Bayesian Networks in Environmental and Resource Management

David N Barton,*† Sakari Kuikka,‡ Olli Varis,∥ Laura Uusitalo,¶ Hans Jørgen Henriksen,# Mark Borsuk,** Africa de la Hera,†† Raziyeh Farmani,§§ Sandra Johnson,## and John DC Linnell||

*Norwegian Institute for Nature Research (NINA), Gaustadalleen 21, NO-0349 Oslo, Norway
†University of Helsinki, Fisheries and Environmental Management Group (FEM), Department of Environmental Sciences, University of Helsinki, Helsinki, Finland
‡Water and Development Research Group, Aalto University, Espoo, Finland
§Finnish Environment Institute (SYKE), Marine Research Centre, Helsinki, Finland
#Geological Survey of Denmark and Greenland (GEUS), Copenhagen, Denmark
||Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire, USA
§§Centre for Water Systems, University of Exeter, Exeter, United Kingdom
##Discipline of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia
|||Norwegian Institute for Nature Research (NINA Head Office), Trondheim, Norway

(Submitted 8 April 2012; Returned for Revision 14 May 2012; Accepted 14 May 2012)

ABSTRACT

This overview article for the special series, “Bayesian Networks in Environmental and Resource Management,” reviews 7 case study articles with the aim to compare Bayesian network (BN) applications to different environmental and resource management problems from around the world. The article discusses advances in the last decade in the use of BNs as applied to environmental and resource management. We highlight progress in computational methods, best-practices for model design and model communication. We review several research challenges to the use of BNs in environmental and resource management that we think may find a solution in the near future with further research attention. Integr Environ Assess Manag 2012;8:418–429. © 2012 SETAC

Keywords: Bayesian networks Risk analysis Ecological economic modeling DPSIR Decision analysis

INTRODUCTION

Environmental and resource managers and decision makers need models to help them understand the effectiveness of alternative management decisions. The closer one looks, though, the more apparent and daunting the challenges become. Models have to be adaptable because the context in which models are applied will never be the same twice. Even if an environment were somehow replicated, the needs of managers and decision makers would be different every time, because the priorities, objectives, knowledge and influence of various stakeholders, and the information available about the environment and resources are always unique. Therefore, models have to be flexible enough to incorporate a wide range of knowledge and operate on multiple scales and levels of resolution.

The ultimate purpose of modeling is to facilitate wise decisions. That means that models should promote social learning, i.e., learning that helps managers and decision makers rally diverse coalitions of stakeholders to support decisions that, without the benefit of models, might seem unacceptable.

In order to promote social learning, modelers have to be able to handle uncertainties well in their models. That means being able to express, differentiate, prioritize, and rigorously resolve uncertainties. It means being able to advise managers and decision makers on whether getting more information would be worth the costs. The challenges of uncertainty are among the greatest challenges facing modelers today. These are the challenges confronted in this special series of Integrated Environmental Assessment and Management (IEAM).

This journal has addressed how to integrate domain-specific knowledge in ecological risk assessment to support decision making (Ohlson and Serveiss 2007), as well as multicriteria analysis methods to analyze decision problems that seem intractable due to inherent trade-offs between sociopolitical, environmental, ecological, and economic factors (Kiker et al. 2005). Authors have called for models that allow environmental and resource managers to consider a broad range of social values in decision making (Costanza 2006). IEAM has also focused on how to structure models and decision-making processes in ecosystem management (Bruins et al. 2010), and how to promote social learning by requiring stakeholders to articulate and reconsider their respective values (Sparrevik et al. 2011; Stahl et al. 2011). Implications of different types of uncertainty for risk modeling and management have also been addressed in watershed management (Ragas et al. 2009).

Bayesian networks (BNs) are models that graphically and probabilistically represent relationships among variables. IEAM has only recently begun to publish articles using influence diagrams and BNs as tools for integrating multiple lines of evidence, including process-related information from existing data and expert judgment to aid ecological risk-based decision making (Carriger and Newman 2011). The articles in
Bayesian Networks in Environmental and Resource Management—Integr Environ Assess Manag 8, 2012 419

this special series aim to continue this initiative, by making BNs more widely known and hopefully contributing to the IEAM readership’s toolkit for environmental and resource management. We first provide a brief history of the use of BNs in environmental and resource management. Although incomplete it shows that the approach is not new, but its use has grown vastly in recent years in tandem with computing power. Also by way of introduction, we illustrate a simple BN example for analyzing an environmental management problem. The main body of this article discusses how the BN case studies in this special series illustrate operational, tactical, strategic, and directive decision-making problems. The last part of the article discusses further methodological challenges for BNs arising from the case studies and research questions for the future.

A BRIEF HISTORY OF BN

The BN approach has been successfully applied for many years in fields such as artificial intelligence, medicine, and epidemiology (Haas 1991a, 1991b; Jensen 1996; Jensen and Nielsen 2007; Olson et al. 1990; Pearl 1988). A recent literature review of BNs during the last decade found that 128 articles (approximately 4.2%) were published within the topics of environmental sciences (Aguilera et al. 2011), dominant fields of application being the computer sciences and mathematics. The authors identified approximately 70% of these environmental science articles as using BNs broadly for “inference” in environmental modeling, whereas the remainder used BNs for special cases of inference for “characterization,” “classification,” or “regression” of particular environmental variables. Aguilera et al. (2011) do not report on which studies related to environmental sciences used BN specifically for decision analysis. However, the majority of BNs used for inference modeling suggest that most applications in the literature were related to prediction, simulation, diagnosis, or learning for the purpose of environmental and resource management.

