

# Accepted Manuscript

Advances in Bayesian network modelling: Integration of modelling technologies

Bruce G. Marcot, Trent D. Penman

PII: S1364-8152(18)30293-7

DOI: [10.1016/j.envsoft.2018.09.016](https://doi.org/10.1016/j.envsoft.2018.09.016)

Reference: ENSO 4312

To appear in: *Environmental Modelling and Software*

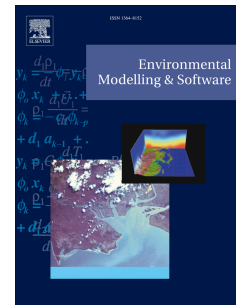
Received Date: 27 March 2018

Revised Date: 22 August 2018

Accepted Date: 22 September 2018

Please cite this article as: Marcot, B.G., Penman, T.D., Advances in Bayesian network modelling: Integration of modelling technologies, *Environmental Modelling and Software* (2018), doi: <https://doi.org/10.1016/j.envsoft.2018.09.016>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Advances in Bayesian network modelling: Integration of modelling technologies

Bruce G. Marcot<sup>a\*</sup>, Trent D. Penman<sup>b</sup>

<sup>a</sup>USDA Forest Service, Portland, Oregon, USA

<sup>b</sup> School of Ecosystem and Forest Sciences, University of Melbourne, Victoria, Australia

\*Corresponding author.

*E-mail address:* [bmarcot@fs.fed.us](mailto:bmarcot@fs.fed.us) (B. G. Marcot).

FOR JOURNAL: Environmental Modelling & Software

#### ARTICLE INFO

*Article history:* \_\_\_\_\_

*Keywords:* Bayesian network, decision modeling, integrated modeling, monitoring

## ABSTRACT

Bayesian network (BN) modeling is a rapidly advancing field. Here we explore new methods by which BN model development and application are being joined with other tools and model frameworks. Advances include improving areas of Bayesian classifiers and machine-learning algorithms for model structuring and parameterization, and development of time-dynamic models. Increasingly, BN models are being integrated with: management decision networks; structural equation modeling of causal networks; Bayesian neural networks; combined discrete and continuous variables; object-oriented and agent-based models; state-and-transition models; geographic information systems; quantum probability; and other fields. Integrated BNs (IBNs) are becoming useful tools in risk analysis, risk management, and decision science for resource planning and environmental management. In the near future, IBNs may become self-structuring, self-learning systems fed by real-time monitoring data. Such advances may make model validation difficult, and may question model credibility, particularly if based on uncertain sources of knowledge systems and big data.

**Keywords:** Bayesian networks; decision models; model integration; machine learning; model validation

## 1. Introduction

Bayesian networks (BNs) are directed acyclic graphs that link variables by conditional probabilities, where model outputs are probabilities calculated using Bayes' Theorem (Fenton and Neil 2012, Koski and Noble 2011). BN modeling is useful for data mining, determining and explicitly displaying the relationship among variables, representing expert knowledge and combining expert knowledge and empirical data, and identifying key uncertainties (Cheon et al. 2009, Hanea et al. 2010, Landuyt et al. 2013). Outputs are typically expressed as probabilities of various states, which lends well to decision-science approaches to risk analysis and risk management (Aalders 2008, Farmani et al. 2012).

General network structure of BN models is highly flexible, leading many researchers to find new areas of application. As examples, BNs have become popular in environmental management for projecting potential impacts of proposed projects (Krüger and Lakes, 2015), forecasting impacts of environmental disturbances such as fire (Dlamini, 2010) and climate change (Sperotto et al., 2017) and providing a basis for making environmental management decisions (Barton et al., 2012). Many examples are available of the use of BNs in a wide variety of other environmental and resource management contexts, such as management of groundwater (Giordano et al., 2013), recreation impacts (Fortin et al., 2016), and green energy production (Carta et al., 2011). If the BN contains no random variable, then the outcome generated is fixed, i.e., deterministic, for a given set of priors, else the outcome is stochastic. BNs can be made stochastic by introducing random deviates as part of formulae within nodes. Variables also can be described with formulae combining values of parent nodes, such as used by Steventon et al. (2006) in assessing viability risk of a rare seabird. Further, variables can be denoted with continuous ranges, rather than discrete state conditions, such as used by Hradsky et al. (2017) to determine impacts of fire and other stressors on the distribution of terrestrial wildlife.

BNs are a highly useful tool for depicting and modeling current knowledge, such as with initial representation of a system or problem to gain a better understanding and perspective on uncertainties and complexities so as to help advise managers and decision-makers. BNs provide a robust statistical framework when little data are available. An example is in the context of environmental modeling with time-critical situations with scant available data, such as active

monitoring of key energy infrastructures (Guerriero et al. 2016) and surveillance of endangered species (Koen et al. 2017).

Increasingly, BNs are being integrated with other modeling constructs and tools, such as geographic information systems (GIS) and remote sensing databases. In this paper, we explore these new avenues of how BN model development and application are being joined with many other tools and model frameworks for a variety of environmental assessment and management objectives. We briefly review the current state and recent advances of BN modeling, and then provide examples of an emerging new era of integrating BN models with other frameworks and tools. Lastly, we present a vision of next advances to come, concluding with a perspective on ensuring scientific and decision-making credibility, with cautions on accelerated model advancements.

## **2. The cutting edge of BN modeling**

The field of BN modeling is advancing swiftly with the number of journal articles using BNs continuing to rise, including a recent era of exponential growth (Marcot 2017). Recent developments in BN modeling are reviewed in the following sections.

### *2.1. Recent advances in BN model structure and applications*

A number of recent advances in BN model structure and application have followed a diverse track of topics, generally related to identifying and exploring system dynamics and aiding decision management (Fig. 1).

*Classification.* One area of resurgence is in new approaches to the classification problem, viz., Bayesian classifiers and machine learning. BN classifiers include a wide variety of algorithms starting with naive Bayes and variants thereof (Bielza and Larrañaga, 2014). Related to this are algorithms for Bayesian learning of probability structures from empirical data (e.g., Tsamardinos et al., 2006; Do and Batzoglou, 2008).

*Latent variables.* A common problem in ecological or environmental modeling is the influence of latent variables, which are effects inferred from the relation among observed variables but which are not directly observed (Marcot 2017). Machine-learning algorithms used in parameterizing the probability values in BN models, such as the expectation maximization algorithm (Do and Batzoglou 2008), can, to some degree, account for the influence of latent variables and missing data (Lauritzen 1995).

A related problem is how to validate BN models developed entirely from expert elicitation with no case-file data by which to structure or parameterize the model. Such models portray logical or causal relations among variables as inferred by expert knowledge, but these relations often are influenced by unspecified, latent variables (de Waal et al. 2016). Pitchforth and Mengersen (2013) proposed methods for evaluating confidence in the validity of such models even in the dearth of empirical data and presence of latent variables, thus providing a validation framework for expert-elicited BNs. de Waal et al. (2016) suggested several approaches to handling latent variables in BNs, including explaining uncertainties associated with latent variables, parameterizing the probability values of BNs so as to directly address the roles of latent variables, and addressing uncertainty in model validation.

