ECOLOGICAL MODELLING 203 (2007) 312–318

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel

Advantages and challenges of Bayesian networks in environmental modelling

*Laura Uusitalo**

Fisheries and Environmental Management Group (FEM), Department of Biological and Environmental Sciences, P.O. Box 65, 00014 University of Helsinki, Finland

ARTICLE INFO

Article history:
Received 24 February 2006
Received in revised form
28 November 2006
Accepted 29 November 2006
Published on line 8 January 2007

Keywords:
Bayesian networks
Bayes nets
Probabilistic networks
Influence diagrams
Modelling

ABSTRACT

Bayesian networks (BNs) are an increasingly popular method of modelling uncertain and complex domains such as ecosystems and environmental management. At best, they provide a robust and mathematically coherent framework for the analysis of this kind of problems. However, there are certain pitfalls as well. In this paper, I summarise the pros and cons of the use of Bayesian networks especially in the context of environmental modelling and management. I will also give references to relevant publications, and introduce some software products that can be used to build Bayesian networks.

1. Introduction

Bayesian networks (BNs), also called belief networks, Bayesian belief networks, Bayes nets, and sometimes also causal probabilistic networks, are an increasingly popular methods for modelling uncertain and complex domains such as ecosystems and environmental management. They emerge from artificial intelligence research and have been applied to a wide range of problems, ranging from text analysis (Dong and Agogino, 1997) to problems in medical diagnoses (Kahn et al., 1997) and the evaluation of scientific evidence (Garbolino and Taroni, 2002). They are also increasingly used in environmental modelling and management (e.g. Varis et al., 1990; Lee and Rieman, 1997; Varis, 1997; Reckhow, 1999; Marcot et al., 2001; Borsuk et al., 2004; Little et al., 2004; Wooldridge and Done, 2004; Bromley et al., 2005; Uusitalo et al., 2005).

Bayesian modelling techniques have several features that make them useful in many real-life data analysis and management questions. They provide a natural way to handle missing data, they allow combination of data with domain knowledge, they facilitate learning about causal relationships between variables, they provide a method for avoiding overfitting of data (Heckerman, 1995), they can show good prediction accuracy even with rather small sample sizes (Kontkanen et al., 1997a), and they can be easily combined with decision analytic tools to aid management (Kuikka et al., 1999; Marcot et al., 2001; Jensen, 2001, Ch. 4). On the other hand, their ability to deal with continuous data is limited (Jensen, 2001, p. 69), and such data generally needs to be discretized, which may cause certain difficulties. Bayesian networks are also a useful tool for expert elicitation and combining uncertain knowledge when used with care. Furthermore, building models forces us

* Tel.: +358 9 19158992; fax: +358 9 19158257.
E-mail address: laura.uusitalo@iki.fi.

doi:10.1016/j.ecolmodel.2006.11.033

to think clearly about the subject, and articulate that thinking in the form of the model. This is often beneficial in and of itself (Marcot et al., 2001; Walters and Martell, 2004, p. 3).

Bayesian networks represent one branch of Bayesian modelling, the other major approach being hierarchical simulation-based modelling (Gilks et al., 1994; Gelman et al., 1995). In simulation-based modelling, the often analytically intractable probability distributions are estimated by generating samples from these distributions by simulation (Gelman et al., 1995), whereas in Bayesian networks, the probability distributions are generally expressed in discrete form and solved analytically. Both of these approaches share the idea of conditional dependence between variables and the updating of knowledge based on Bayes's theorem. Despite their similarity in aims and ideas, the practical modelling work is quite different, however, and the ideal method depends on the modelling needs in each case. Hierarchical modelling is especially suitable for cases with relatively abundant knowledge of complicated interactions between the model variables especially if this knowledge can be expressed with parametric distributions, and time-sliced models, while Bayesian networks are at their best with discrete domains and when reviewing and comparing different management choices or other courses of action. For many applications, either approach is appropriate.

In this paper I give an overview of the advantages and weak points of Bayesian networks, especially in relation to environmental research, and try to summarise the practical issues that often arise when applying BNs to the field. I review the current use of BNs in environmental research, and give some pointers to those who wish to apply BNs but do not know where to start. All along the way, I give references to books and articles that might prove useful in getting to know BNs.