The first examples of BNs to environmental and resource management problems are more than 20 years old (Varis et al. 1990). In 1999, Ecological Modelling published what is widely cited as the first review of BN applications to environmental and resource management problems (Varis and Kuikka 1999).

In 2006, the Canadian Journal of Forest Research (McCann et al. 2006) ran a special issue with a number of case studies of BN applications to ecological modeling, including guidelines for BN modeling. During the last decade or so, BNs have been increasingly used in environmental modeling and integrated assessment, such as studies of surface and groundwater management (Reckhow 1999; Borsuk et al. 2003, 2004; Henriksen 2004; Bromley et al. 2005; Castelletti and Soncini-Sessa 2007; Ticheurst et al. 2007; Barton et al. 2008; Henriksen and Barlebo 2008), ecology and wildlife population viability (Marcot et al. 2001; Marcot 2006, 2007, 2008; Borsuk et al. 2006; Marcot, Hohenlohe et al. 2006), fisheries and ecosystem management (Varis and Kuikka 1997b; Kuikka et al. 1999; Hammond and O’Brien 2001; Hammond 2004; Uusitalo et al. 2005), and climate change (Varis and Kuikka 1997a). Lately, the BN methodology has been expanded to model the social dynamics of fishermen (Haapasaari et al. 2007; Haapasaari and Karjalainen 2010). Web-based model interfaces have recently made it possible to deploy Bayesian belief networks online for wider dissemination of the main model features (Madsen et al. 2012). Uusitalo et al. (this issue) provide an example of a BN distributed on a web-based platform (http://demo.hugin.com/index.php/Coastal_Fish_Production).

The term BN (or sometimes Bayesian belief network) is widely used to mean an acyclic-directed graph of probability distributions. “Acyclic directed” means that the model assumes a sequential flow of information among variables and has no dynamic feedback loops. In an acyclic directed graph, a variable may condition or be conditioned on another variable, but not both. Advances have recently been made in how to do dynamic modeling in BNs (Nicholson and Flores 2011). Johnson and Mergersen (this issue) also discuss how feedback loops similar to dynamic modeling can be implemented using BNs despite their acyclical properties.

This special series provides a number of examples of how BNs are used to link the knowledge represented in different kinds of models (e.g., expert systems, statistical models, and simulation models). A few different varieties of BN are presented in this issue, including models using a framework known as driver-pressure-state-impact-response (DPSIR). DPSIR has been used to frame models used to help solve many environmental and resource management problems (Borja et al. 2006; Agyemang et al. 2007; Atkins et al. 2011; Tscherning et al. 2012). A DPSIR can be set up as a causal probabilistic network. A causal probabilistic network is a relatively generic term used to describe networks whose nodes represent cause and effect variables and their states. The nodes in causal probabilistic networks are linked by conditional probability tables (CPTs). CPTs may be continuous probability density functions (PDFs) where uncertainty is simulated using a Monte Carlo approach. An example is given in Borsuk et al. (this issue). BNs may also be used to determine correlation structure among a set of variables without pre-establishing causal connections; for example see the article by Uusitalo et al. in this issue.

Discretized, rather than continuous, CPTs are the more common variety of BN used in the articles presented in this issue. All varieties of BN use Bayes’ theorem to “update” the probabilities in discretized networks based on new information that is fed into the network. Bayes’ theorem is easily derived from the axioms of probability theory and provides a rigorous way to use new information to update model uncertainties. Bayes’ theorem allows modelers to use new information to update probabilities associated with particular outcomes of interest to managers and decision makers (e.g., the probability of achieving a cleanup target concentration for each of 3 remedial alternatives) as new information becomes available. The next section of this article elaborates on how BNs may be used to help solve environmental and resource management problems, using a stylized example.

A STYLIZED BAYESIAN NETWORK FOR AN ENVIRONMENTAL MANAGEMENT EXAMPLE

We present a simple example of a BN for a lake eutrophication management problem in Figure 1. This problem is set in a DPSIR framework. The “driver” of change in this example is a decision about alternative nutrient abatement measures (D) with abatement costs, which abate the pressure of nutrient run-off from the catchment (P), which in turn affects the state of algal blooms in the lake (S), with impacts on bathing suitability (I) for lake users who derive recreational benefits. Costs and benefits of abatement measures are interpreted as direct and indirect responses (R) in this framing.
of the example. A manager doing an impact or scenario analysis may be interested in determining the posterior probability of a water quality state, given information on nutrient pressure and the states of the context variables such as precipitation. A policy-maker may be interested in the expected net benefits of uncertain nutrient abatement costs in light of uncertain recreational benefits. A number of modeling approaches, including BNs, allow for this kind of deductive ‘‘top-down’’ reasoning by using Monte Carlo simulation to link functions in a DPSIR causal chain to propagate the effects of management actions.

BNs also facilitate using Bayes’ theorem for inductive or ‘‘bottom-up’’ reasoning in the causal chain to determine the likelihood of different levels of nutrient pressure in a lake given knowledge about the states of the lake, the decision and context variables. Inductive reasoning should be natural to environmental managers with diagnostic questions such as, ‘‘given that I know environmental quality to be below a threshold with X% of confidence, what is the likelihood of a pollutant pressure, given by a specific combination of measures?’’ Posterior probabilities for deductive reasoning and likelihood for inductive reasoning are calculated by BN software using Bayes’ Theorem. The information needed to set up the conditional probability tables for each node in the BN is based on results from simulation modeling, data analysis, or expert judgment.

