Quantifying model structure uncertainty in a Bayesian network for aquatic plants using a goodness-of-fit test

William Barnett[1], Kathryn M. Irvine[1*], Richard S. Sojda[2]

[1] *Montana State University, Department of Mathematical Sciences,*

PO Box 172400 Bozeman, MT 59717

[2] *U S Geological Survey, Northern Rocky Mountain Science Center,*

2327 University Way, Suite 2, Bozeman, MT 59715

[*] *Corresponding author at,*
Address: U S Geological Survey, Northern Rocky Mountain Science Center, 2327 University Way, Suite 2, Bozeman, MT 59715
Tel.: 011 406 994 7492;
E-mail address: kirvine@usgs.gov.

Abstract

Bayesian networks are commonly used by researchers and wildlife managers as decision support tools. Driven primarily by conceptual models and expert knowledge, simplifications to the linkage structure are often preferred to a more complex, but perhaps better fitting, model. We demonstrate a method for evaluating the goodness-of-fit of a Bayesian network linkage structure using path analysis techniques. Currently, there is not an established quantitative method available to assess the fit of a Bayesian network's dependency structure for a given ecological system. For our application, the goal is to develop a Bayesian network to inform managers at the Red Rock Lakes National Wildlife Refuge on the potential benefits of artificial water level manipulation. Using our methodology, we investigate the association between water level and northern water milfoil and sago pondweed abundance, as well as other soil characteristics. We find that high water levels are associated with more northern water milfoil and less sago pondweed. In terms of managing for waterfowl within the refuge, the manipulation of water levels should be based on which aquatic plant species is preferred forage of key water birds.

Keywords:

Decision support tool, conceptual model, graphical models, *Myriophyllum exalbescens*, path analysis, *Stuckenia pectinatus*

Prepared for Submission to Ecological Modelling

1. Introduction

Conceptual models are graphical illustrations that depict ecological processes in a qualitative fashion (Barrows and Allen, 2007; King and Hobbs, 2006). Essentially box-and-arrow diagrams that show linkages between variables of importance, such models are used to develop a consensus among researchers and managers on an ecosystem's function (e.g., Ogden et al., 2005). They are useful as decision support tools in the absence of data, since they can be driven entirely by expert knowledge (Rudnick et al., 2005). Often a conceptual model is used as the starting point for developing a more quantitative model such as a Bayesian network (BN). A BN allows for quantification of key ecosystem relationships and evaluation of different management scenarios (e.g., Dlamini, 2010; Peterson, 2008; Mead et al., 2006).

The use of BNs to describe ecological systems has increased in the past few years (Aalders, 2008; Johnson et al., 2010; Renken et al., 2009), as researchers realize the utility of such a modeling approach. Advantages of BNs include a user friendly decision support tool, predicting ecological responses to system disturbances, and the incorporation of expert knowledge in the absence of data. Simple model structures are often chosen over more complex ones, due to the heavy reliance on expert knowledge for BN model development and the fact that conceptualizing conditional probabilities is difficult, especially as the number of conditioning variables increases (Uusitalo, 2007). This is a potential drawback with the use of BNs to describe ecosystem relationships; a simplified structure could miss key relationships or pathways. Therefore, providing a measure of fit for a given model’s linkage structure is a useful additional piece of information to incorporate into the decision making process.

Frequently, ecological papers using BNs follow guidelines suggested in Marcot et al. (2006) for developing and testing the model (e.g., Johnson et al., 2010; Peterson et al., 2008; Smith et al., 2007); currently, little attention is placed on verifying the hypothesized linkage structure or “causal web” of environmental variables driving species responses. Marcot et al. (2006) suggests an extensive peer review by experts and frequent updating of the conditional probability tables (CPTs) as more data become available. The prediction accuracy of a BN is sometimes evaluated using receiver operator characteristic (ROC) curves (Marcot et al., 2006). Another suggested method for refining a BN is sensitivity analysis (Marcot et al., 2006). This process involves systematically changing the conditional probabilities for each variable by a small amount, and measuring the effect of that change on the response or terminal node. If the observed sensitivities of the BN do not yield an expected result, the model can be parameterized again or even restructured. However, a model with an incorrect linkage structure may still behave as expected at the terminal node or produce accurate predictions (for example, see burglary alarm network p. 494 in Russell and Norvig, 2002). As BNs are often used to explain ecosystem interrelationships, the linkage structure of the model should be validated as well.