2. Bayesian networks and their advantages

Bayesian networks are mathematical models presented graphically so that each variable is presented as a node with the directed links forming arcs between them. The information content of each variable is represented as one or several probability distributions. If a variable has no incoming arcs and is hence not dependent on any other variables in the model universe (i.e. has no *parents*), it has one probability distribution, and if it has parents, it has one probability distribution per each combination of possible values of the parents. Bayesian networks use probability as a measure of uncertainty: Beliefs about values of variables are expressed as probability distributions, and the higher the uncertainty, the wider is the probability distribution. As information accumulates, knowledge of the true value of the variable usually increases, i.e. the uncertainty of the value diminishes and the probability distribution grows narrower (Gelman et al., 1995; Sivia, 1996).

The probabilistic presentation of the interactions is one of the key points of BNs (Reckhow, 1999), and it allows for the estimation of risks and uncertainties better than models that only account for expected values. The probabilistic presentation of knowledge also prevents overconfidence in the strength of responses obtained by manipulating certain parts of the ecosystem. This is an important improvement to deterministic models which may work well in theoretical examinations but remain fraught with uncertainty when applied to problems with real data (Reckhow, 1999; Wikle, 2003).

The usefulness of BNs lies in the fact that by using Bayes's theorem (after Reverend Thomas Bayes, 1702–1762), one can calculate not only the probability distributions of children given the values of their parents, but also the distributions of the parents given the values of their children. That is, one can proceed not only from causes to consequences, but also deduce the probabilities of different causes given the consequences. Ellison (2004) gave a comprehensive account on the differences of Bayesian and frequentist (i.e. classical statistical) philosophies.

Bayesian networks are used for the analysis of data and expert knowledge especially in fields that are fraught with uncertainty, since they make it possible to treat uncertainty explicitly. They are also used to create "expert systems" that model and include expert knowledge about a complicated domain such as medicine and medical research (Kahn et al., 1997; Papaconstantinou et al., 1998; Lucas et al., 2000). Bayesian networks can also be supplemented with decision support tools (Kuikka et al., 1999; Jensen, 2001), which are a natural addition to the ability to treat uncertainty in the first place.

Charniak (1991) gave an introduction to Bayesian networks. Van der Gaag (1996) gave a review of the historical development of BNs as well as an introduction to their formalism and use. Jensen (2001) introduced BNs and their decision support extensions. Constant development occurs in the BN software products (reviewed in Ch. 5) as well as in their use in several fields, but the theoretical development was strongest in the 1980s and 1990s.

2.1. *Suitable for small and incomplete data sets*

A very useful aspect of BNs is that technically, there is no such thing as "too little data". There are no minimum sample sizes required to perform the analysis, and BNs take into account all the data there is (Myllymäki et al., 2002). Furthermore, Kontkanen et al. (1997a) demonstrate that Bayesian networks can show good prediction accuracy even with rather small sample sizes.

The conditional probabilities of the model can be estimated from data using an Expectation-Maximization (EM) algorithm (Spiegelhalter et al., 1993; Laurizen, 1995). It requires only the model structure to be known beforehand, and iteratively calculates maximum likelihood estimates for the parameters given the data and the model structure.

Environmental data often include missing values, since problems in sampling may mean that some unique event or point in time is missed. Unlike many estimation methods, EM algorithms can handle situations with missing observations, whether the data is missing randomly or the absence of an observation is dependent on the states of other variables (Heckerman, 1995). The distributions for the incomplete data can be approximated using Dirichlet distributions.

2.2. *Structural learning possible*

In addition to defining the model structure based on subject matter knowledge and using the data to define the conditional probability distributions, it is also possible to use data to learn also the structure of BN. This is an area of active research, and although the statistical theory is well understood, the methods are still under development, since their computational requirements are hard (Jensen, 2001, p. 81; Myllymäki et al., 2002). Finding the optimal model structure is computationally a very hard procedure (Myllymäki et al., 2002), and approximation methods are generally used instead.

There are two main approaches to structural learning (Steck and Tresp, 1999): Bayesian and constraint-based. In the Bayesian approach (Heckerman et al., 1995), the user first constructs a BN with which she encodes her knowledge of the subject and her confidence in this network. This prior network is then combined with data to find the most likely model structure. This can be computationally very demanding (Steck and Tresp, 1999). The constraint-based algorithms search for conditional dependences between each pair of variables, and build the model structure based on them (Steck and Tresp, 1999). They are computationally easier, and therefore more common. Constraint-based learning requires no prior knowledge or input from the user.