In BN modeling, particular care must be paid to correlations between seemingly unrelated exogenous variables. In the example, precipitation (X) is simultaneously correlated with the environmental pressure (P), state of the environment (S), and the impact (I). Failure to identify such correlation can lead to overstating uncertainty regarding outcomes of management measures (Reichert and Borsuk 2005; Barton et al. 2008). BNs facilitate reasoning about and explicit identification of correlation structures with empirical data.

In modeling system dynamics an environmental impact leads to a system response, which in turn (drives actions that) lead to changes in pressure in the next time period. In such a system, the decision D to implement a management measure is often influenced by prior knowledge such as monitoring of the state of the environment from previous periods. Because BNs are acyclical, a decision node can therefore be interpreted as either a response to an impact on the environment, or as a driver of change to a pressure. Advances have recently been made in how to do dynamic modeling in BNs (Nicholson and Flores 2011). Johnson and Mengersen (this issue) also discuss how feedback loops similar to dynamic modeling can be implemented using BNs despite the acyclical properties.

Influence diagrams are mentioned here for completeness. When decision nodes (square in Figure 1) and utility nodes representing costs and benefits of actions and outcomes (triangles in Figure 1) are included in a BN, the resulting network is often called an ‘‘influence diagram’’ (Lauritzen and Nilsson 2001; Kjaerulff and Madsen 2008). Identifying states of decision and utility nodes makes it possible to evaluate the expected utility of decision alternatives and carry out optimal decision analysis in BN software. There are no articles applying ‘‘influence diagrams’’ to environmental and resource management in this special series, and few examples in the literature (Barton et al. 2008). However, Farmani et al. (this issue) discuss possible solutions to finding global optima for combinations of management measures in a BN.

BNs ADDRESSING DIFFERENT DECISION CONTEXTS IN ENVIRONMENTAL AND RESOURCE MANAGEMENT

Sutherland (1983) discusses how the role of uncertainty is strongly dependent on the type of decision context being modeled, categorizing contexts as either (1) directive, (2) strategic, (3) tactical, or (4) operational. In this special series
we bring together articles covering all 4 of Sutherland’s categories, from different environmental and resource management fields and spatial scales of analysis (Figure 2). Assigning articles to Sutherland’s categories is not clear cut because BNs may serve multiple objectives or evolve over time from directive to strategic, strategic to tactical, and so on as higher resolution data become available. Models may, for example, be scaled up from shorter term smaller scale tactical management applications to longer term, larger scale strategic applications. In BNs there is no clear cut distinction between ex ante hoc (before the fact) assessment and ex post hoc (after the fact) evaluation. BN evaluation is in media res (in the middle of things), subject to continual updating and iteration during a decision-making process.

**Directive**

The 2 articles that fall under the directive category focus on exploring long-term effects of decision options. In the process of building BNs, alternative model structures are explored, and the type of reasoning used is dominated by relational thinking, i.e., identifying causalities. Exploring causal model structure is a key element of directive decision making, and less emphasis is given on the detailed estimation of conditional probabilities by the data. Expert knowledge is mainly focused on relational rather than logical reasoning (Rowe and Watkins 1992).

Long time horizons often also involve consideration of larger spatial scales in environmental and resource management. For example, Varis et al. (this issue) demonstrate an approach to modeling integrated water resource management and development objectives in the Ganges Basin, at temporal and spatial scales that involve highly complex interconnected socioeconomic and biophysical problems. They propose using a matrix-based approach to explore causal structure and identify major driving forces of complex systems.

Alternative models are representations of reality and specific value-based choices about which features of nature and society are important in environmental and resource management. Different stakeholders have different values and understandings, so they have different ideas about what’s important to consider in the models used to support environmental and resource management decisions. Henriksen et al. (this issue) use BNs for communicating divergent views and understandings in adaptive groundwater management in the Guardiana Basin, Spain. They demonstrate how group model building and shared design of BNs can promote social learning, and clarification, if not a shared vision, of the trade-offs in the management of groundwater abstraction. A BN can help to focus discussion on the most relevant parts of the decision problem.

**Strategic**

When working in a strategic context, several important decisions often are in play and the time frame for solving them is often medium-term. BNs are used to explore the cumulative, and jointly uncertain impacts of the many different possible outcomes to these several important decisions, with the aim of finding, evaluating, and selecting (socially) acceptable outcomes. When multiple uncertain outcomes to combinations of several decisions must be compared, and outcomes are measured in different units that are not readily comparable, the process can be tedious. One consequence is that strategic decision analyses may be more prone to being influenced by analysts’ preferences and biases. Farmani et al. (this issue) demonstrate how an iterative multi-objective optimization algorithm can be used to search a BN for a “Pareto optimal” set of decisions among a large number of alternatives with uncertain outcomes. They demonstrate its usefulness for groundwater management decisions in Denmark.

Johnson and Mengersen (this issue) use BNs to evaluate different scenarios that affect cheetah population viability in Namibia. The article demonstrates how disciplinary models of factors affecting the cheetah population are integrated in a hierarchical object-oriented Bayesian network (OObN). OObNs are used to integrate knowledge from a number of expert domains, within which there is often consensus about causal structure, but where there is no common modeling platform across disciplines.