Here we present an analytical approach that allows researchers to build a BN whose linkage structure can be verified by a traditional goodness-of-fit (GOF) test, provided that some data are available. We use recently developed theory from statistical graphical models (Shipley, 2009) to first assess the GOF for a proposed graph structure, and then we convert the result to a Bayesian network. Using data from Red Rock Lakes National Wildlife Refuge, we develop and verify the linkage structure of a BN which can be used to evaluate the potential effects of water level manipulation on certain species of submergent vegetation.

2. Case Study

Red Rock Lakes National Wildlife Refuge is a large montane wetland complex along the Idaho-Montana border that provides crucial habitat for trumpeter swans and other important waterbirds. A key food source for some species is submergent vegetation in the wetland system, of which *Stuckenia pectinatus* (sago pondweed) seems particularly important. Its abundance appears to be inversely correlated with that of *Myriophyllum exalbescens* (northern water milfoil) (Anderson, 1978). Recently, however, certain areas of the wetland have shifted in abundance to other species of submergent vegetation, especially northern water milfoil. In the interest of managing critical habitat for trumpeter swans and other waterbirds, research was initiated in 2003 at Red Rock Lakes NWR to quantify the abundance of various submergent vegetation species within the wetland system, and evaluate the effects of varying water levels on the abundance of the species in question. Our goal was to develop a Bayesian network for the refuge managers to use as a decision support tool in evaluating the potential impacts of water level manipulations on submergent plant abundance.

From 2003-2007, abundance estimates of various submergent vegetation species were collected at 32 randomly chosen open water and pond sites throughout the wildlife refuge (one site was omitted in the final analysis due to missing data). Each site was divided into multiple 1 m^2 plots (usually 10, 15 or 25, depending on the site) and vegetation cover measurements were recorded every year at each plot according to a modified Daubenmire scale ("0" was included as a class; Daubenmire, 1959). Relevant environmental variables were also recorded at varying time scales. Soil characteristics were recorded at each site in 2003 and were assumed to be constant throughout the study. These variables included calcium (Ca), calcium carbonate ($CaCO_3$), sodium (Na), magnesium (Mg), inorganic carbon (IC), organic carbon (OC), electrical conductivity (EC), cation exchange capacity (CEC), soil texture type, and substrate composition (percent sand/silt/clay) (Table 1). Water levels (surface depths) were also recorded each year at varying time points from late May to mid-September. These measurements were taken by hand for the first two years, but became automated in 2005 after piezometers and wells with electronic loggers were installed.

Table 1 Abbreviations for variables recorded

Variable	Abbreviation
Calcium	Ca
Calcium Carbonate	$CaCO_3$
Sodium	Na
Magnesium	Mg
Inorganic Carbon	IC
Organic Carbon	OC
Electrical Conductivity	EC
Cation Exchange Capacity	CEC
Percent Sand in Soil	PerSand
Percent Clay in Soil	PerClay

A general conceptual model of the wetland provided a basic outline of the hypothesized relationships between different components of the ecosystem (Fig. 1). Soil characteristics and water level are both thought to be important drivers of vegetation abundance. Climate variables likely affect submergent vegetation through temperature (evaporation) and wind (water turbidity), but due to a lack of long term data, no climate effects were considered in this analysis. We did not have an *a priori* causal structure involving all of the specific measured variables in Table 1. However, we did believe, based on the literature review of Kantrud (1990), that the previous year's water level was associated with the current year's vegetation abundance. Sago pondweed and northern water milfoil both develop winter buds or turions during late summer for the next year's growth, and water levels are thought to have an effect on these biological mechanisms.