*Depicting model confidence.* Although BN models explicitly incorporate uncertainty, there is no common method for depicting and quantifying the degree of confidence in the underlying probability values of the model and in the resulting posterior probability calculations. That is, uncertainty measures can be inferred from the probability distributions of states calculated in the model, but these are not necessarily the same as the degree of confidence (level of certainty) of that probability distribution. Pitchforth and Mengersen (2013) characterized confidence in BN model behavior as consisting of three components of structure confidence, discretization confidence, and parameterization confidence. For use in BN model validation, they further adduced 7 common dimensions of validity as used in psychometry: nomological, face, content, concurrent, predictive, convergent, and discriminant validity, which, collectively, pertain to the degree of concordance within accepted norms and credibility within a particular discipline.

A more quantitative method of depicting BN model confidence, as developed by Van Allen et al. (2008), entails estimating error bars around posterior probability calculations from

BNs, which then depict the degree of uncertainty (or confidence) in model outcomes. Error bars for BNs are termed credible intervals, which provide the range of model outcomes within a specified probability level (Marcot 2012), but which are not often reported in BN modeling projects. Such error bars should not be confused with frequentist confidence intervals. Hamilton et al. (2015) used credible intervals to measure the strength of the relationship between suitability of habitat of a crayfish and environmental predictor variables.

*Links to GIS.* Probability outcomes from BN models for evaluating local conditions have been used as input to GIS systems to create maps depicting habitat quality of wildlife species (Raphael et al. 2001, Havron et al. 2017). Kininmonth et al. (2014) presented a model which combined spatial datasets, spatial models, and expert opinion in an integrated BN-GIS structure for evaluating boating damage to the Great Barrier Reef of eastern Australia. Dlamini (2010) developed a BN-GIS model that uses geographically-referenced remote sensing MODIS data to analyze wildfire in Swaziland. BN programs that integrate or intersect with GIS include GeoNetica® (Norsys Inc.), HUGIN® (HUGIN Expert A/S), and Ecosystem Management Decision Support (EMDS, Reynolds et al. 2014) that integrates the GeNIe BN modeling platform with two GIS components of ArcMap® (Esri) and open-source QGIS. Gonzalez-Redin et al. (2016) BNs linked to GIS to map trade-offs of ecosystem services in the French Alps to inform planning decisions. Several projects have explicitly integrated GIS and BN modeling frameworks, such as the QGIS plug-in for BNs developed by Landyut et al. (2015) and the integration of the GeNIe BN modeling framework into the ArcGIS-based Ecosystem Management Decision Support system (emds.mountain-viewgroup.com). *Dynamic Bayesian networks.* Other recent variations on the traditional BN modeling theme include dynamic Bayesian networks (DBNs) that model a time series of conditions and contingencies, such as with oscillating predator-prey dynamics (Fig. 2). DBNs typically contain feedback loops which are not allowed in the directed acyclic graph structure of BNs, but can be modeled when BNs are time-expanded so that the entire BN structure is replicated for different time periods so that the links become acyclic. In some cases, DBNs have been made spatially explicit by integrating with geographic information systems (GIS; e.g., Chee et al., 2016). A variant of DBNs are those operating in real time in response to discrete or continuous inputs, such as for predicting highway crashes (Hossain and Muromachi, 2012) and in analyzing gene networks (Kim et al. 2003). New approaches to structuring DBNs combine methods from static and dynamic networks (Vlasselaer

et al. 2016). Uusitalo et al. (2018) used DBNs with hidden variables to model major structural changes of a Baltic Sea food web, and Orphanou et al. (2014) used "temporal Bayesian networks" (TBNs, a synonym for DBN) to evaluate temporal relationships in clinical data for medical diagnosis and prognosis. Their hidden variables represented unobserved processes contributing to the changes and resulted in DBN models that reflected known dynamics of the food web system.

*Bayesian decision networks.* Bayesian decision networks (BDNs) extend BN models by explicitly including decision and utility nodes (e.g., Barton et al. 2008). Decision nodes are deterministic nodes that depict unique management decisions, and utility nodes are continuous nodes that estimate a cost or benefit of a given outcome resulting from a decision. BDNs use utility nodes to calculate overall expected values of all costs or benefits of alternative management decisions, given the probability structure of the model, and can be highly useful in risk analysis and risk management arenas. For example, Loyd and DeVore (2010) developed a BDN to advise on alternatives for management of feral cats in the United States. Catenacci and Guipponi (2013) used a BDN as a basis for adaptation planning to sea-level rise. A further variation of BDNs is with dynamic decision networks (DDNs) that essentially merge decision networks with time-expanded dynamic networks. DDNs were developed by Murray et al. (2004) to guide selection of teaching tutorials, and by Penman et al. (2015a) to advise on reducing risk of loss of homes to wildfire. BNs have also been used to assess value of information to optimize resource use decisions, such as the fisheries industry (Kuikka et al. 1999).

*Depicting causality in structural equation models.* BN modeling has been compared to structural equation modeling (SEM) in that both can be used to depict causal networks and influences and can analyze degree of causality (Pearl 1998, 2000). The two approaches differ in that SEM is a general suite of statistical tools usually using frequentist, multivariate approaches (although some SEM approaches also support Bayesian estimation), whereas BNs use conditional probabilities and Bayes' theorem. A main difference is that SEMs are purely statistical tools developed, for example, to test hypotheses or to test whether an assumed causal relation in the graph is significant, whereas BNs are probabilistic models (trainable by data) mainly for investigating the consequences of conditions or events on outcomes, or deducing causal conditions resulting in an outcome.



More recently, Li et al. (2018) compared and combined BN and SEM modeling to evaluate the interactive influence of land use and climate change on stream macroinvertebrates. They first used SEM to develop a conceptual influence diagram of causal effects, that is, the network structure of variables and their linkages, and then they built prediction and diagnostic BNs from the same conceptual model. Using SEM helped identify and justify the links used in the BN models. Their results suggested that modeling all causal factors together in the SEM conceptual model and in the subsequent BN model provided a more robust understanding of how positive effects from climate change could mollify negative influences from land use.

*Neural networks.* A somewhat different variant of BN modeling appears in Bayesian neural nets, that is, using Bayesian learning to determine neural network node weights. Neural networks are typically trained using a variety of approaches including variants of gradient descent and least-squares methods to minimize loss functions of variables singly or in conjugation. Using BNs can bring greater efficiency in adjusting node weight parameters based on prior knowledge. Bayesian neural networks have been used to forecast energy load requirements (Lauret et al., 2008), solar irradiation (Yacef et al., 2012), stock market performance (Ticknor 2013), and internet traffic loads (Auld et al. 2007).

*Continuous Bayesian networks.* Another area of recent interest and progress is in developing continuous BN models where quantitative variables are not discretized into exclusive state ranges but instead are represented by continuous values such as equations or statistical distributions. Continuous BNs can be constructed using programming tools such as UNINET (Cooke et al. 2007, Delgado-Hernández et al. 2012), WinBUGS (Kery 2010), AgenaRisk® (Neil et al. 2007), Hugin® (Madsen et al. 2003), and GeNIe (Druzel 1999). Other BN modeling and graphical modeling software also can deal with continuous nodes<sup>1</sup> if the multivariate normal is assumed (A. Hanea, pers. comm.)