**Tactical**

Supporting decisions in the tactical context requires models that account for repeated observations and help managers react to short-term predictions. In many cases, the
models may include inductive elements, i.e., the models’ representation of what’s known about the state of the system is updated by using observations related to the decision at hand. Borsuk et al. (this issue) combine a number of biophysical and economic models to anticipate the outcomes of planned or recently completed river reach rehabilitation measures in Switzerland. In their integrated model there is less uncertainty about the dimension and cost of management measures than in a strategic modeling application. The smaller spatial scale of a river reach and specific management measures allow the authors to construct more accurate conditional probability distributions. They use causal networks using Monte Carlo based simulation arguing that these are more accurate in describing uncertainty of continuous variables than BNs, which often require discretization of probability distributions. This comes at the cost of not being able to conduct inference-based model assessment. They state that their modeling approach could also be scaled up and used for ex ante strategic planning purposes at the watershed level.

Although dynamic models are common to water quality issues, the complex ecology and multiple limiting nutrients of algal blooms and population dynamics are not analytically well understood. Johnson et al. (this issue) demonstrate the use of a ‘time-sliced’ OOB network to make month by month predictions of cyanobacteria blooms in Deception Bay, Australia. Object-oriented modeling offers an approach to making the otherwise acyclic BN models temporally explicit, offering a modeling tool with dynamic feedback loops for short-term forecasting. This article also discusses the different ways in which BNs may interact with geographical information systems (GIS). GIS may be used to obtain input data, to illustrate BN output, to model spatial interactions, or as part of a larger integrated modeling framework. As such GIS data may be used for both providing input data and illustrating impacts in strategic and tactical applications of BNs.

Uusitalo et al. (this issue) aim at providing a tool for operational understanding of risk in fisheries management, given significant environmental variation beyond managerial control. They use the expectation-maximization (EM) algorithm implemented in the Hugin Expert\textsuperscript{TM} software (www.hugin.com) (Laurizen 1995; Madsen et al. 2005) given available monitoring data and model structure. Because their model is based on spatially explicit fisheries and environmental monitoring data, it could conceivably be applied in both strategic and tactical decisions in fisheries management. Depending on the time resolution of monitoring data, a repeatedly updated BN could also be used “on-line” in operational fisheries management.

**Operational**

Operational decisions characteristically have known causal structure making it feasible to develop computer routines for analyzing alternatives and recommending actions, based on criteria provided by decision makers. Artificial intelligence and operations research approaches are used to conduct fast calculus in identification, classification, and diagnosis problems. This special series does not contain any illustrations of using BNs for real-time environmental and resource management. However, Uusitalo et al. (this issue) demonstrate the use of data learning algorithms that are common to diagnostic artificial intelligence applications of BNs.

Within each decision context proposed by Sutherland (1983), BNs with decision variables can be used to prioritize between environmental and resource management measures. These approaches are familiar in the decision theory and economics literature (Oliver and Smith 1990), including cost-effectiveness, benefit–cost, and multiple criteria analysis. Borsuk et al. (this issue) use their causal probabilistic network to conduct a benefit–cost analysis based on river rehabilitation cost and local economic benefits in terms of gross output and job creation. They model benefits and costs using Monte Carlo simulation in Analytica\textsuperscript{TM} software (www.lumina.com). BN software such as Hugin Expert\textsuperscript{TM} can also use “utility nodes”—expressed in the same utility or monetary units—to calculate maximum expected utilities of combinations of different decision alternatives. Farmani et al. (this issue) use an algorithm for evolutionary multi-objective optimization (EMO) for costs and benefits of management measures that are not expressed in the same units. The other articles in the special series represent examples of environmental risk assessment, viability, and scenario analysis but do not carry out economic decision analysis as such.

**ADAPTING BN MODELING TO THE PROBLEM STRUCTURE**

Sutherland (1983) points out the correlation in efficient and effective modeling between methodology and normative problem structure. Appropriate BN modeling for operational decision making is positivistic. It uses deduction processes derived from direct observation as inputs to optimization processes. Appropriate modeling for tactical decision making is predominantly inductive using empirical (historical) data as a basis for estimation (extrapolation and/or projection) exercises. “Probabilities” are interpreted as objective and amenable to analysis using statistical decision theory.

Appropriate modeling for strategic decision making is predominantly deductive, using “subjective” probabilities consistent with decision premises derived from a priori sources (general theory, axiomatic, or logical constructs). Subjective probabilities are interpreted as stochastic-state and decision making is predominantly for pre-adaption rather than expected utility maximization.

Appropriate modeling for directive decision making is heuristic where initial decision premises are of methodological significance. Modeling is designed to discipline the discovery processes under some sort of action research protocol.

Sutherland (1983) argues that modeling efforts may be inefficient or ineffective. At one extreme, modeling is ineffective if operational research theory is applied to a problem where the structure of causalities is still indeterminate. At the other extreme, a prescriptive model structuring process is inefficient if the problem is already well-known and deterministic.

Readers of this special series must make up their own minds as to whether they find the BN modeling examples appropriate for the decision problem and scale in each case. We make the case below that BNs are a generic modeling platform that can be adapted to available knowledge more easily and have a greater chance of being both effective and efficient.
HOW CAN BN APPROACHES CONTRIBUTE TO ENVIRONMENTAL AND RESOURCE MANAGEMENT?

In this section we highlight what Bayesian modeling approaches have to offer to environmental and resource management, as illustrated by the articles in this special series. We also question how these advantages may be specific to environmental and resource management contexts.