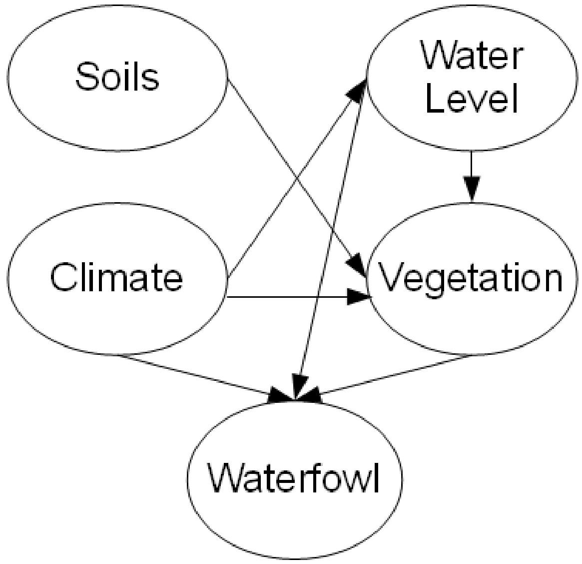

Fig. 1. Conceptual model of wetland ecosystem

We used expert opinion and a review of available literature to determine which of the site level covariates to include in a preliminary model. Kantrud (1990) suggests that calcium (Ca) and inorganic carbon (IC) are important factors in sago pondweed growth, while soil composition plays a less distinct role. On the idea that nutrient availability is also important in plant establishment and growth, cation exchange capacity (CEC) and electrical conductivity (EC) were also included as potential variables. Water turbidity seems to also negatively affect sago pondweed production (Kantrud 1990), but we did not have reliable estimates of turbidity across all five years due to equipment failure. Much less is known about the nutritional requirements of northern water milfoil, so we included the same covariates for this response variable (terminal node). The arbitrary time steps at which water levels were measured prevented us from calculating weekly averages for each site. Instead, for every year we calculated monthly water level averages for each of the 31 sites. We selected July of each year as an indicator of the season's water level.

3. Statistical Methods

3.1 Introduction to graphical models and Bayesian networks

A graphical model is an illustrative way to represent processes in terms of measurable data and possibly latent (un-measured) factors. Formally, a graphical model is defined by two sets **V** and **E**,

1. The set **V** represents the random variables (also called nodes or vertices) whose probability distributions can be discrete or continuous, and

2. A set **E** represents the arrows (also called directed edges) between variables that represent pairwise relationships. A unidirectional arrow denotes causation while a bi-directional arrow denotes correlation (i.e., undirected edge).

In this paper, we use the terminology linkage structure in reference to the edge set (**E**) of a graph.

Together, these two sets allow one to construct a box-and-arrow diagram that depicts the relationships among variables in a system. Bayesian networks (Jensen, 2001; Korb and Nicholson, 2004; Russell and Norvig, 2002) comprise a subset of graphical models in which the model is both directed and acyclic (meaning there are no feedback loops and only directed edges). In addition, Bayesian networks are typically defined as having discrete or categorical variables; path analysis and structural equation models have no such requirement. This latter constraint in BNs implies that the conditional probability distribution of a node is expressed as a conditional probability table (Jensen, 2001). We restrict ourselves to directed acyclic graphs (DAGs) for the rest of the paper.

One of the most important concepts in graphical model theory is d-separation (the “d” stands for dependency), which allows one to write the joint probability distribution for the graph in a much simpler manner. It is a formalization of a few common sense rules about how information can or cannot pass between variables in a causal system, and it represents the independencies that result from this flow of information. Several texts provide in-depth explanations of d-separation, for those interested in the details (see p. 10 Jensen, 2001 for BN and p. 29 in Shipley, 2002 for path models). The concept of d-separation is fundamental to graphical models, and hence to BNs as well. Assuming the causal relationships portrayed in a graph are true, d-separation implies that if two nodes are d-separated in a graph, the corresponding random variables are statistically conditionally independent. The observed data are just one realization of the true causal process which is unknown, but of interest. Essentially, the inferential goal is to determine whether observed data are consistent with being generated by the causal process depicted in the hypothesized graph.

3.2 Goodness-of-fit test

Shipley (2002) proposes a method for testing model implied d-separation statements for continuous variables and expands these ideas for both a mixture of discrete and continuous variables in Shipley (2009). The main conclusions are that one may form a "basis set" of d-separation statements (i.e., $A \perp\!\!\!\perp B \mid \{C\}$) implied by the proposed graph, and then one can test these implied statistical conditional independencies by using p-values for the corresponding partial regression coefficients in an appropriate statistical model (e.g., multiple linear regression, logistic regression, mixed model).