Structural learning, especially in the constraint-based form, has only limited use in environmental research, however. Often there are knowledge, theories, and hypotheses on the causal interactions between the variables, and it is then only reasonable to make use of them by building the model structures accordingly. Experience and simulations have shown that environmental interaction, which often includes a lot of variation and uncertainty, cannot be reliably estimated based on the available data sets. Theories about causal connections generally result in better models.

2.3. *Combining different sources of knowledge*

An important feature of Bayesian methods is the use of prior information. Priors reflect our knowledge of the subject before the research is conducted, and can be either highly informative and detailed, in case there is a lot of knowledge about the subject already, or very uninformative, if not much is known. These priors are then updated with data, to obtain a synthesis of old knowledge and new data. This synthesis can then be used as a prior in a new study. This mechanism makes the scientific learning process explicit, and also makes the assumptions made by the scientists transparent and open to discussion.

Bayesian network models also have the advantage that they can easily and in a mathematically coherent manner incorporate knowledge of different accuracies and from different sources. Expert knowledge can be combined with data (Marcot et al., 2001) regarding variables on which no data exist. Data measured on different levels of accuracy (e.g. absence/presence and quantity data) can be also combined.

If desired, BNs can also be used in conjunction with other Bayesian analysis methods such as Markov chain Monte Carlo (MCMC). MCMC results can be imported into BN, for instance in order to create a meta-model incorporating different scenarios in an uncertain framework (Kuikka et al., 1999).

2.4. *Explicit treatment of uncertainty and support for decision analysis*

Bayesian networks can easily be supplemented with variables encoding managerial decisions that in their turn affect the natural variables of the model, and with variables encoding costs and utilities related to these decisions and their outcomes (Jensen, 2001). These models naturally focus on the relationship between actions, knowledge and uncertainty; the consequences of various management decisions can be studied not only from the perspective of expected values, but also with regard to the risks of highly undesirable outcomes (Kuikka et al., 1999). This aspect of BNs has been found very useful, and a large fraction of the environmental applications of BNs focuses on the management of natural resources, and the relationship between nature, society and economics (see Ch. 4).

2.5. *Fast responses*

Because BNs are solved analytically, they can provide fast responses to queries once the model is compiled. The compiled form of a BN contains a conditional probability distribution for every combination of variable values, and can thus provide any distribution instantly, in contrast to simulation models in which the results need to be simulated, which can take very long. Due to the fastness is responses, BNs can be useful in teaching the principles of Bayesian modelling as well as in communicating e.g. the results of simulation-based models.

3. Challenges in the use of Bayesian networks

3.1. *Discretization of continuous variables*

In environmental research as well as in many other fields, data and parameters often have continuous values. Bayesian networks can, however, deal with continuous variables in only a limited manner (Friedman and Goldszmidt, 1996; Jensen, 2001, p. 69). The usual solution is to discretize the variables and build the model over the discrete domain. There is a trade-off, however, as the discretization can only capture rough characteristics of the original distribution (Friedman and Goldszmidt, 1996), and we may loose statistical power if the relationship between the variables is, in fact, linear (Myllymäki et al., 2002). On the other hand, we gain the ability to use the reasoning machinery of BNs, which is especially efficient if the relationships between the variables are non-linear and complex (Myllymäki et al., 2002).

How to discretize the data is more difficult a question. Automatic data discretization techniques have been developed and discussed (Dougherty et al., 1995; Friedman and Goldszmidt, 1996; Kontkanen et al., 1997b; Kozlov and Koller, 1997; Monti and Cooper, 1998; Kurgan and Cios, 2001; Hammond, 2004), but no satisfactory automatic discretization methods for Bayesian networks have been found (Myllymäki et al., 2002; Prof. P. Myllymäki, pers. comm.). Thus, finding a discretization that can be reasonably interpreted in terms of the study problem, generally given by domain experts, remains the best solution.

Finding the discretization is a task that deserves attention, since the way the data is discretized – the number of intervals and the division points – can make a notable difference in the resulting model (Myllymäki et al., 2002). Generally, we are likely to find more dependencies when discretizing the data to only few intervals than when using many intervals (Myllymäki et al., 2002). The discretization should thus consider the number of intervals, the ecological/domain significance of the breakpoints, and preferably try to guarantee that each of the intervals has a reasonable amount of observations. This is a task that may require time and examination on the part of the expert team working on it.