Social learning

Although numerous advanced numerical tools exist for covering the physical-technical system, these tools lack social acceptability and economic appraisal criteria—a socio-economic and policy interface as it were—that allow a wide range of stakeholders to identify with and participate in modeling. In this respect, Henriksen et al. (this issue) argue that the novelty of BNs is not as a technical tool, but rather as a process-structuring tool for dealing with adaptive challenges and “wicked problems” in groundwater management. They show how a BN graphical user interface was used to explore different causal structures reflecting the different values and interests held by stakeholders. The causal structure was repeatedly updated in a group learning process. Furthermore, determining numbers to enter in probability tables in a BN in a participatory fashion helped participants to learn the perceived importance of positions held by other participants. Henriksen et al. (this issue) also demonstrate that a participatory model-updating approach is one of the few routes available to validation of large integrated causal probabilistic networks that cannot be calibrated and validated jointly in the same sense as numerical simulation models.

Causal structure

Varis et al. (this issue) demonstrate how a link matrix approach (Varis and Lahtela 2002) can be an alternative in participatory modeling to the graphical user interface of BNs when the number of variables and links become prohibitively large for graphical representation. Their fully connected belief network (FC BeNe) was constructed with groups of domain experts and scientists who could be asked to participate in quite conceptual modeling required by a matrix interface. This may be the only approach to modeling with a large number of variables and long distributed causal chains. However, one could question whether the lack of a graphical representation limits the participation of nontechnical stakeholders in the kind of learning process discussed by Henriksen et al. (this issue).

Distributed modeling of domain knowledge

Several articles propose using OOBNs to take advantage of graphical representation of even complex models (Figure 3). Subdividing the network into submodels that can be independently specified by domain experts facilitates communication of complex nested model structure (Borsuk et al. this issue; Johnson and Mengersen this issue). OOBNs are also used to introduce time by “slicing” or replicating model structure over several linked time periods (Johnson and Mengersen this issue). Furthermore, OOBNs are used to replicate or distribute a model over several geographical locations in a watershed (Johnson et al. this issue). In reviewing these articles it is pertinent to ask why spatially explicit and temporally dynamic watershed simulation models were not used instead of probabilistic causal networks. The main reasons would seem to be a lack of knowledge of causal structures (a directive context) or a lack of data that would allow for calibration of dynamic models. In some cases a probabilistic causal network is used to express the main correlation patterns of underlying simulation or statistical models (Borsuk et al. this issue).

Resolution of uncertainty

Borsuk et al. (this issue) and Varis et al. (this issue) argue that the lack of parametric probability distribution functions in most BN algorithms, and the resulting need to specify discretized and/or discontinuous conditional probability tables is both cumbersome for large networks, and introduces additional uncertainty in particular nodes. This is particularly important in directive modeling exercises, where relational reasoning needs to be supported and the structuring of the
problem is particularly challenging and complicated. Borsuk et al. (this issue) use causal probability network in which continuous conditional probability distributions can be represented by Monte Carlo simulation in Analytica software. Varis et al. (this issue) use a spreadsheet matrix-based approach where continuous distributions can also be implemented.

An empirical question to both articles would be to what extent the parameters of continuous distributions—themselves the results of estimation—are a more accurate description of the data than a discrete probability distribution. The answer would depend on the number of data observations available to parameterize a distribution and the number of discrete intervals used to describe the data. The fewer the data observations and the more discrete intervals that are used, the less of a difference would be expected in “artificial noise” introduced by the modeler. In discretizing probability distributions, BN modelers need to find a balance between accuracy and precision. A smaller number of discrete intervals might increase model accuracy and decrease model precision and vice versa. The appropriate number of discrete states of a distribution depends on the modeling objectives. For example, when BNs are used with large data sets generated across a large number of locations or time periods, discretization may be both cumbersome and introduce unwanted noise.

Data learning

With a large data set discretization assumptions have a significant impact on the strength of causal links. Uusitalo et al. (this issue) use the data learning expectation-maximization (EM) algorithm of the Hugin Expert software to explore the correlations of environmental factors and fisheries productivity over a large number of multi-annual monitoring locations. They show how modeling causal structure and probability tables can be entirely defined by the data even if no priors are available. An advantage of BN learning algorithms is the use of all available monitoring data, rather than rejecting records with partial and/or missing information as in (some) classical statistical techniques. Uusitalo and colleagues also show how published literature can be used as prior information in a first BN to be updated by the monitoring data, or as a data set for validation of a BN estimated on the monitoring data alone. The choice of literature for external and independent validation of BN model results is affected by a publication bias toward model precision and vice versa. The appropriate number of discrete states of a distribution depends on the modeling objectives. For example, when BNs are used with large data sets generated across a large number of locations or time periods, discretization may be both cumbersome and introduce unwanted noise.

Decision analysis and optimization

Constrained optimization is well known in economics, but for largely deterministic problems using linear programming. Farmani et al. (this issue) use an evolutionary multi-objective optimization (EMO) algorithm to search a BN with a probabilistic and nonlinear structure for Pareto-optimal combinations of management measures. The EMO approach shows how optimization algorithms can be used to conduct an “external” assessment of decision alternatives specified in a BN. The authors show how EMO can be used to evaluate the overall coherence of decision alternatives relative to constraints on management outcomes.