*Hybrid Bayesian networks.* Some researchers have developed BN models with both discrete and continuous variables where the latter are not discretized (Aguilera et al. 2010, Castillo et al. 1998, Driver and Morrell 1995). These types of models are referred to as hybrid BNs (HBNs; Hanea et al. 2006). A special case of HBNs, called non-parametric BNs (NPBNs), was reviewed by Hanea et al. (2015). NPBNs were initially devised for continuous-only BNs

---

<sup>1</sup> <https://www.cs.ubc.ca/~murphyk/Software/bnsoft.html>

but are used in situations of HBNs as well. Hanea et al. (2010) developed a NPBN methodology with data mining to develop prediction models. Hradsky et al. (2017) used NPBNs to model a presence-absence continuous response of wildlife to fire age classes and terrestrial vegetation classes depicted with discrete variables.

*Object-oriented Bayesian networks.* A further area of recent exploration is with object-oriented Bayesian networks (OOBNs) and dynamic object-oriented Bayesian networks (DOOBNs) (Bangsø et al. 2004, Benjamin-Fink and Reilly 2017). For example, OOBNs and DOOBNs have been used to model health impacts of cyanobacteria blooms (Johnson et al. 2010), viability of populations of cheetahs (*Acinonyx jubatus*) in Namibia (Johnson et al. 2013), and issues of water resource management (Phan et al. 2016). The tools Hugin® and AgenaRisk® provide for true OOBN modeling, and GeNIe® also can be used as such although it is not a true OOBN framework.

*Agent-based modeling.* Related to OOBNs is the merging of agent-based models and BNs. Agent-based models (An 2012) are simulations of the dynamics, such as movement patterns, of individual objects. Nielsen and Parsons (2007) developed a model of consensus-building where individual agents were represented by BNs that expressed a range of possible agreements. Sun and Müller (2013) presented a BN agent model to explore the economics of ecosystem services and land-use decision-making.

*State-and-transition modeling.* Another area of integration is with state-and-transition models (STMs) that are used to project future proportions or amounts of conditions, such as landscape vegetation conditions and species responses, under known or hypothesized rates of change (Mason et al. 2017). STMs project future conditions, such as area covered in vegetation type categories, by multiplying a matrix of current area in each category by a matrix of probabilities depicting transitions to the same or other categories (e.g., Jorgenson et al. 2015). In a hybrid STM-BN model, transition probabilities are estimated from calculations in the BN network that account for environmental influences on each vegetation type category. Bashari et al. (2009) developed an integrated STM-BN model, as expanded upon by Nicholson and Flores (2011), to inform management decisions in rangelands of Queensland, Australia. Chee et al. (2016) integrated STMs and BNs in a geographic information system (GIS), with object-oriented

concepts, to model spatial and temporal changes in an Australian woodland and a wetland in Florida.

*Quantum Bayesian networks.* BNs are being increasingly used in the area of quantum information theory as quantum Bayesian networks (QBNs; Tucci 1995). QBNs are constructed to represent outcomes that deviate from, and are paradoxical to, classical probability calculations. Examples include when outcomes are dependent on the sequence of inputs (priors and parent nodes in a BN); when human decision-making deviates from dominant probability outcomes in a BDN; when a system can result in  $>1$  dominant probability outcome state; and other situations. Such outcomes could be modeled in traditional BNs by including latent variables but such models quickly become overly complex and serve only to describe specific conditions and outcomes, not to serve as predictive and explanatory models.

Generally, in QBNs, classical conditional probability tables using Bayes calculus are replaced by quantum probability amplitudes (a complex number function that describes the behavior of a system). Moreira and Wichert (2018) developed a decision-based QBN with quantum probability amplitudes to demonstrate how prediction of some aspects of human decision-making is more efficient than with a traditional BN with latent variables. Similarly, Trueblood et al. (2016) used a QBN approach to model how human judgment can deviate from classical probability in the face of high uncertainty and imperfect information about causality of a system. Busemeyer and Trueblood (2009) explored the use of QBNs in quantum theory to model how different sequences of measurements can affect the probabilities of system outcomes. Leifer and Spekkens (2013) developed a QBN framework for depicting how quantum conditional states can result from the influence of two systems at one time or from one system at two times. Other formulations and applications of QBNs are found in the literature, although at present there does not seem to be any generally available software by which QBNs can be constructed.

*Power PC theory and causal BNs.* Power PC ("probabilistic contrast") theory states that a system outcome is the sum effect of the relative power of observed and unobserved causal relations which can be depicted and partitioned mathematically (Cheng 1997, Norick and Cheng 2004). In applying power PC theory, Lu et al. (2008) demonstrated how the relative influence of

different, independent causes can be determined empirically and can be represented in Bayesian causal networks.

Here, we have covered a range of significant recent advances in application of BNs. Still other variations and new approaches to BN modeling continue to appear in the literature.

## 2.2. *Beyond the network: a new era of integration*

BN models of various forms have been increasingly used in a variety of applications. They are also being specifically integrated with other modeling constructs, which we refer to here as integrated Bayesian networks (IBNs; Johnson et al. 2010). IBNs can be defined as BN structures that are explicitly embedded within the framework of other modeling constructs, instead of just being applied to some area of inquiry as reviewed in the previous section. Examples of IBNs include assimilating BNs in structured decision-making frameworks, agent-based models, and state and transition models (Table 1). IBNs are crafted generally to apply the probabilistic basis of BNs to new areas of application and research such as dynamic and stochastic simulation modeling.

IBNs are also becoming useful tools in risk analysis, risk management, and decision science, such as in environmental resource planning (Johnson and Mengersen 2012, Fraser et al. 2017) and for depicting how deep uncertainty affects policy decisions (Aven 2013, Cox 2012). Janssens et al. (2006) developed an IBN that combined BNs and decision trees for developing decision rules in transportation management. QGeNIe Modeler® (BayesFusion, LLC) has a graphical user interface that provides for rapid prototyping of decision models in a BN environment. Some BN modeling platforms such as Netica® (Norsys Inc.) and Hugin® provide application program interfaces (APIs) to facilitate integration links to other programs such as geographic information systems which can facilitate their use in integrated risk analysis and structured decision-making under uncertainty (e.g., Barton et al. 2008).

An emerging area is the development of IBNs operating from real-time monitoring data. For example, Maglogiannis et al. (2006) have proposed a patient-health risk analysis system

using IBNs operating from vital sign monitoring data. Their vision is to produce a real-time system for homecare telemedicine. Penman et al. (2015b) developed a fire danger rating system that updated daily with meteorology forecasts, and as new fires appeared in the landscape the model automatically updated the risk projections. Vagnoli et al. (2017) proposed a real-time, IBN-based system to monitor the structural integrity of railway bridges in Europe. Koen et al. (2017) developed an IBN model to monitor the poaching (illegal hunting) of rhinoceroses in Kruger National Park, South Africa, which is being implemented as a real-time management tool.

### 3. Things to come

The field will continue to advance more rapidly into uncharted territory as IBNs become more sophisticated. Future IBN research foci will include substantial cutting-edge advancements in the many areas of integration reviewed above. Currently, IBNs are essentially static; however, we expect that in the near future, IBNs will emerge that are self-updating and self-improving, and that will learn from real-time continuous input of environmental monitoring data. Being able to dynamically update BN conditional probability values with new data (e.g. using an expectation maximization algorithm), and to continually recalculate posterior probabilities, are quite feasible with existing software and hardware. Some current IBNs do automatic recalculation of posterior probability outcomes with new input data feeding the models, but here we are referring also to the basic structure and underlying conditional probability tables themselves being created anew and updated with machine-learning tools. The purpose of this would be to continually refine the context, accuracy, and robustness of the models, which is a fundamental precept and advantage of Bayesian statistical approaches that improve model predictions from prior data. Some aspects of this have appeared in the application of BNs in areas of finance (Giudici and Spelta 2016, Garvey et al. 2015), and medical diagnosis and decision-support (Constantinuo et al. 2016). In the future, we envision self-creation of IBNs based on emergent information from crowdsourced data (Park and Budescu 2015).