Using discretized values that are treated as categorical allows us to capture non-linear relationships between the variables (Myllymäki et al., 2002), as well as complex distributions such as bi- or multimodal distributions that may be difficult to capture into parametric distributions. We may need a large number of bins to be able to present these complex distributions, however. The bigger the number of intervals, the more data is needed for it to be likely to find interesting dependencies (Myllymäki et al., 2002). In practice, even large ecological data sets are rarely large enough to allow a high number of intervals per variable; ecological studies that report their models in such detail have included 2–10 intervals (Lee and Rieman, 1997; Marcot et al., 2001; Wooldridge and Done, 2004). The problem of the sufficiency of data is multiplied if the model structure is complicated, i.e. many of the variables have several parents, since the number of conditional distributions in the child is the product of the number of intervals in the parents. These distributions become weakly defined if the data has to be divided into a large amount of conditional distributions and there are only a few data points per distribution. In practice, this means that in order to build meaningful BNs, we will often have to restrict the number of intervals, which diminishes the benefits of theoretically being able to capture complex empirical distributions. So, while there are no theoretical minimum limits for the amount of data in the context of BNs, in practical applications the amount of data may well be the limiting factor in the modelling.

3.2. *Collecting and structuring expert knowledge*

While Bayesian models are a useful way to model expert knowledge, it may prove difficult to get the knowledge out of the experts in a form that can be converted into probability distributions. There are two main reasons for this. Firstly, many ecology researchers are used to working with real sampling or experimental data, and may find it exceedingly difficult to provide any numbers without relying on data. Secondly, they may be used to classical statistical analyses and feel uncertain when trying to think about their knowledge in terms of distributions rather than point estimates and confidence intervals. This uncertainty together with only superficial knowledge about the methodology may also lead to distrust towards the BNs, which easily leads to reluctance to provide the estimates.

The task of estimating probabilities, especially those of rare events, is a difficult one, and people naturally rely on a set of heuristic procedures that often do serve them well but may also result in biased outcomes (Morgan and Henrion, 1990, p. 102). Studies of estimation processes have also revealed that regardless of elicitation technique, human estimators are prone to *overconfidence*, that is, giving estimates that are too near to zero or one (Morgan and Henrion, 1990, pp. 112–116). On the other hand, experts' judgments tend to be rather under- than overconfident (Morgan and Henrion, 1990, pp. 130–131). Morgan and Henrion (1990) give a good account on these procedures and the psychology of probability assessment.

How to organise the expert elicitation then? The first thing to be considered is whether the experts are asked to provide both the model structure and the probability distributions or whether the model structure will be pre-mediated and the expert knowledge is just needed for the probability distributions. In some study problems, the model structure that reflects which variables affect others can be defined with very little uncertainty based on, e.g. established theory or natural laws. In some other problems, it may be far from clear which variables are the causes and which the consequences, and even if certain variables are relevant at all. This can be the situation in the study of complex phenomena such as the global change. In such cases, it is advisable to ask the experts to create first the model structure according to their own beliefs, and then give estimates of the relevant probability distributions related to the model. This type of approach was used by Kuikka and Varis (1997), whereas Pellikka et al. (2005) constructed and analysed models separately for each expert. Not much literature exist on this topic, however. Various methods have been developed for the elicitation process with multiple experts, and Morgan and Henrion (1990) give an extensive overview of them. Clemen and Winkler (1999) discuss and review combining distributions from experts.

If the model structure is pre-mediated, the role of experts is to provide estimates of the parameters, i.e. the (conditional) probability distributions. In order to be able to do this, the experts must be familiar with the variables and agree with the model structure; if the model does not make sense to them, they cannot be expected to produce sensible probability distributions either. Furthermore, mathematical methods used to combine expert judgements assume that the experts agree on the definitions and on what is to be assessed (Clemen and Winkler, 1999).

It has to be remembered in designing the model structure that people generally have cognitive difficulty in thinking of conditional distributions with several conditioning factors (Morgan and Henrion, 1990, p. 163). Hence, it is advisable to restrict the amount of conditioning factors to a very low number by restructuring the model or by omitting less important variables, if possible. Bayesian networks are a good tool for expert elicitation in the sense that breaking the problem down to lower-dimension sub-problems is natural in Bayesian networks, and tends to provide more accurate estimates than direct assessments of probabilities (Armstrong et al., 1975). In addition to restructuring, the use of parametric distributions can be considered (Druzdzel and van der Gaag, 2000). The aim is to reach balance between rich, detailed models, and the costs associated with building and using such a model (Druzdzel and van der Gaag, 2000).