Summarizing from Shipley (2002, 2009), for a hypothesized linkage structure, the basis set of d-separation statements implied by the graph is the set of all pairs of variables without an edge between them, conditioned on the direct causes of either node in the pair (known as the parent nodes). Each independence relationship in the basis set can be tested statistically through the use of an appropriate statistical model. Regressing one variable on another and including the parents of each node as explanatory variables, the p-value of the partial regression coefficient for one of the two "independent" variables gives the evidence for the relationship between them. Assuming the proposed graph structure is correct (the null hypothesis), the p-values corresponding to the basis set of independence claims are mutually independent. Fisher (1932) provided the original statistical framework to combine multiple independent p-values. Shipley (2002) used Fisher's work to develop a goodness-of-fit test for a proposed causal graph. The goodness-of-fit statistic is $C = -2\sum_{i=1}^{k}\log(p_i)$, where p_i is the p-value associated with the i^{th} conditional independency claim in the basis set. Assuming the causal graph is correct, this statistic follows a Chi-square distribution with 2k degrees of freedom. A large p-value may be interpreted as evidence that the data are consistent with actually being generated by the proposed causal model. A small p-value below an accepted threshold, results in the causal model being rejected. If the causal model is rejected, there is no way to know if the causal model is incorrect or if the data are just one of the unlikely realizations possible, but unlikely, to be observed from the assumed causal process depicted in the graph.

Using the Red Rock Lakes NWR dataset, we develop a preliminary causal graph using previously described scientific knowledge of sago pondweed and northern milfoil and the search algorithms available in TETRAD. Then we use the GOF test to provide a measure of uncertainty of the proposed linkage structure. We then convert this graph into a BN to use as a decision support tool for wetland managers.

4. Results and Discussion

4.1 Model structure

As described previously, we reduced the set of possible soil variables to include in our graph structure based on expert opinion and the available literature. We then used the available search algorithms in the TETRAD IV program (Spirtes et al., TETRAD Project) to suggest the linkage structure among the soil covariates (Fig. 2). This program searches for linkage structures that have implied conditional independencies consistent with those in the dataset and then orients the directions of the arrows as a last step. One should be aware that multiple graphs can be found with equivalent conditional independences. There are a few search algorithms available in the program, the PC, FCI and GES algorithms. We tried both the GES (Meek, 1997) and FCI (Spirtes, 2001) search algorithms. The GES algorithm starts with all pairwise edges included in the model and pares them down individually according to a scoring function, while the FCI algorithm starts without any edges and adds them sequentially (similar to the PC algorithm). The FCI algorithm produced essentially empty DAGs without any useful edges, likely due to our small sample size, so we used the GES algorithm instead.

A complication we encountered when trying to determine the relationships among soils measurements and plant abundance was the search algorithms available in TETRAD require either continuous or discrete random variables, but our data was a mixture of continuous and ordinal data. Northern water milfoil and sago pondweed were measured in the field as visually estimated cover classes (i.e., an integer value corresponding to a cover class was recorded for each plot within a site), but water level and the soils measurements were roughly continuous variables. Our solution was to calculate for each site, the proportion of plots (within a site) greater than or equal to the third Daubenmire class for each species. The third Daubenmire class corresponds to a plot having 50 percent or more plant cover. We then used a logit-transformation to convert the proportions (defined between 0 and 1) to a roughly continuous variable (defined between negative and positive infinity) in order to use the TETRAD search algorithm.

We used the GES algorithm in TETRAD on yearly subsets of the aquatic plant data to determine which soil variables were associated with sago pondweed and/or northern water milfoil abundance each year. That is, we always included the same set of soil variables PerClay, Ca, CEC, EC, IC, and PerSand [see Table 1] in each of the five yearly searches (2003, 2004, 2005, 2006, and 2007). The only difference amongst the searches was the sago pondweed and water milfoil abundance measurements. For sago pondweed, TETRAD's GES algorithm suggested that percent clay, CEC, calcium level, and inorganic carbon levels were associated with vegetation abundance at least once during the five year span. In the case of water milfoil, percent clay, CEC, and calcium were associated with abundance in at least one year. For clarification, see Table 2.