Also to come will be the further integration of BNs with expert system knowledge bases, particularly using fuzzy-logic or neural-net forms of knowledge-representation and machine-learning algorithms. In the past, control-rule-based expert systems used confidence factors or

scoring rules (Zohar and Rosenschein 2008), e.g., indices as provided by the expert and scaled [1,10], to rank the likelihood or the credibility of a particular inference or outcome. Today, the BN modeling shell BayesiaLab® (Bayesia S.A.S.; Conrady and Jouffe 2015) averages conditional probability values from multiple experts and weights the values by the product of confidence in, and credibility of, the values as scored by each expert.

#### **4. Conclusions and perspectives**

What are some of the main cautions and caveats in this new era of IBNs? Here we address quandaries of ensuring validity and credibility of IBNs that are becoming increasingly complex in construction and interpretation, particularly as they are induced from automated and big data sources.

##### *4.1. Ensuring a future of validity and credibility*

It will be important to ensure the validity and credibility of increasingly-complex expert-based IBN systems (Kleeman et al. 2017), particularly as they interact with human social systems, as they guide resource management decisions, and as they operate with greater autonomy. As IBNs become more complex, it will become more problematic to create and test simple, intuitive, and understandable influence diagrams and mind maps that chart their structures, logic, and operation. A way around this could entail decomposing such complex systems into simpler component submodels, and testing and updating each submodel. BNs are generally constructed as Markov processes so that they can be dissected and reassembled without loss of information, particularly using cutpoints in the network graph (nodes in the BN whose removal would separate the graph).

Direct and indirect causes in process models will be increasingly difficult to clearly identify, particularly with models having multiple interaction terms, feedback loops, latent variables, and synergistic functions among covariates and response variables. This also means increasing difficulty in parsing out sensitivity and influence effects to specific drivers, that is, clearly identifying key factors that most influence outcomes and for which management might

have greatest uncertainty and greatest (or least) control. The result may limit their acceptance and adoption in management situations if users do not understand and trust the models and if the models have limited face validity. This can be a major issue with decision networks and use of models in real time, and with models intended to guide planning and management of resources with high opportunity costs. Support for the model, as with advances in validation methods and frameworks as discussed above, is needed. Some work advancing methods of sensitivity analysis of BN models may show promise for further development and application on increasingly complex model structures (e.g., global sensitivity analysis methods of Li and Mahadevan 2017).

Increasing complexity of IBNs will carry increased difficulty in model calibration, testing, validation, and updating, eventually necessitating new heuristic approaches and algorithms that can wade the swamp of big data (Spiegelhalter 2014, LaDeau et al. 2017, Lv et al. 2014). Lewis et al. (2018) further warned that developing wildlife biology models even with high-quality big data sets raises concerns for how opaque modeling algorithms can lead to complex problems of data management, exploratory data analysis, data-sharing, and reproducibility. Paradoxically, big data sets with many variables often are sparse in terms of replicated conditions and variable combinations (Hastie et al. 2015). Although BN models induced from big data may fit well, that is, with high calibration accuracy, they may lack robustness and perform poorly because of this sparsity, that is, the models may have low independent validation accuracy and be overfit. Also, BNs built on big data collected through citizen science initiatives, which are becoming increasingly popular in many topic areas, should be carefully vetted for the accuracy of those data (Kosmala et al. 2016), or else the models and their interpretations may be incorrect, biased, and misleading.

Particularly with self-generating and self-updated systems, IBNs will be no more trustworthy than the data on which they are based and on the opacity of the algorithms used to structure their networks and parameterize their probabilities. In the end, the validity and credibility of IBNs are inextricably linked to those of the data and the methods on which they are based and created. At stake is clear demonstration of expert knowledge, empirical data, and model simulations, and ultimately of model validity, operational robustness, and the credibility of the modeling science itself.



Despite potential worries over the testing and validity of big-data-sourced IBNs, and the quickly-evolving environments in which IBNs are constructed, their strengths will remain in "what if?" scenario exploration, and in being able to combine soft and hard evidence, that is, expert knowledge and empirical data. Even with an exponential increase in publicly available information, combining expert knowledge with data is best conducted so as to detect and counteract any spurious correlations that may be indicated by machine learning algorithms. Models that are structured and parameterized solely by use of "blind," automatic methods such as unguided machine learning algorithms, without human oversight, lose the key advantage of the Bayesian approach to updating human knowledge and insight.

#### *4.2. Cautions and caveats in the new era of integration*

We have tried here to chart some routes that development of IBN systems may be taking, and some cautions indicated in this fast-evolving era (Salmond et al. 2017). Soon to come will be dynamic and self-modifying and even self-creating IBNs using "big data" garnered from automated monitoring (e.g., environmental and Earth Science remote sensing data; Li et al. 2016), citizen science initiatives, and even crowdsourced information sources. We ask, to what degree should we trust such models? For, in the end, the knowledge source and expertise that may serve to generate and fuel IBNs may become a fully realized global, emergent artificial intelligence. If so, how shall the models' veracity be determined when their source becomes non-human?

BNs developed from empirical data often can be tested using existing cross-validation and jackknifing algorithms. However, the validity of BNs developed from expert knowledge is more difficult to determine if independent data sets by which to test the models are unavailable, or if specific combinations of variables in big data sources are sparse; in such cases, peer review of the model at various stages of development is essential to establish credibility and conformity with accepted precepts. But determining the external validity of self-organizing and self-updating BNs that are developed with deep-learning algorithms in real time from big data sources, including citizen-science and crowd-sourced data (especially the multitude of Internet blog sources and news posts), without independent testing or review, may be most problematic



and will require new approaches to scientifically evaluating their veracity. This will become one of the greater challenges in the fast-evolving new era of BN model integration.

### **Acknowledgments**

Inspiration for this paper comes from a keynote address given by the senior author in 2017 at the Joint Conference of the Australasian Bayesian Network Modelling Society and the Society for Risk Analysis Australia and New Zealand, at the University of Melbourne, Australia. We thank Tom Bruce, Anca Hanea, Sandra Johnson, and two anonymous reviewers for their helpful reviews of the manuscript.