Keith (1996) notes that combining divergent expert judgements is problematic because the fraction of experts with a certain opinion is not likely to be proportional to the probability of that opinion being true. While this is true, it may still be that this is the best available estimate of the probability, and while the resulting uncertainty caused by this has to be taken into account, it is hardly worthwhile to abandon the whole estimation process. The question of combining expert judgments is discussed widely in literature (e.g. Morgan and Henrion, 1990; Titus and Narayanan, 1996).

3.3. *No support for feedback loops*

Bayesian networks are acyclic, and thus do not support feedback loops (Jensen, 2001, p. 19) that would sometimes be beneficial in environmental modelling. Temporal or spatial dynamics can be modelled in BNs using a separate network for each time slice (e.g. Jensen, 2001, Ch. 2); however, this is often very tedious.

4. Environmental applications of Bayesian networks

Data analysis performed with BNs is quite rare in the field of environmental sciences. Varis and Kuikka (1997) built a BN into which they embedded computational and regression models describing Baltic salmon stock dynamics. Wooldridge and Done (2004) predicted coral bleaching in the Great Barrier Reef using a BN based on various data.

Models summarising and incorporating simulation models are more common. Lee and Rieman (1997) created a Bayesian network model for the assessment of fish population viability. They ran Monte Carlo simulations over all possible combinations of the parameters, and used these probabilities in the BN. Kuikka et al. (1999) combined a Bayesian network with MCMC simulations to look for the most information-robust management methods for Baltic cod; that is, they tried to find a management strategy that works despite having insufficient knowledge about the stock dynamics. This model was used to predict ecosystem responses under different management scenarios. Little et al. (2004) created a hypothetical, simulated fishery based on a real fishery on the Great Barrier Reef, and used it to examine the effect of information flow among fishing vessels. Borsuk et al. (2004) developed a Bayesian network representing the eutrophication process by merging several previously published sub-models.

Environmental BN applications often include decision analysis and encoded expert knowledge. Varis et al. (1990) created a Bayesian network with decision and utility variables (an influence diagram) to evaluate management strategies of a eutrophic lake, especially regarding the choice between gathering more data and acting. The quantification of expert knowledge using BNs also has applications in the field of environmental sciences. Kuikka and Varis (1997) interviewed experts to estimate the effect of climatic change on watersheds, and Marcot et al. (2001) combined some data with their mostly expert-opinion-based analysis of population viability. Pellikka et al. (2005) used the approach to examine the effect of local game management on the uncertainty of the future population sizes, and Uusitalo et al. (2005) used it to obtain an estimate on the maximum salmon smolt production capacity in the Baltic Sea, which is an important management variable on which no actual data exist.

5. State of the art: how to apply Bayesian networks?

It is advisable to get familiar with the basics of the reasoning behind BNs before starting to build a model. Very good introductions to the topic are written by, e.g. Charniak (1991), Heckerman (1995), Heckerman and Wellman (1995) and Jensen (2001). They also provide a look at some modelling technique issues.

There are a number of software packages for building BNs on desktop computers, and many of those have been ported to several common operating systems. These packages are shortly introduced here so that readers can assess which of them would be most suitable for their purposes. This list is not comprehensive, however, and the reader is advised to search the Internet for further packages. Not all the features of the software packages are presented here; the aim is to highlight the most important aspects of each software package.

Hugin (http://www.hugin.com/, Madsen et al., 2005) is a family of software products designed for building Bayesian networks on Windows, Linux, Solaris and Mac operating systems and a free, limited demo version on is provided on their web site. The demo version is capable of handling up to 50 states (that is, e.g. 5 variables with 10 states in each or 25 variables with 2 states in each) and learning from up to 500 cases. There are no restrictions in saving the files etc.

The Hugin engine lets the used choose between designing the model structure manually and learning the model structure from data using constrain-based PC or NPC algorithms; also supervised data-based learning is possible. The conditional probability tables can then be learned from data using EM algorithm of Laurizen (1995). All data given to Hugin must be previously discretized, or discretized manually in the GUI dialogue, as Hugin is not capable of sorting continuous-type data into bins. Hugin can handle missing data, though data missing massively and systematically causes some problems (Laurizen, 1995).

Hugin supports pure BNs as well as influence diagrams, i.e. BNs with decision and utility nodes. In the latter case the software computes expected utilities related to each decision option given the rest of the network.

Hugin has extensive tutorials. Bromley et al. (2005) present the use of Hugin in developing an integrated water resource planning tool.

Netica (http://www.norsys.com/) is another widely used Bayesian network software package that has Windows and Mac OS versions. A free, limited demo version is available. The demo version imposes no limitations in saving or the use period, but only networks with up to 15 variables can be saved.