The combination of articles demonstrates that BNs have been used across environmental and resource management issues as diverse as watershed development, water quality, groundwater abstraction, wildlife and fisheries management, and habitat restoration. Although the special series only represents a handful of BN articles in a rapidly growing field, the articles give a sense that BNs in environmental and resource management have been most widely used in mid-to-long-term strategic decision making contexts at intermediate spatial scales (e.g., catchments, habitats) and for problems where knowledge of causal structures is available (although perhaps contested). Furthermore, strategic environmental and resource management does not require high spatial or temporal resolution, often allowing the entire system to be described by aggregate state indicators (e.g., average annual water quality, groundwater level, population size, etc.). As computing power increases, BNs will increasingly be able to run in a spatially distributed “batch mode” with explicit GIS integration. On-line updating is also technically possible with commercially available software (e.g., Hugin). However, a limitation to these applications of BNs will continue to be the availability of spatially and temporally disaggregated monitoring data.

LIMITATIONS TO BNS IN MODELING DRIVER-PRESSURE-STATE-IMPACT-RESPONSE CHAINS IN ENVIRONMENTAL AND RESOURCE MANAGEMENT

The articles in this special series document a wide range of potential uses for BNs. However, the authors have also noted situations in which BNs may be inferior to alternative modeling approaches.

Different submodel domains have different spatial and temporal resolution. Borsuk et al. (this issue) illustrate how cross-scale interfacing of a low resolution subnetwork with a fine resolution subnetwork model requires “variable speed splitting” to make the states of each model’s conditional probability distributions compatible. This might be typical of “driver”–“pressure” model interfaces where e.g., watershed management measures determined by annual budgets, driving seasonal changes in landcover at farm level, which affects daily plot level run-off and streamflow. At the other state-impact-response end, the interface problem is reversed: a dynamic model may predict daily water quality states, but perceptions of recreational suitability can only be recorded on a seasonal basis, whereas responses by households in terms of willingness-to-pay for management measures are on an annual basis. Here an averaging over spatial and temporal units is required. Both steps require some manipulation of model input-output resolution, which introduces “artificial” variability to the causal chain. The multiscale issue is widely acknowledged in many disciplines and a growing number of Monte Carlo and other solutions are proposed for this purpose. It is of ongoing interest to investigate how these may be translated to the BN context.

Marcot and colleagues have discussed best practice guidelines for the size and structure of BNs for communication purposes, emphasizing modeling parsimony in the number of nodes and states of each node (Marcot, Stevenson et al. 2006). When using BNs as metamodels to integrate a number of...
different underlying model simulation results in a DPSIR approach, causal chains easily become longer than the length advised by Marcot and colleagues. Although this may not be such a problem for communication if an object-oriented BN is used, long causal chains combined with low resolution (few discrete intervals) of conditional probability distributions, and cross-scaling of input-output probability distributions between models, can lead to uncertainty attenuation between the “driver” management measures at the head of the DPSIR causal chain and “response” at the end of it (Barton et al. 2008). In these cases, cost utilities immediately linked to decisions will have a higher probability weighting than benefit utilities at the far end of the DPSIR chain. This may be a realistic expression of uncertainty. However, in a highly variable system it may be difficult to distinguish modeling “noise” from natural variability. The longer the causal chain, the more likely that systemic latent variables (e.g., climate) jointly affect several parts of the causal chain (see Figure 1). If these latent variable correlations are ignored, joint uncertainty in impact–response variables will be overestimated. In this case, the expected net benefits of management measures may be underestimated. It is an empirical question when the length of a causal chain, ignorance of latent variables, and low resolution of probability distributions combine to make environmental and resource management actions look cost-inefficient. Modeling uncertainty in a BN does not automatically resolve these issues, but it does offer a transparent way of thinking about such omissions.

There are some computational limitations that limit the role of economics in using BNs to support environmental and resource management. Utility nodes in influence diagrams must currently be expressed in a deterministic manner in commercially available software (Barton et al. 2008). This reduces the flexibility of BNs for evaluating uncertainty stemming from market and survey data on behavior of consumer populations.

Economists often explain variation in environmental costs and benefits across populations using regression techniques that produce Gaussian normal parameters. A continuous chance node cannot as yet be parent node of a discrete chance node in commercially available BN software using influence diagrams. These limitations have encouraged Borsuk et al. (this issue) to use Monte Carlo based simulation approaches to evaluate the economics of their river rehabilitation problem.

A RESEARCH AND DEVELOPMENT AGENDA FOR BNs APPLIED TO ENVIRONMENTAL AND RESOURCE MANAGEMENT PROBLEMS

Varis et al. (this issue) argue that no modeling approach is superior to others in all respects and a modeler’s toolbox should include a number of complementary approaches, including Bayesian tools. We note here that BNs are 1 form of Bayesian hierarchical model in the suite of Bayesian methods and models. In the words of Sutherland (1983): “[…] there are several different scientifically valid analytical approaches because there are several fundamentally different categories of problems.” We argue that this also holds perfectly for the application of BNs for environmental and resource management purposes.

Uusitalo (2007) reviewed the benefits and shortcomings of BNs, especially in relation to environmental modeling. Drawing on their findings and the articles in this special series we highlight some research questions that we feel are specific for BN applications to environmental and resource management.

Complexity and the role of adaptive management

Most of the published BN articles are stationary in the sense that only 1 time step is included. Usually this is simply the current state of the ecosystem, conditional on a single (set of) decisions. Truly dynamic models (several times steps and several decisions, like annual quota decisions in fisheries) are rare. Although BNs are acyclical, BN methodology as such allows for a dynamic setting (see Johnson and Mengersen this issue).