## References

- Aalders, I. 2008. Modeling land-use decision behavior with Bayesian belief networks. *Ecology and Society* 13(1):16.
- Aguilera, P. A., A. Fernández, F. Reche, and R. Rumi. 2010. Hybrid Bayesian network classifiers: application to species distribution models. *Environmental Modelling & Software* 25(12):1630-1639.
- An, L. 2012. Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling* 229:25-36.
- Auld, T., A. W. Moore, and S. F. Gull. 2007. Bayesian neural networks for Internet traffic classification. *IEEE Transactions on Neural Networks* 18(1):223-239.
- Aven, T. 2013. On how to deal with deep uncertainties in a risk assessment and management context. *Risk Analysis* 33(12):2082-2091.
- Barton, D. N., T. Saloranta, S. J. Moe, H. O. Eggestad, and S. Kuikka. 2008. Bayesian belief networks as a meta-modelling tool in integrated river basin management — Pros and cons in evaluating nutrient abatement decisions under uncertainty in a Norwegian river basin. *Ecological Economics* 66:91-104.
- Barton, D. N., S. Kuikka, O. Varis, L. Uusitalo, H. J. Henriksen, M. Borsuk, A. de la Hera, R. Farmani, S. Johnson, and J. D. C. Linnell. 2012. Bayesian networks in environmental and resource management. *Integrated Environmental Assessment and Management* 8(3):418-429.
- Bashari, H., C. Smith, and O. J. H. Bosch. 2009. Developing decision support tools for rangeland management by combining state and transition models and Bayesian belief networks. *Agricultural Systems* 99:23-34.
- Benjamin-Fink, N., and B. K. Reilly. 2017. A road map for developing and applying object-oriented Bayesian networks to “WICKED” problems. *Ecological Modelling* 360:27-44.
- Bielza, C., Larrañaga, P., 2014. Discrete Bayesian network classifiers: a survey. *ACM Computing Surveys* 47(1) doi:10.1145/2576868.

- Busemeyer, J. R., and J. Trueblood. 2009. Comparison of quantum and Bayesian inference models. Pp. 29-43 in: P. Bruza, D. Sofge, W. Lawless, K. van Rijsbergen, and M. Klusch, editors. *Quantum Interaction. QI 2009. Lecture Notes in Computer Science*, volume 5494. Springer, Berlin, Heidelberg.
- Carta, J. A., S. Velázquez, and J. M. Matias. 2011. Use of Bayesian networks classifiers for long-term mean wind turbine energy output estimation at a potential wind energy conversion site. *Energy Conversion and Management* 52:1137-1149.
- Castillo, E., J. M. Gutierrez, and A. S. Hadi. 1998. Modeling probabilistic networks of discrete and continuous variables. *Journal of Multivariate Analysis* 64:48-65.
- Catenacci, M., Guipponi, C., 2013. Integrated assessment of sea-level rise adaptation strategies using a Bayesian decision network approach. *Environmental Modelling & Software* 44 87-100.
- Chee, Y.E., Wilkinson, L., Nicholson, A.E., Quintana-Ascencio, P.F., Fauth, J.E., Hall, D., Ponzio, K.J., Rumpff, L., 2016. Modelling spatial and temporal changes with GIS and Spatial and Dynamic Bayesian Networks. *Environmental Modelling & Software* 82 108-120.
- Cheng, P. W. 1997. From covariation to causation: a causal power theory. *Psychological Review* 104(2):367-405.
- Cheon, S.-P., S. Kim, S.-Y. Lee, and C.-B. Lee. 2009. Bayesian networks based rare event prediction with sensor data. *Knowledge-Based Systems* 22(5):336-343.
- Conrady, S., and L. Jouffe. 2015. *Bayesian networks & BayesiaLab*. Bayesia USA, Franklin, Tennessee. 385 pp.
- Constantinuo, A. C., N. Fenton, W. Marsh, and L. Radlinski. 2016. From complex questionnaire and interviewing data to intelligent Bayesian network models for medical decision support. *Artificial Intelligence in Medicine* 67:75-93.
- Cooke, R. M., D. Kurowicka, A. M. Hanea, O. Morales, D. A. Ababei, B. Ale, and A. Roelen. 2007. Continuous/discrete non parametric Bayesian belief nets with UNICORN and

- UNINET. In: Proceedings of Mathematical Methods in Reliability MMR, 1-4 July 2007. Glasgow, UK.
- Cox, L. A., Jr. 2012. Confronting deep uncertainties in risk analysis. *Risk Analysis* 32(10):1607-1629.
- Delgado-Hernández, D.-J., O. Morales-Nápoles, D. De-León-Escobedo, and J.-C. Arteaga-Arcos. 2012. A continuous Bayesian network for earth dams' risk assessment: an application. *Structure and Infrastructure Engineering* 10(2):225-238.
- de Waal, A., H. Koen, P. de Villiers, H. Roodt, N. Moorosi, and G. Pavlin. 2016. Construction and evaluation of Bayesian networks with expert-defined latent variables. Pp. 774-781 in: 19th International Conference on Information Fusion (FUSION), Heidelberg, Germany.
- Dlamini, W.M., 2010. A Bayesian belief network analysis of factors influencing wildfire occurrence in Swaziland. *Environmental Modelling & Software* 25(2) 199-208.
- Do, C.B., Batzoglou, S., 2008. What is the expectation maximization algorithm? *Nature Biotechnology* 26 897-899.
- Driver, E., and D. Morrell. 1995. Implementation of continuous Bayesian networks using sums of weighted Gaussians. in: P. Besnard and S. Hanks, ed. *Uncertainty in artificial intelligence. Proceedings of the Eleventh Conference (1995)*. Morgan Kaufmann Publishers, San Francisco, CA. 134-140 pp.
- Druzdzel, M. J. 1999. GeNIe: A development environment for graphical decision-analytic models. P. 1206 in: *Proceedings of the 1999 Annual Symposium of the American Medical Informatics Association (AMIA-1999)*, November 6-10, 1999. Washington, D.C.
- Farmani, R., H. J. Henriksen, D. Savic, and D. Butler. 2012. An evolutionary Bayesian belief network methodology for participatory decision making under uncertainty: An application to groundwater management. *Integrated Environmental Assessment and Management* 8(3):456-461.

- Fenton, N., and M. Neil. 2012. Risk assessment and decision analysis with Bayesian networks. CRC Press, Boca Raton, FL. 524 pp.
- Fortin, J. K., K. D. Rode, G. V. Hilderbrand, J. Wilder, S. Farley, C. Jorgensen, and B. G. Marcot. 2016. The impacts of human recreation on brown bears (*Ursus arctos*): a review and new management tool. PLoS ONE 11(1):e0141983. doi:10.1371/journal.pone.0141983.
- Fraser, H., L. Rumpff, J. D. L. Yen, D. Robinson, and B. A. Wintle. 2017. Integrated models to support multiobjective ecological restoration decisions. Conservation Biology 31(6):1418-1427.
- Garvey, M. D., S. Carnovale, and S. Yeniyurt. 2015. An analytical framework for supply network risk propagation: A Bayesian network approach. European Journal of Operations Research 243(2):618-627.
- Giordano, R., D. D'Agostino, C. Apollonio, N. Lamaddalena, and M. Vurro. 2013. Bayesian Belief Network to support conflict analysis for groundwater protection: The case of the Apulia region. Journal of Environmental Management 115:136-146.
- Giudici, P., and A. Spelta. 2016. Graphical network models for international financial flows. Journal of Business & Economic Statistics 34(1):128-138.
- Gonzalez-Redin, J., S. Luque, L. Poggio, R. Smith, and A. Gimona. 2016. Spatial Bayesian belief networks as a planning decision tool for mapping ecosystem services trade-offs on forested landscapes. Environmental Research 144(Part B):15-26.
- Guerriero, M., F. Wheeler, G. Koste, S. Dekate, and N. Choudhury. 2016. Bayesian data fusion for pipeline leak detection. Pp. 278-285 in: 19th International Conference on Information Fusion (FUSION), Heidelberg, Germany.
- Hanea, A., O. M. Napoles, and D. Ababei. 2015. Non-parametric Bayesian networks: improving theory and reviewing applications. Reliability Engineering and System Safety 144:265-284.
- Hanea, A., D. Kurowicka, and R. Cooke. 2006. Hybrid method for quantifying and analyzing Bayesian belief nets. Quality and Reliability Engineering International 22(6):613-729.