Table 2 Number of years, out of five (2003-2007), that each covariate was detected by TETRAD as a "cause" of vegetation abundance.

	Sago	Milfoil
Ca	1	2
IC	2	0
CEC	2	1
EC	0	0
PerClay	1	1
PerSand	0	0

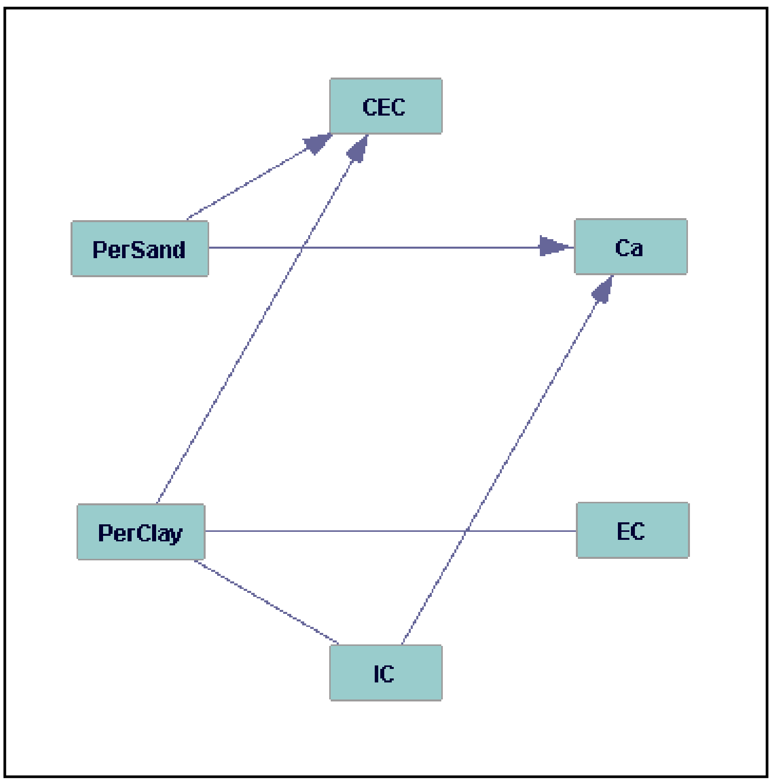

Fig. 2. Soil covariate dependency structure, as determined by TETRAD's GES search algorithm. PerSand and PerClay were not associated with each other due to the particular soil types present in the wildlife refuge.

By combining the soil covariate structure with the variables that were found to be associated with vegetation abundance and the *a priori* belief about water level effects, we hypothesized the complete dependency structure in Fig. 3. In this graph, we assume the relationships between the previous year's and current year's water level and vegetation abundance do not vary among years, an assumption of stationarity. We provide the point estimates and standard errors for the edges in Fig. 3 in Table 3. These estimates are based on similar procedures as described in Section 4.3 for the p-values. One consideration is that the assumption of stationarity in the water-vegetation dynamics might not be valid for future years if climatic shifts cause the underlying relationship to change. However, our model allows for explanation to wetland managers about the manipulations of water levels and their effects on submergent plant abundance.