Netica provides EM algorithm for learning the conditional probability tables from a data set (Spiegelhalter et al., 1993), but does not perform structural learning; the user has to define the model structure. Netica can, however, allocate continuous-type of data into correct bins once the bins are defined in the network structure and does not require for the input data file to be discretized.

Netica supports the use of decision and utility variables; Marcot et al. (2001) use Netica and its influence diagram features in their work. Netica offers single-finding sensitivity analysis, which determines how much our perception of values in other variables will change given different values of a certain variable. This is done to only one variable at a time. However, any information currently entered into the network is taken into account in the sensitivity analysis, which provides a (slightly tedious) way to perform sensitivity analysis over several observed variables. Wooldridge and Done (2004) apply the sensitivity analysis in their work.

SamIam (http://reasoning.cs.ucla.edu/samiam/) has been developed at the University of California, and can be downloaded for free from their website. There are versions of the software for Windows, Linux, Mac and Solaris operating systems.

SamIam reads and writes Hugin, Netica, Ergo, GeNIe, and Microsoft BN toolkit network file formats and Hugin-type data files. It provides EM learning but not structural learning from data, so the user has to define the model structure. The algorithms SamIam uses are exceptionally well documented in the help files.

SamIam provides sensitivity analysis different to that one implemented in Netica (Chan and Darwiche, 2002). It allows the user to set constraints to certain parameters and identifies the minimal changes that are required in the network to satisfy those constraints. This kind of sensitivity analysis is very useful in building systems that are intended to mimic experts' reasoning and help people in diagnosing and decision-making.

B-Course (http://b-course.hiit.fi/, Myllymäki et al., 2002) is a web-based tool for Bayesian network modelling, developed by the Helsinki Institute for Information Technology. It supports most web-browsers, and may be used freely for research and educational purposes.

B-Course provides a structural learning engine, and in fact cannot be used by building the model structure manually. B-course can deal with missing data, and it automatically discretizes continuous data. The user cannot affect how the data is discretized, however, and often the better solution is to discretize the data separately before submitting it. B-course has a tutorial-type interface, and it might serve as a good starting point for studying dependence modelling.

Bayesian network toolkit by Microsoft (http://research.microsoft.com/adapt/MSBNx/) runs on Windows OS. It may be freely used for non-commercial purposes. It does not support structural learning or learning the conditional probabilities from data, but model structures and conditional probabilities must be defined manually. Bayesian network toolkit does not support influence diagrams, i.e. the inclusion of decision and utility variables. It does provide value-of-information analysis that helps define whether knowing a state of a certain variable would improve the expected value of a decision.

Ergo (http://www.noeticsystems.com/ergo/) runs on Windows and Mac, and free trial version that does not allow saving can be downloaded from their website. Ergo does not support any learning from data, so any models need to be defined manually. It has some built-in features that make it easy to construct easy-to-use interfaces for diagnostic purposes. Ergo does not support influence diagrams.

GeNIe (http://genie.sis.pitt.edu/) has been developed in the University of Pittsburgh and it runs on Windows. It is provided free of charge for any use including commercial use. GeNIe does not provide learning model structures or conditional probabilities from data, but it does allow the user to construct influence diagrams. It also offers value-of-information analysis.

6. Conclusions

Bayesian networks can be a useful addition to the toolkit of environmental scientists, especially if their work is related to environmental management. Explicit accounting for uncertainty can add substantial insight to many real-life problems, and the graphical representation of model structures and probability distributions is very useful in communicating theories and results to colleagues, students, and decision-makers. The readily available BN development packages are relatively advanced and easy to use, but it has to be remembered that the development of BN methodology is still ongoing, and not all algorithms that have been proposed are implemented in these programs. Bayesian networks are not a household name in the environmental research field yet. They are gaining popularity in the field, however, and are likely to establish their position as one of the standard methods of analysis especially in problems dominated by uncertainty. Therefore, getting to know them at least on a general level will be beneficial to all environmental scientists.

Acknowledgements

This work was funded by the Finnish Biological Interactions Graduate School and Jenny and Antti Wihuri Foundation. I wish to thank Prof. Sakari Kuikka for introducing me to Bayesian networks and Sanna Koulu, Dr. Samu Mäntyniemi, Dr. Jani Pellikka, and an anonymous reviewer for constructive comments on the manuscript.