- Hanea, A., D. Kurowicka, K. Cooke, and D. Ababei. 2010. Mining and visualising ordinal data with non-parametric continuous BBNs. *Computational Statistics and Data Analysis* 54(3):668-687.
- Hastie, T., R. Tibshirani, and M. Wainwright. 2015. Statistical learning with sparsity: the Lasso and generalizations. *Monographs on statistics and applied probability* 143. CRC Press, Chapman and Hall, 367 pp.
- Havron, A., C. Goldfinger, S. Henkel, B. G. Marcot, C. Romsos, and L. Gilbane. 2017. Mapping marine habitat suitability and uncertainty using Bayesian networks: a case study of northeastern Pacific benthic macrofauna. *Ecosphere* 8(7):e01859. doi: 10.1002/ecs2.1859.
- Hossain, M., Muromachi, Y., 2012. A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. *Accident Analysis & Prevention* 45 373-381.
- Hradsky, B.A., Penman, T.D., Ababei, D., Hanea, A., Ritchie, E.G., York, A., Di Stefano, J., 2017. Bayesian networks elucidate interactions between fire and other drivers of terrestrial fauna distributions. *Ecosphere* 8(8) e01926 doi:10.1002/ecs2.1926
- Janssens, D., G. Wets, T. Brijs, K. Vanhoof, T. Arentze, and H. Timmermans. 2006. Integrating Bayesian networks and decision trees in a sequential rule-based transportation model. *European Journal of Operations Research* 175(1):16-34.
- Johnson, S., F. Fielding, G. Hamilton, and K. Mengersen. 2010. An Integrated Bayesian Network approach to *Lyngbya majuscula* bloom initiation. *Marine Environmental Research* 69:27-37.
- Johnson, S., and K. Mengersen. 2012. Integrated Bayesian network framework for modeling complex ecological issues. *Integrated Environmental Assessment and Management* 8(3):480-490.
- Johnson, S., Marker, L., Mengersen, K., Gordon, C.H., Melzheimer, J., Schmidt-Küntzel, A., Nghikembua, M., Fabiano, E., Henghali, J., Wachter, B., 2013. Modeling the viability of

- the free-ranging cheetah population in Namibia : an object-oriented Bayesian network approach. *Ecosphere* 4(7) doi: 10.1890/ES1812-00357.00351.
- Jorgenson, M. T., B. G. Marcot, D. K. Swanson, J. C. Jorgenson, and A. R. DeGange. 2015. Projected changes in diverse ecosystems from climate warming and biophysical drivers in northwest Alaska. *Climatic Change* 130(2):131-144.
- Kery, M. 2010. Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses. Academic Press, 320 pp.
- Kim, S. Y., S. Imoto, and S. Miyano. 2003. Inferring gene networks from time series microarray data using dynamic Bayesian networks. *Briefings in Bioinformatics* 4(3):228-235.
- Kininmonth, S., Lemm, S., Malone, C., Hatley, T., 2014. Spatial vulnerability assessment of anchor damage within the Great Barrier Reef World Heritage Area, Australia. *Ocean & Coastal Management* 100 20-31.
- Kleemann, J., E. Celio, and C. Fürst. 2017. Validation approaches of an expert-based Bayesian Belief Network in Northern Ghana, West Africa. *Ecological Modelling* 365:10-29.
- Koen, H., J. P. de Villiers, H. Roodt, and A. De Waal. 2017. An expert-driven causal model of the rhino poaching problem. *Ecological Modelling* 347:29-39.
- Koski, T., and J. Noble. 2011. Bayesian networks: An introduction. Wiley, 366 pp.
- Kosmala, M., A. Wiggins, A. Swanson, and B. Simmons. 2016. Assessing data quality in citizen science. *Frontiers in Ecology and Evolution* 14(10):551-560.
- Krüger, C., Lakes, T., 2015. Bayesian belief networks as a versatile method for assessing uncertainty in land-change modeling. *International Journal of Geographical Information Science* 29(1) 111-131.
- Kuikka, S., N. Hildén, H. Gislason, S. Hansson, H. Sparholt, and O. Varis. 1999. Modeling environmentally driven uncertainties in Baltic cod (*Gadus morhua*) management by Bayesian influence diagrams. *Canadian Journal of Fisheries and Aquatic Sciences* 56:629-641.
- LaDeau, S. L., B. A. Han, E. J. Rosi-Marshall, and K. C. Weathers. 2017. The next decade of big data in ecosystem science. *Ecosystems* 20(2):274-283.

- Landuyt, D., S. Broekx, R. D'hondt, G. Engelen, J. Aertsens, and P. L. M. Goethals. 2013. A review of Bayesian belief networks in ecosystem service modelling. *Environmental Modelling & Software* 46:1-11.
- Landuyt, D., K. Van der Biest, S. Broekx, J. Staes, P. Meire, and P. L. M. Goethals. 2015. A GIS plug-in for Bayesian belief networks: Towards a transparent software framework to assess and visualise uncertainties in ecosystem service mapping. *Environmental Modelling & Software* 71:30-38.
- Lauret, P., Fock, E., Randrianarivonyh, R.N., Manicom-Ramsamy, J.-F., 2008. Bayesian neural network approach to short time load forecasting. *Energy Conversion and Management* 49(5) 1156-1166.
- Lauritzen, S. L. 1995. The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis* 19:191-201.
- Leifer, M. S., and R. W. Spekkens. 2013. Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference. *Physical Review A* 88(5):<https://doi.org/10.1103/PhysRevA.88.052130>.
- Lewis, K. P., E. Vander Wal, and D. A. Fifield. 2018. Wildlife biology, big data, and reproducible research. *Wildlife Society Bulletin* 42(1):172-179.
- Li, C., and S. Mahadevan. 2017. Sensitivity analysis of a Bayesian network. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering* 4(1):doi: 10.1115/1.4037454.
- Li, M., A. Stein, W. Bijker, and Q. Zhan. 2016. Urban land use extraction from Very High Resolution remote sensing imagery using a Bayesian network. *ISPRS Journal of Photogrammetry and Remote Sensing* 122:192-205.
- Li, X., Y. Zhang, F. Guo, X. Gao, and Y. Wang. 2018. Predicting the effect of land use and climate change on stream macroinvertebrates based on the linkage between structural equation modeling and Bayesian network. *Ecological Indicators* 85:820-831.
- Loyd, K. A. T., and J. L. DeVore. 2010. An evaluation of feral cat management options using a decision analysis network. *Ecology and Society* 15(4):10. [online] URL: <http://www.ecologyandsociety.org/vol15/iss4/art10/>.