An obvious area of dynamic modeling would be to model approach of adaptive management (Holling 1978; Berkes et al. 2000; Nyberg et al. 2006), i.e., a management scheme where one of the aims is explicitly to learn more about the system managed. Certain management actions may lead to better future estimates of e.g., fish productivity, if monitoring is planned to provide better estimates. Henriksen et al. (this issue) argue that BNs are a tool for group model building, engagement of stakeholders, integration, and for dealing with ambiguity that are important issues in adaptive management. BNs can here be used for supporting stakeholder engagement, common dialogue and communication between stakeholders, authorities, and researcher communities.

Distinguishing between uncertainty and variability

It is important to distinguish between the different sources of variability in a model, which may arise due to the statistical representation of the sampling distribution of the data, differences between groups of observations, etc., and other sources of uncertainty in the model due to lack of information, missing data and so on. As BNs are especially developed to model uncertain features and uncertain relationships, it is a requirement that the uncertainty estimation is carried out by scientifically sound methods. Currently MCMC methodology (Markov chain Monte Carlo) is widely applied in fisheries analysis (Hilborn 2012) and hierarchical Bayesian models are used in many areas such as water quality modeling (Borsuk et al. 2001; Malve et al. 2007). Especially in cases where risk-averse attitudes are adopted by environment and resource managers (like the precautionary approach in fisheries), scientifically supported management actions should use uncertainty as 1 decision criteria. In particular, the tail of a probability distribution should play an important role. Therefore, the scientifically justified uncertainty estimate is important. We are not aware of any environmental and resource management applications of BNs that explore the significance of accurate modeling of the tails of probability distributions, and the implications for the expected utility and choice of management actions.

A usual approach in the field of artificial intelligence (AI), where BN are widely applied, is to use the observed data directly to learn the conditional probabilities. This may be justified in cases where there is a need to react directly to observed values (like operational decisions in the 4 categories above), and the interest is not to estimate the real states of nature, like the amount of fish in a lake or the total amount of nutrients entering a lake. However, in many cases the data do not accurately reflect the variability that is in the underlying biophysical processes. The sample average includes more
variation than the average value in nature, but the samples also includes less variation than the actual, realized single-point-in-time-and-space values in nature. For example, if we collect samples of P, N, and chlorophyll-A in a number of lakes, calculating a conditional probability distribution for the relationships of the 3 variables directly from observed data may include more (or less) uncertainty than there actually is in nature. A model may then give too pessimistic (or too optimistic) a view on our possibilities to manage the system by nutrient reductions to achieve a desired state of chlorophyll-A concentration in the water. Probabilistic process models and related likelihood functions are needed to provide realistic uncertainty estimates of the processes.

Recognizing prior knowledge in modeling

A fundamental advantage of Bayesian inference is that it allows for the use of informative prior distributions on uncertain model parameters. For example, literature data may be used to define priors, with study-specific data being used to update those priors. Using noninformative priors in decision analysis is important in that modelers are formalizing that they have no prior knowledge for a particular variable—from other studies or expert opinion—that has information value for the decision being modeled.

The importance of the use of informative priors should be noticed by the peer-reviewed journals as well. Although it is common to conduct literature reviews as a basis for hypothesis formulation, it is less common in environmental modeling to use prior model findings as a starting point for calibration in new applications. Publication systems could encourage the use of posterior values of earlier studies as priors for future studies (Kuikka et al. 2011). In many cases, the variance–covariance structures of the parameters may also play an important role in the effective scientific learning processes, and they should also be published. This is especially true in the cases of expensive data sets and cases where stakes of the end users are high. Having said this, it is important that there are formal, careful, and transparent protocols for eliciting and using informative priors to maintain scientific integrity.

Improving BN model graphical user interfaces

Available BN software can be improved in a number of ways to better facilitate communication of the science behind the network. The likely impact of the management actions is based on the strength of real causalities in the modeled system, for example how strongly the nutrients impact the chlorophyll concentration. For instance, this can be described graphically by the thickness of links (Varis and Fraboulet-Jussila 2002); this is implemented in Genie software (http://genie.sis.pitt.edu/). Similar information is provided in different ways in other BN software packages. Such “flows of causal strength” may help to explain the results of decision analysis to an end user, i.e., why some decisions dominate and why some are of less importance. Another improvement to the graphical user interface of BNs would be tools to distinguish real causalities from ones are just supposed to describe informative dependencies (likelihood dependencies between the reality and the observations).

It is also important to make clear in the model formulation, whether the model is used to predict uncertain features, such as in learned networks, or for the prediction of impacts of actions and decision analysis contexts as described by Sutherland (1983). In the first case, models can include a mixture of real causalities and links of purely correlative nature, whereas in the latter case the description of impacts of intervention should be based on documented causal relationships or strong theories of causation.

Improving model validity through communication and acceptability

The final aim of decision analysis is to have an impact on the behavior of humans, i.e., to convince them with scientific methods that a certain behavior is likely more useful than another behavior. To do this, the end user must understand the scientific message correctly. Commitment of stakeholders to the suggested actions is a necessity, and one may claim that an improved understanding of the process of how the scientific conclusions were made should play an important role (Haapasari and Karjalainen 2010).

