and data sources and reporting the limitations and capabilities of the model. Every stage of the modelling process should be open to critical review and revision.

The development of a BN is often seen as an ongoing process. One of the major advantages of BN models is their ability to be easily updated with new information. This can be especially valuable in the context of environmental systems, where data and knowledge on processes is often limited. It provides modellers the opportunity to proceed with building a BN even with poor or incomplete knowledge, which can be updated when new data or improved system knowledge becomes available. The models can be updated using case files by applying data learning algorithms. BN modelling shells such as Netica, allow users to specify the weight of single or sets of cases (Marcot et al. 2006). BN models are therefore suitable for supporting the iterative process of learning and updating that serves adaptive management. The easily updatable nature of BNs also gives it a longer life span than most other models.

Conclusion

In order to create meaningful and credible BNs, the model objective and scope must first be clearly defined, followed by a compilation of knowledge about the modelled system and articulation of this knowledge into a conceptual model. This conceptual model can help form the structure of the BN, which should be parsimonious. The number of layers between input and final nodes should be kept to as few as possible, and as even as possible. The states for each node should be meaningful and also kept to as few as possible, especially if the CPTs are populated through expert elicitation. All assumptions used to build the BN must be documented, including uncertainties, descriptions and reasoning for each node and linkage, and information and data sources. This provides transparency in the modelling process, and enables model users to fully understand the basis of the model and its assumptions. It also makes the process of evaluating the model, as well as reconstructing, reproducing or altering it, much easier (Refsgaard and Henriksen 2004). Furthermore, it is important for the modeller to be able to defend the model, and minor details such as information sources can easily be lost or forgotten over time. The sensitivity analysis and testing steps must be properly documented, to demonstrate that the model has been rigorously evaluated. Following these guidelines will help ensure that the modelling process and the final model itself is transparent, credible and robust, within its given limitations (i.e. knowledge, data limits).

References

Borsuk, M. E., P. Reichert, A. Peter, E. Schager, and P. Burkhardt-Holm. 2006. Assessing the decline of brown trout (Salmo trutta) in Swiss rivers using a Bayesian probability network. Ecological Modelling **192**:224-244.

Borusk, M. E., C. A. Stow, and K. H. Reckhow. 2004. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecological Modelling **173**:219-239.

Bromley, J., N. A. Jackson, O. J. Clymer, A. M. Giacomello, and F. V. Jensen. 2004. The use of Hugin® to develop Bayesian networks as an aid to integrated water resource planning. Environmental Modelling and Software **20**:231-242.

Brugnach, M., C. Pahl-Wostl, K. E. Lindenschmidt, J. A. E. B. Janssen, T. Filatova, A. Mouton, G. Holtz, P. van der Keur, and N. Gaber. 2008. Complexity and uncertainty: rethinking the modelling activity. Pages 49-68 *in* A. J. Jakeman, A. A. Voinov, A. E. Rizzoli, and S. H. Chen, editors. Environmental Modelling, Software and Decision Support. Elsevier, Amsterdam.

Burgman, M., F. Fidler, M. McBride, T. Walshe, and B. Wintle. 2006. Eliciting expert judgements: literature review. Australian Centre of Excellence for Risk Analysis. University of Melbourne. ACERA Project 0611.

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, Wallingford, UK.

Cansado, A., and A. Soto. 2008. Unsupervised anomaly detection in large databases using Bayesian networks. Applied Artificial Intelligence **22**:309-330.

Castelletti, A., and R. Soncini-Sessa. 2007. Bayesian Networks and participatory modelling in water resouce management. Environmental Modelling and Software **22**:1075-1088.

Cheng, J., and R. Greiner. 2001. Learning Bayesian belief network classifiers: algorithms and system. Pages 141-151 *in* E. Stroulia and S. Matwin, editors. Advances in Artificial Intelligence. Lecture Notes in Computer Science 2056. Springer-Verlag, Berlin Heidelberg.

Crout, N., T. Kokkonen, A. J. Jakeman, J. P. Norton, L. T. H. Newham, R. Anderson, H. Assaf, B. F. W. Croke, N. Gaber, J. Gibbons, D. Holzworth, J. Mysiak, J. Reichl, R. Seppelt, T. Wagener, and P. Whitfield. 2008. Good modelling practice. Pages 15-31 *in* A. J. Jakeman, A. A. Voinov, A. E. Rizzoli, and S. H. Chen, editors. Environmental Modelling, Software and Decision Support. Elsevier, Amsterdam.