- Lu, H., A. L. Yuille, M. Liljeholm, P. W. Cheng, and K. J. Holyoak. 2008. Bayesian generic priors for causal learning. *Psychological Review* 115(3):955.
- Lu, Y., Y. Duan, W. Kang, Z. Li, and F.-Y. Wang. 2014. Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems* 16(2): DOI: 10.1109/TITS.2014.2345663
- Madsen, A. L., M. Lang, U. B. Kjærulff, and F. Jensen. 2003. The Hugin tool for learning Bayesian networks. Pp. 594-605 in: T. D. Nielsen and N. L. Zhang, editors. *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*. 7th European Conference, ECSQARU 2003 Aalborg, Denmark, July 2-5, 2003.
- Maglogiannis, I., E. Zafiropoulos, A. Platis, and C. Lambrinoudakis. 2006. Risk analysis of a patient monitoring system using Bayesian Network modeling. *Journal of Biomedical Informatics* 39:637-647.
- Marcot, B. G. 2017. Common quandaries and their practical solutions in Bayesian network modeling. *Ecological Modelling* 358:1-9.
- Mason, T. J., D. A. Keith, and A. D. Letten. 2017. Detecting state changes for ecosystem conservation with long-term monitoring of species composition. *Ecological Applications* 27(2):458-468.
- Moreira, C., and A. Wichert. 2018. Are quantum-like Bayesian networks more powerful than classical Bayesian networks? *Journal of Mathematical Psychology* 82:73-83.
- Murray, R.C., VanLehn, K., Mostow, J., 2004. Looking ahead to select tutorial actions: A decision-theoretic approach. *International Journal of Artificial Intelligence in Education* 14(3-4) 235-278.
- Neil, M., M. Tailor, and D. Marquez. 2007. Inference in hybrid Bayesian networks using dynamic discretization. *Statistics and Computing* 17(3):219-233.
- Nicholson, A.E., Flores, M.J., 2011. Combining state and transition models with dynamic Bayesian networks. *Ecological Modelling* 222(3) 555-566.

- Nielsen, S. H., and S. Parsons. 2007. An application of formal argumentation: Fusing Bayesian networks in multi-agent systems. *Artificial Intelligence* 171(10-15):754-775.
- Norick, L. R., and P. W. Cheng. 2004. Assessing interactive causal influence. *Psychological Review* 111(2):455-485.
- O'Donoghue, M., Boutin, S., Krebs, C.J., Hofer, E.J., 1997. Numerical responses of coyotes and lynx to the snowshoe hare cycle. *Oikos* 80(1) 150-162.
- Orphanou, K., A. Stassopoulou, and E. Keravnuo. 2014. Temporal abstraction and temporal Bayesian networks in clinical domains: A survey. *Artificial Intelligence in Medicine* 60:133-149.
- Park, S., and D. V. Budescu. 2015. Aggregating multiple probability intervals to improve calibration. *Judgment and Decision Making* 10(2):130-143.
- Pearl, J. 1998. Graphs, causality, and structural equation models. *Sociological Methods & Research* 27(2):226-284.
- Pearl, J. 2000. *Causality: models, reasoning, and inference*. Cambridge University Press, Cambridge. 384 pp.
- Penman, T.D., Nicholson, A.E., Bradstock, R.A., Collins, L., Penman, S.H., Price, O.F., 2015a. Reducing the risk of house loss due to wildfires. *Environmental Modelling & Software* 67 12-25.
- Penman, T. D., K. A. Parkins, S. Mascaro, D. Chong, and R. A. Bradstock. 2015b. National fire danger rating system probabilistic framework project. *Bushfire and Natural Hazards, Cooperative Research Centres Programme, Australia*. 27 pp.
- Phan, T. D., J. C. R. Smart, S. J. Capon, W. L. Hadwen, and O. Sahin. 2016. Applications of Bayesian belief networks in water resource management: A systematic review. *Environmental Modelling & Software* 85:98-111.
- Pitchforth, J., and K. Mengersen. 2013. A proposed validation framework for expert elicited Bayesian Networks. *Expert Systems with Applications* 40(1):162-167.

- Raphael, M. G., M. J. Wisdom, M. M. Rowland, R. S. Holthausen, B. C. Wales, B. G. Marcot, and T. D. Rich. 2001. Status and trends of habitats of terrestrial vertebrates in relation to land management in the interior Columbia River Basin. *Forest Ecology and Management* 153(1-3):63-87.
- Reynolds, K. M., P. F. Hessburg, and P. S. Bourgeron, ed. 2014. Making transparent environmental management decisions. Springer, Berlin Heidelberg. 337 pp.
- Salmond, J. A., M. Tadaki, and M. Dickson. 2017. Can big data tame a “naughty” world? *The Canadian Geographer* 61(1):52-63.
- Sperotto, A., Molina, J.-L., Torresan, S., Critto, A., Marcomini, A., 2017. Reviewing Bayesian Networks potentials for climate change impacts assessment and management: A multi-risk perspective. *Journal of Environmental Management* 202 320-331.
- Spiegelhalter, D. J. 2014. The future lies in uncertainty. *Science* 345(6194):264-265.
- Steventon, J.D., Sutherland, G.D., Arcese, P., 2006. A population-viability based risk assessment of Marbled Murrelet nesting habitat policy in British Columbia. *Canadian Journal of Forest Research* 36 3075-3086.
- Sun, Z., Müller, D., 2013. A framework for modeling payments for ecosystem services with agent-based models, Bayesian belief networks and opinion dynamics models. *Environmental Modelling & Software* 45 15-28.
- Trueblood, J. S., P. K. Mistry, and E. M. Pothos. 2016. A quantum Bayes net approach to causal reasoning. Pp.449-464 in: E. Dzhamalov, S. Jordan, R. Zhang, and V. Cervantes, editors. *Advanced series on mathematical psychology. Volume 6. Contextuality from quantum physics to psychology.* World Scientific.
- Tsamardinos, I., Brown, L.E., Aliferis, C.F., 2006. The max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning* 65(1) 31-78.
- Tucci, R. R. 1995. Quantum Bayesian nets. *International Journal of Modern Physics B* 9(3):295-337.
- Uusitalo, L., M. T. Tomczak, B. Müller-Karulis, I. Putnis, and N. Trifonova. 2018. Hidden variables in a Dynamic Bayesian Network identify ecosystem level change. *Ecological Informatics* 45:9-15.

- Vagnoli, M., R. Remenyte-Priscott, and J. Andrews. 2017. Towards a real-time structural health monitoring of railway bridges. Pp. 208-218 in: I. Žutautaitė, M. Eid, K. Simola, and V. Kopustinskias, editors. Proceedings of the 52nd ESReDA Seminar on Critical Infrastructures: Enhancing Preparedness & Resilience for the Security of Citizens and Services Supply Continuity, 30-31 May 2017. Kaunas, Lithuania.
- Van Allen, T., A. Singh, R. Greiner, and P. Hooper. 2008. Quantifying the uncertainty of a belief net response: Bayesian error-bars for belief net inference. *Artificial Intelligence* 172(4-5):483-513.
- Vlasselaer, J., W. Meert, G. Van den Broeck, and L. De Raedt. 2016. Exploiting local and repeated structure in Dynamic Bayesian Networks. *Artificial Intelligence* 232:43-53.
- Wilson, D.S., Stoddard, M.A., Puettmann, K.J., 2008. Monitoring amphibian populations with incomplete survey information using a Bayesian probabilistic model. *Ecological Modelling* 214(2-4) 210-218.
- Yacef, R., Benghanem, M., Mellit, A., 2012. Prediction of daily global solar irradiation data using Bayesian neural network: A comparative study. *Renewable Energy* 48 146-154.
- Zohar, A., and J. S. Rosenschein. 2008. Mechanisms for information elicitation. *Artificial Intelligence* 172(16-17):1917-1939.