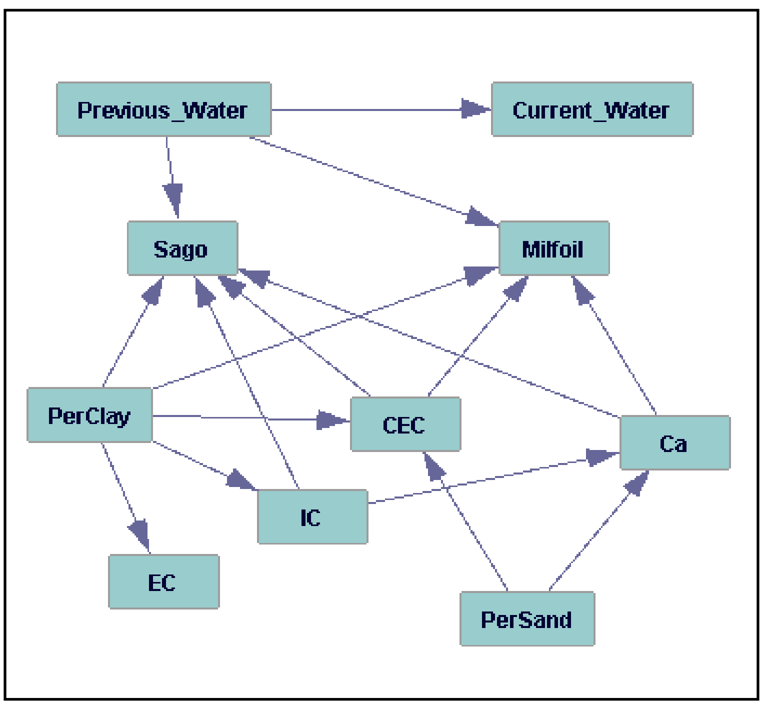

Fig. 3. Hypothesized dependency structure from TETRAD and expert knowledge. Abbreviations in Table 1.

Table 3 Edge coefficient estimates and associated standard errors for path analysis diagram

Path Analysis Diagram Edges		
Edge	Coefficient	SE
PrevWater → Milfoil	0.5840	0.2548
Ca → Milfoil	1.9511	0.5162
CEC → Milfoil	-0.2139	0.3083
PerClay → Milfoil	0.3545	0.2619
PrevWater → Sago	-2.9526	0.6861
Ca → Sago	-2.4456	0.6143
CEC → Sago	-0.8848	0.3180
IC → Sago	1.9544	0.6209
PerClay → Sago	-2.3292	0.6862
PerClay → EC	-0.4562	0.1651
PerClay → IC	0.6156	0.1460
PerClay → CEC	0.4271	0.1547
PerSand → CEC	0.3813	0.1549
PrevWater → CurrWater	0.9576	0.0313
IC → Ca	0.5457	0.1415
PerSand → Ca	0.3333	0.1415

4.2 D-separation statements

According to Shipley (2002), the basis set of d-separation statements is all the pairs of variables without edges between them, conditioned on the parent nodes of the two variables in the pair. Using the complete hypothesized dependency structure (Fig. 3), there were seventeen such statements that comprised the basis set for the hypothesized BN (Table 4). We then tested whether our data were consistent with these dependency statements using regression techniques appropriate for multilevel data, here multiple time scales, as suggested in Shipley (2009).

Table 4 Basis set of d-separation statements for proposed dependency structure (Fig. 4) and the associated p-values.

D-Separation Statement	P-Value for Conditional Independence Test
PerSand, PerClay	1.000
PerSand, EC | PerClay	0.447
PerSand, IC | PerClay	0.501
CEC, Ca | PerSand	0.103
CEC, EC | PerClay	0.840
CEC, IC | PerClay	0.489
Ca, EC | PerSand, IC, PerClay	0.604
Ca, PerClay | IC, PerSand	0.373
EC, IC | PerClay	0.784
Sago, Milfoil | Water-Prev, Ca, CEC, IC, PerClay	0.322
Sago, Water-Curr | Water-Prev, Ca, CEC, IC, PerClay	0.170
Milfoil, Water-Curr | Water-Prev, Ca, CEC, PerClay	0.797
Milfoil, IC | Water-Prev, Ca, CEC, PerClay	0.682
Milfoil, EC | Water-Prev, Ca, CEC, PerClay	0.068
Milfoil, PerSand | Water-Prev, Ca, CEC, PerClay	0.351
Sago, EC | Water-Prev, Ca, IC, CEC, PerClay	0.093
Sago, PerSand | Water-Prev, Ca, IC, CEC, PerClay	0.574

4.3 Conditional independencies and GOF test

To produce the GOF statistic, we calculated a p-value for the statistical conditional independence implied by each d-separation statement in Table 4. We used three different approaches to account for the temporal misalignment of the soils data, measured only once, with the water level and vegetation data, measured yearly. For the first nine statements in Table 4 (rows 1-9) referring to statistical independences between soils variables there were no temporal mismatches, since soil covariates were measured once and during the same time period. The independences were tested with multiple linear regression models, we used the lm() function in R to regress one variable of each pair onto the second variable and the conditioning set.