REFERENCES

Armstrong, J.S., Denniston Jr., W.B., Gordon, M.M., 1975. The use of the decomposition principle in making judgments. Org. Behav. Hum. Perform. 14, 257–263.

Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2004. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecol. Model. 173, 219–239.

Bromley, J., Jackson, N.A., Clymer, O.J., Giacomello, A.M., Jensen, F.V., 2005. The use of Hugin® to develop Bayesian networks as an aid to integrated water resource planning. Environ. Model. Software 20, 231–242.

Chan, H., Darwiche, A., 2002. When do numbers really matter? J. Artif. Intell. Res. 17, 265–287.

Charniak, E., 1991. Bayesian networks without tears. AI Mag. 12 (4), 50–63.

Clemen, R.T., Winkler, R.L., 1999. Combining probability distributions from experts in risk analysis. Risk Anal. 19 (2), 187–203.

Dong, A., Agogino, A.M., 1997. Text analysis for constructing design representations. Artif. Intell. Eng. 11, 65–75.
Dougherty, J., Kohavi, R., Sahami, M., 1995. Supervised and unsupervised discretization of continuous features. In: Proceedings of the 12th International Conference on Machine Learning (ICML). Morgan Kaufmann Publishers, San Francisco, CA.
Druzdzel, M.J., van der Gaag, L.C., 2000. Building probabilistic networks: where do the numbers come from?—a guide to the literature. Trans. Knowledge Data Eng. 12 (4), 481–486.
Ellison, A.M., 2004. Bayesian inference in ecology. Ecol. Lett. 7, 509–520.
Friedman, N., Goldszmidt, M., 1996. Discretizing continuous attributes while learning Bayesian networks. In: Proceedings of the 13th International Conference on Machine Learning (ICML). Morgan Kaufmann Publishers, San Francisco, CA, pp. 157–165.
Garbolino, P., Taroni, F., 2002. Evaluation of scientific evidence using Bayesian networks. Forensic Sci. Int. 125, 149–155.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian data Analysis. In: Texts in Statistical Science. Chapman & Hall, ISBN 0-412-03991-5.
Gilks, W.R., Thomas, A., Spiegelhalter, D.J., 1994. A language and program for complex Bayesian modelling. The Statistician 43 (1), 169–177.
Hammond, T.R., 2004. A recipe for Bayesian network driven stock assessment. Can. J. Fish. Aquat. Sci. 61, 1647–1657.
Heckerman, D., 1995. A tutorial on learning with Bayesian networks. Technical report MSR-TR-95-06, Microsoft Research.
Heckerman, D., Wellman, M.P., 1995. Bayesian networks. Commun. ACM 38 (3), 27–30.
Heckerman, D., Geiger, D., Chickering, D.M., 1995. Learning Bayesian networks: the combination of knowledge and statistical data. Machine Learning 20, 197–243.
Jensen, F.V., 2001. Bayesian Networks and Decision Graphs. Springer-Verlag, New York, ISBN 0-387-95259-4.
Kahn Jr., C.E., Roberts, L.M., Shaffer, K.A., Haddawy, P., 1997. Construction of a Bayesian network for mammographic diagnosis of breast cancer. Computers Biol. Med. 27 (1), 19–29.
Keith, D.W., 1996. When is it appropriate to combine expert judgments? Clim. Change 33, 139–144.
Kontkanen, P., Myllymäki, P., Silander, T., Tirri, H., 1997a. Comparing predictive inference methods for discrete domains. In: Proceedings of the sixth International Workshop on Artificial Intelligence and Statistics, Ft. Lauderdale, USA, pp. 311–318.
Kontkanen, P., Myllymäki, P., Silander, T., Tirri, H., 1997b. A Bayesian approach to discretization. In: Proceedings of the European Symposium on Intelligent Techniques, Bari, Italy, pp. 265–268.
Kozlov, A.V., Koller, D., 1997. Nonuniform dynamic discretization in hybrid networks. In: Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI-97), August 1–3, 1997, pp. 314–325.
Kuikka, S., Varis, O., 1997. Uncertainties of climatic change impacts in Finnish watersheds: a Bayesian network analysis of expert knowledge. Boreal Environ. Res. 2, 109–128.
Kuikka, S., Hildén, M., Gislason, H., Hansson, S., Sparholt, H., Varis, O., 1999. Modeling environmentally driven uncertainties in Baltic cod (*Gadus morhua*) management by Bayesian influence diagrams. Can. J. Fish. Aquat. Sci. 56, 629–641.