Table 1. Recent advances in integration of Bayesian network (BN) modeling with other modeling constructs. Acronyms listed here are as used in the literature (see text).

Recent integration advances	Purpose
Geographic information systems Bayesian networks (BN-GIS)	Map geographically-referenced posterior probabilities generated from the BN
Dynamic Bayesian networks (DBNs)	Replicate a BN structure over simulated time period to incorporate time-dynamic feedback loops and lag effects
Bayesian decision networks (BDNs)	Determine potential effects and expected values of alternative management decisions in a probabilistic framework
Dynamic decision networks (DDNs)	Evaluate effects and expected values of a sequence of management decisions
Structural equation modeling (SEM) Bayesian networks	Use SEM to determine appropriate causal network structures for a BN
Bayesian neural networks	Use BNs to determine neural network node weights
Continuous-variable Bayesian networks	Avoid simplification of ratio-scale data into discretized range states
Hybrid Bayesian networks (HBNs)	BNs containing both discrete and continuous (non-discretized) variables; include non-parametric BNs (NPBNs)
Object-oriented Bayesian networks (OOBNs)	Treat BN variables as "objects" that can combine methods and data structures
Dynamic object-oriented Bayesian networks (DOOBNs)	Conduct OOBNs in a dynamic simulation where object parameters can vary over simulated time
Agent-based Bayesian networks	Treat BN variables as agents or individual entities with dynamic interactions with their environment
State-and-transition Bayesian networks (STM-BNs)	Project changes in amounts and dispersions of conditions under probability distributions
Quantum Bayesian networks (QBNs)	Model non-classical probability outcomes using quantum probability amplitude functions

Power PC theory in causal BNs	Model causal probability structures with observed and unobserved influences, partitioning out independent causes as additive effects
-------------------------------	--

Figure 1. Examples of various Bayesian network (BN) modeling objectives, and associated categories of recent advances in BN model applications and integration (see Table 1 for abbreviations).

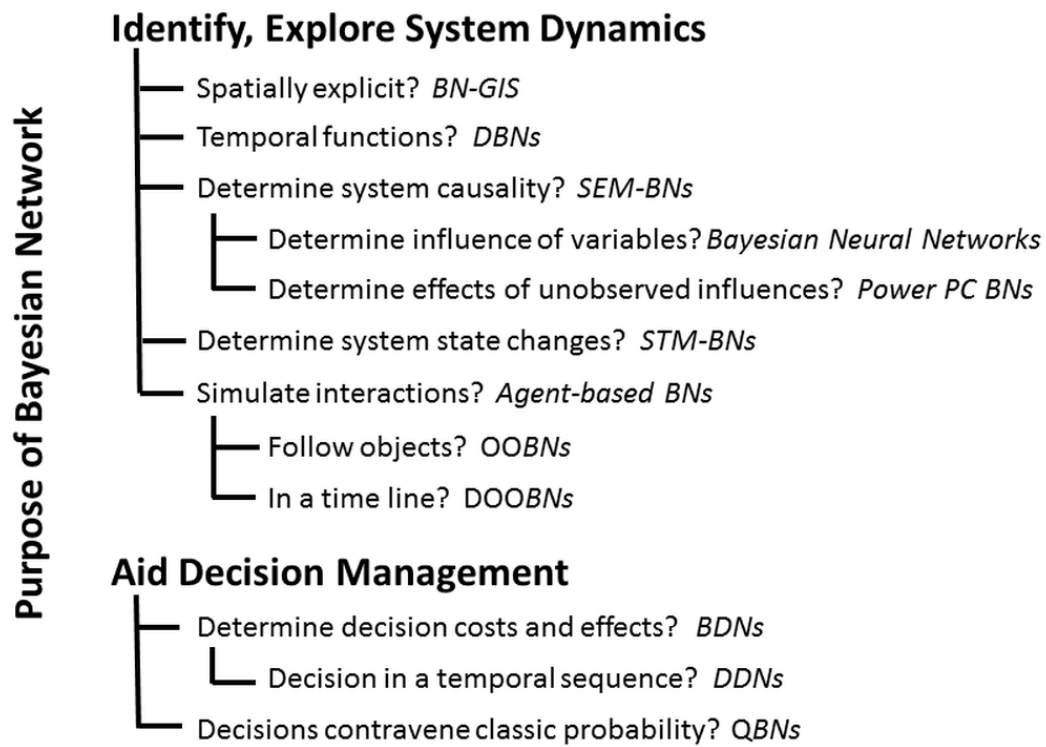
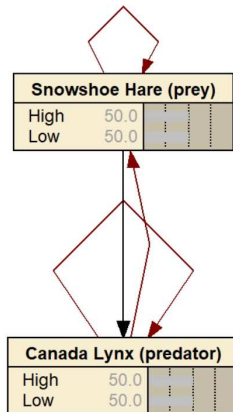
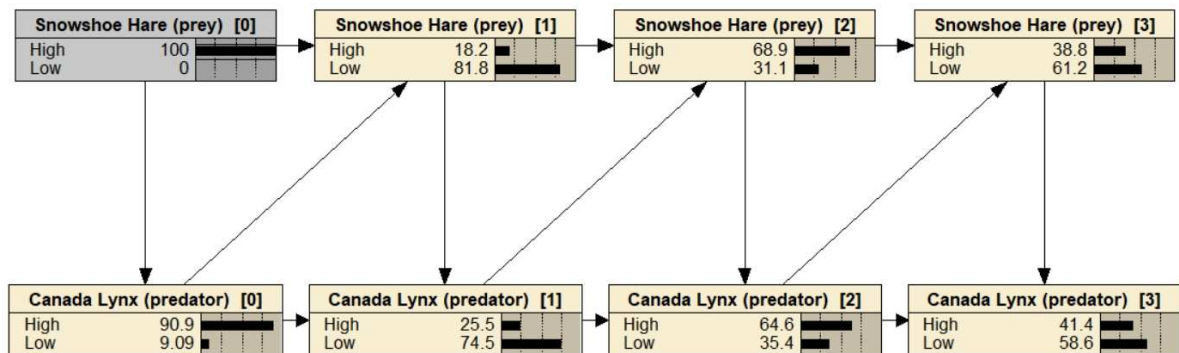


Figure 2. Example of a dynamic Bayesian network of a dampened oscillating predator-prey system of Canadian lynx (*Lynx canadensis*) and snowshoe hare (*Lepus americanus*), demonstrating (a) the collapsed model with feedback loops and (b) the model expanded to 4 time intervals showing population interactions (based on O'Donoghue et al., 1997).

(a)



(b)





## Advances in Bayesian network modelling: Integration of modelling technologies

Bruce G. Marcot and Trent D. Penman

### Highlights

- Bayesian network (BN) modeling has become a popular tool in ecological science and environmental management.
- Many innovations and variations on the basic probability network structure have been devised in recent years, including improvements in Bayesian classifiers and machine-learning algorithms for model structuring and parameterization, and development of time-dynamic and decision models.
- We review the wide array of advances in BN modeling, particularly how they are being used and integrated in management decision networks; structural equation modeling of causal networks; Bayesian neural networks; combined discrete and continuous variables; object-oriented and agent-based models; state-and-transition models; geographic information systems; quantum probability; and other fields.
- We close with cautions regarding validity and credibility of self-structuring and self-learning BN models developed from uncertain sources of knowledge systems and big data.