Gibert, K., J. Spate, M. Sànchez-Marrè, I. N. Athanasiadis, and J. Comas. 2008. Data mining for environmental systems. Pages 205-228 *in* A. J. Jakeman, A. A. Voinov, A. E. Rizzoli, and S. H. Chen, editors. Environmental Modelling, Software and Decision Support. State of the Art and New Perspectives. Elsevier, Amsterdam.

Haag, D., and M. Kaupenjohann. 2001. Parameters, prediction, post-normal science and the precautionary principle - a roadmap for modelling for decision-making. Ecological Modelling **144**:45-60.

Henderson, C., C. A. Pollino, and B. T. Hart. 2009. Appendix A: The Review - Workshop discussion document. Pages 33-83 *in* B. T. Hart and C. A. Pollino, editors. Bayesian modelling for risk-baed environmental water allocation. Waterlines Report Series No. 14, Feb 2009, Canberra.

Jakeman, A. J., R. A. Letcher, and S. H. Chen. 2007. Integrated assessment of impacts of policy and water allocation changes across soical, economic and environmental dimensions. Pages 97-112 *in* K. Hussey and S. Dovers, editors. Managing Water for Australia: the Social and Institutional Challenges. CSIRO Publishing, Collingwood, Victoria.

Jakeman, A. J., R. A. Letcher, and J. P. Norton. 2006. Ten iterative steps in development and evaluation of environmental models. Environmental Modelling & Software **21**:602-614.

Kjærulff, U. 1995. dHugin: a computational system for dynamic time-sliced Bayesian networks. International Journal of Forecasting **11**:89-111.

Marcot, B. G., R. S. Holthausen, M. G. Raphael, M. M. Rowland, and M. J. Wisdom. 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. Forest Ecology and Management **153**:29-42.

Marcot, B. G., J. D. Steventon, G. D. Sutherland, and R. K. McCann. 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere **36**:3063-3074.

Mead, R., J. Paxton, and R. Sojda. 2006. Applications of Bayesian Networks in ecological modelling. *in*

Murphy, K. P. 2002. Dynamic Bayesian Networks. PhD Dissertation. University of California, Berkeley.

Norsys. 2003. Test net using cases. *in*

Park, Y. S., P. F. M. Verdonschot, and S. Lek. 2005. Review of modelling techniques. Pages 21-40 *in* S. Lek, M. Scardi, P. F. M. Verdonschot, J. P. Descy, and Y. S. Park, editors. Modelling Community Structure in Freshwater Ecosystems. Springer-Verlag Berlin Heidelberg, Berlin.

Pollino, C. A., B. T. Hart, and B. R. Bolton. 2008. Modelling ecological risks from mining activities in a tropical system. Australasian Journal of Ecotoxicology **14**:119-128.

Pollino, C. A., A. K. White, and B. T. Hart. 2007a. Examination of conflicts and improved strategies for the management of an endangered Eucalypt species using Bayesian networks. Ecological Modelling **201**:37-59.

Pollino, C. A., O. Woodberry, A. Nicholson, K. Korb, and B. T. Hart. 2007b. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. Environmental Modelling & Software **22**:1140-1152.

Refsgaard, J. C., and H. J. Henriksen. 2004. Modelling guidelines - terminology and guiding principles. Advances in Water Resources **27**:71-82.

Robson, B. J., D. P. Hamilton, I. T. Webster, and T. Chan. 2008. Ten steps applied to development and evaluation of process-based biogeochemical models of estuaries. Environmental Modelling and Software **23**:369-384.

Sadoddin, A., R. A. Letcher, A. J. Jakeman, and L. T. H. Newham. 2005. A Bayesian decision network approach for assessing the ecological impacts of salinity management. Mathematics and Computers in Simulation **69**:162-176.

Smith, C. S., A. L. Howes, B. Price, and C. A. McAlphine. 2007. Using a Bayesian belief network to predict sutiable habitat of an endangered mammal - the Julia Creek Dunnart (Sminthopsis douglasi). Biological Conservation **139**:333-347.

Ticehurst, J. L., L. T. H. Newham, D. Rissik, R. A. Letcher, and A. J. Jakeman. 2007. A Bayesian network approach to assess the sustainability of coastal lakes. Environmental Modelling and Software **22**:1129-1139.

Tighe, M., C. A. Pollino, and S. M. Cuddy. 2007. A Bayesian approach to assessing regional climate change pressures on natural resource conditions in the Central West of NSW, Australia. Pages 233-239 *in* L. Oxley and D. Kulasiri| Secondary Title|. Publisher|, Place Published|.

Uusitalo, L. 2007. Advantages and challenges of Bayesian networks in environmental modelling. Ecological Modelling **203**:312-318.