Kurgan, L., Cios, K.J., 2001. Discretization algorithm that uses class-attribute interdependence maximization. In: Proceedings of the 2001 International Conference on Artificial Intelligence (IC-AI 2001), Las Vegas, Nevada, pp. 980–987.
Laurizen, S.L., 1995. The EM algorithm for graphical association models with missing data. Computat. Stat. Data Anal. 19, 191–201.
Lee, D.C., Rieman, B.E., 1997. Population viability assessment of salmonids by using probabilistic networks. N. Am. J. Fish. Manage. 17, 1144–1157.
Little, L.R., Kuikka, S., Punt, A.E., Pantus, F., Davies, C.R., Mapstone, B.D., 2004. Information flow among fishing vessels modelled using a Bayesian network. Environ. Modell. Software 19, 27–34.
Lucas, P.J.F., de Bruijn, N.C., Schurink, K., Hoepelman, A., 2000. A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU. Artif. Intell. Med. 19, 251–279.
Madsen, A.L., Jensen, F., Kjærulff, U.B., Lang, M., 2005. The Hugin tool for probabilistic graphical models. Int. J. Artif. Intell. Tools 14 (3), 507–543.
Marcot, B.G., Holthausen, R.S., Raphael, M.G., Rowland, M., Wisdom, M., 2001. Using Bayesian Belief Networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. Forest Ecol. Manage. 153, 29–42.
Monti, S., Cooper, G.F., 1998. A latent variable model for multivariate discretization. Proceedings of 14th Conference of Uncertainty in AI, Madison, WI, pp. 404–413.
Morgan, M.G., Henrion, M., 1990. Uncertainty. In: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, ISBN 0-521-36542-2.
Myllymäki, P., Silander, T., Tirri, H., Uronen, P., 2002. B-Course: a web-based tool for Bayesian and causal data analysis. Int. J. Artif. Intell. Tools 11 (3), 369–387.
Papaconstantinou, C., Theocharous, G., Mahadevan, S., 1998. An expert system for assigning patients into clinical trials based on Bayesian networks. J. Med. Syst. 22 (3), 189–202.
Pellikka, J., Kuikka, S., Lindén, H., Varis, O., 2005. The role of game management in wildlife populations: uncertainty analysis of expert knowledge. Eur. J. Wildlife Res. 51, 48–59.
Reckhow, K.H., 1999. Water quality prediction and probability network models. Can. J. Fish. Aquat. Sci. 56, 1150–1158.
Sivia, D.S., 1996. Data Analysis. A Bayesian Tutorial. Oxford Science Publications, ISBN 0-19-851889-7.
Spiegelhalter, D.J., Dawid, A.P., Laurizen, S.L., Cowell, R.G., 1993. Bayesian analysis in expert systems. Stat. Sci. 8 (3), 219–247.
Steck, H., Tresp, V., 1999. Bayesian belief networks for data mining. In: Proceedings of The 2nd Workshop on Data Mining und Data Warehousing als Grundlage Moderner Entschidungsunterstuezender Systeme, DWDW99, Sammelband, Universität Magdeburg, September 1999.
Titus, J.G., Narayanan, V., 1996. The risk of sea level rise. Clim. Change 33, 151–212.
Uusitalo, L., Kuikka, S., Romakkaniemi, A., 2005. Estimation of Atlantic salmon smolt carrying capacity of rivers using expert knowledge. ICES J. Marine Sci. 62 (4), 708–722.
Van der Gaag, L.C., 1996. Bayesian belief networks: odds and ends. Computer J. 39, 97–113.
Varis, O., 1997. Bayesian decision analysis for environmental and resource management. Environ. Modell. Software 12 (2–3), 177–185.
Varis, O., Kettunen, J., Sirviö, H., 1990. Bayesian influence diagram approach to complex environmental management including observational design. Computat. Stat. Data Anal. 9, 77–91.
Varis, O., Kuikka, S., 1997. Joint use of multiple environmental assessment models by a Bayesian meta-model: the Baltic salmon case. Ecol. Model. 102, 341–351.
Walters, C.J., Martell, S.J.D., 2004. Fisheries Ecology and Management. Princeton University Press, ISBN 0-691-11545-1.
Wikle, C.K., 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology 84 (6), 1382–1394.
Wooldridge, S., Done, T., 2004. Learning to predict large-scale coral bleaching from past events: a Bayesian approach using remotely sensed data, in-situ data, and environmental proxies. Coral Reefs 23, 96–108.