Aguilera et al. (2011) also find that approximately 45% of the BN applications used for inference or classification did not conduct any model validation. The existing validation methods use mainly data as a test case, e.g., how well a model can predict a data set that was excluded from the data set that was used to estimate the parameters. See Marcot (2012) for a recent wide review of validation methods for BNs. However, it is quite common that a model is used to predict something that has never taken place, or at least never observed in existing data sets, such as the results of completely new policy. For example, the EU Water Framework Directive requests the return of water quality in lakes and coastal waters to at least “good ecological status,” which in some catchment may require a combination of total P, chlorophyll and cyanobacteria states which have rarely or never been observed (Barton et al. 2008). In such cases, experimental information, theoretical knowledge, or informative priors from other water areas are required to make realistic predictions. Validity of the model must be seen in a wider context, such as how well it respects the findings of earlier publications and existing understanding of causal processes in watersheds.

Furthermore, in model approaches that explicitly express ambiguity in model structure and includes subjective probabilities, “confidence” needs to be seen both as a statistical fact and a subjective belief, and validation must also demonstrate “buy-in” and acceptability by stakeholder. Henriksen et al. (this issue) document how development of a BN as a social learning process increases acceptability of stakeholders.

A related, but poorly studied, area is the connection of BN and cognitive processes of the human mind. Antal and Hukkinen (2010) discussed the challenge of communication of climatic change risks, and similar challenges are likely common in most environmental risk analyses. There is a need to consider which of those variables that can be modeled are the most relevant for the stakeholders and to help people to understand the key dependencies between their own safety and environmental quality (e.g., safe swimming without a threat of toxic cyanobacteria [Barton et al. 2008]). BN methodology offers a flexible framework for such studies, but it also meets the challenge to be able to communicate the results in an understandable way.
**Decision analysis and optimization**

Commercial BN software such as Hugin Expert\textsuperscript{TM} use the single policy updating (SPU) algorithm to search an influence diagram’s decision and utility nodes for the combination of decisions that maximize expected utility across decision alternatives. It is possible that the SPU algorithm sometimes settles in a local maximum (Hugin Expert Manual). Future research could therefore compare the advantages of add-in EMO algorithms discussed by Farmani et al. (this issue) to the utility maximization algorithms such as SPU that are integrated in commercial software.

**Value of information versus value of control**

In the support of decision making, value-of-information analysis (VOI) offers a tempting method to decide which parts of the uncertain system need further studies (Oliver and Smith 1990; Yokota and Thompson, 2004). Kuikka et al. (2011) suggested that one should first construct a model that is based on prior information only (publications, databases, experts, available theories), and by applying VOI analysis, decide the allocation of the research to alternative study domains. This offers possibilities to learn in a more systematic way.

However, there are also potential weaknesses in the suggested approach. Perhaps the most important one is that VOI must be based on an existing model structure, i.e., the network of causal relationships that is understood to potentially exist. However, basic research has one essential aim—to find new causalities that may offer totally new ways to understand and manage a socio-ecological system. BNs offer methodology to test how likely a causal connection needs to be to revise the suggested decision order, but naturally, if there is no prior probability for a certain causal connection, the justification of the study topic to funding organizations is problematic.

Environmental problems have several VOI applications (Varis et al. 1990; Kuikka et al. 1999; Yokota and Thompson 2004). However, another (even more important) feature of decision analysis is what kind of additional utilities could be achieved by adding new control mechanisms. Value-of-control analysis (VOC) is discussed in Oliver and Smith (1990), but applications are rare. VOC could be a valuable tool in the planning of new elements of environmental legislation. A case in point is the scarcity of probabilistic benefit-cost analysis of environmental measures and legislation, in particular decisions about adopting the precautionary principle in the context of high consequence low probability events.

**CONCLUSIONS**

We set out to provide an analysis of the place of BNs in the environment and resource management literature as exemplified by the articles in this special series. We conclude that BNs have mainly found their application to strategic decision-making, although the review provides examples of directive, tactical and operational contexts as well. The contribution of BNs to environmental and resource management are in particular through

- better organization of distributed modeling of domain specific knowledge;
- the explicit acknowledgment of the sensitivity of models to different levels of resolution of probability distributions;
- the ability to combine prior knowledge of causal structures with data learning algorithms; and
- the facilitation of context dependent decision analysis.

As reviewed by several authors previously, the same limitations that apply to the use of BNs in environmental and resource modeling on its own apply to modeling applied to management and decision making:

- the acyclical structure of BNs make specification of dynamic problems cumbersome; and
- BNs as applied in commercially available software are constrained to use either probability tables or Gaussian normal distributions, with some loss of accuracy.

Specific to environmental management of DPSIR modeling is the problem that causal chains can become long enough to lead to bias toward concluding that management actions are “ineffective” due to signal dissipation, as well as making communication of the causal structure of integrated models more challenging. Models adapted to the decision problem context, with good modeling and validation practices can to some extent address these problems. Recognizing that BNs are a form of hierarchical Bayesian model, modelers may also use alternative Bayesian models using MCMC or other computational methods for analysis.

We conclude by outlining some future research challenges for BNs: addressing complexity by modeling adaptive management; distinguishing uncertainty and variability; recognizing prior knowledge in modeling; greater focus on model communication and stakeholder acceptability as criteria for model validity; and increasing the use of BNs for benefit–cost analysis of monitoring and management measures, also known as value of information and value of control analysis.

**Acknowledgment**—The authors would like to thank the Strategic Institute Project “Nature 2020 +” and the Research Council of Norway for funding. We also thank Bruce Marcot, Kerrie Mengersen, Anders Madsen, and an anonymous reviewer for valuable comments and improvements to the manuscript. All errors or omissions are the responsibility of the authors.

**REFERENCES**