For the next three conditional independence relationships (Table 4 rows 10-12), the two variables in question were recorded every year (e.g., Sago and Milfoil), but the conditioning set included both soil covariates, measured only once, and previous year's water level, measured each year. Since all three statements included sago pondweed or northern water milfoil, we used quasi-likelihood estimation of the parameters in a logistic regression model for binomial counts since there was evidence of over-dispersion in the plant abundance variables. This is not surprising given that plots were nested within sites. To deal with the fact that soils were only measured once, we used an augmented dataset that included the soil variables repeated 5 times (31*5 rows for each soil variable) to match the data gathered on a yearly basis.

For the final five conditional independence statements (Table 4 rows 13-17), the two variables in question were measured at different temporal scales requiring a multi-step modeling approach. For instance, consider the test of independence between northern water milfoil, which varied each year, and inorganic carbon (IC), measured only once, in row 13 of Table 4. First, we fit a generalized mixed model with northern water milfoil as the response variable, the previous year's water level as the explanatory variable, and a random intercept for each site. Then, we fit a linear regression model with the random site effect estimates as the response variable, the soil covariate in question as an explanatory variable, and the remaining soil variables in the conditioning set as additional covariates. The conditional independence between Milfoil and IC implied by the graph was not falsified by finding an insignificant statistical relationship between IC and the random site effects after accounting for Ca, CEC, and PerClay (P-value=0.682, Table 4 row 13).

Based on the p-values in Table 4, the GOF test statistic was 33.160, with 34 degrees of freedom, which resulted in a GOF p-value of 0.509 for the proposed model. Therefore, the probability of obtaining this GOF statistic or one more extreme based on observing many realizations of data generated by the assumed causal model is 0.509. Based on this, the data appeared consistent with the hypothesized linkage structure being true and we proceeded with developing a BN based on this graph

(Fig. 3). Again, because we had data available, we feel this additional information concerning the uncertainty of the linkage structure is useful to provide a better context for decision-making on the relative merits of water level manipulation for aquatic plants in this wetland complex.

4.4 Bayesian belief network

A BN commonly assumes that variables, represented as nodes in the graph, take on a fixed number of categorical states (e.g., high, medium, low). Consequently, we divided each variable in Fig. 3 into simple “high” and “low” categories, using the median as a threshold. We used TETRAD's expectation-maximization (EM) algorithm to estimate the conditional probabilities in our data set. Table 5 displays the conditional probabilities for sago pondweed and water milfoil.

Table 5 Conditional probability tables for sago pondweed and water milfoil

Previous Water	Calcium	CEC	Percent Clay	IC	Probability of High Milfoil
High	High	High	High	---	1
High	High	High	Low	---	0.93
High	High	Low	High	---	0.92
High	High	Low	Low	---	0.25
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
Low	Low	High	High	---	0
Low	Low	High	Low	---	0
Low	Low	Low	High	---	0
Low	Low	Low	Low	---	0

Previous Water	Calcium	CEC	Percent Clay	IC	Prob. Of High Sago Pondweed
High	High	High	High	High	0
High	High	High	Low	High	0
High	High	Low	High	High	0
High	High	Low	Low	High	0
*	*	*	*	*	*
*	*	*	*	*	*
*	*	*	*	*	*
Low	Low	Low	High	High	0.82
Low	Low	High	High	Low	0
Low	Low	High	Low	Low	0.6
Low	Low	Low	Low	Low	0.38

The resulting BN will be used to inform the refuge manager of the potential benefits of artificial water level manipulation, as well as understanding the potential effects of changing hydrology from such things as diversions, climate change, and the like. Generally, we found that sago pondweed abounds in shallow sites and water milfoil grows in deeper sites, similar to that of Anderson (1978). One hypothesis was that sago pondweed turion production is negatively affected during years with high water (Kantrud, 1990), and this seems to be supported by both the model and time series plots of vegetation abundance and water levels (Fig. 4).

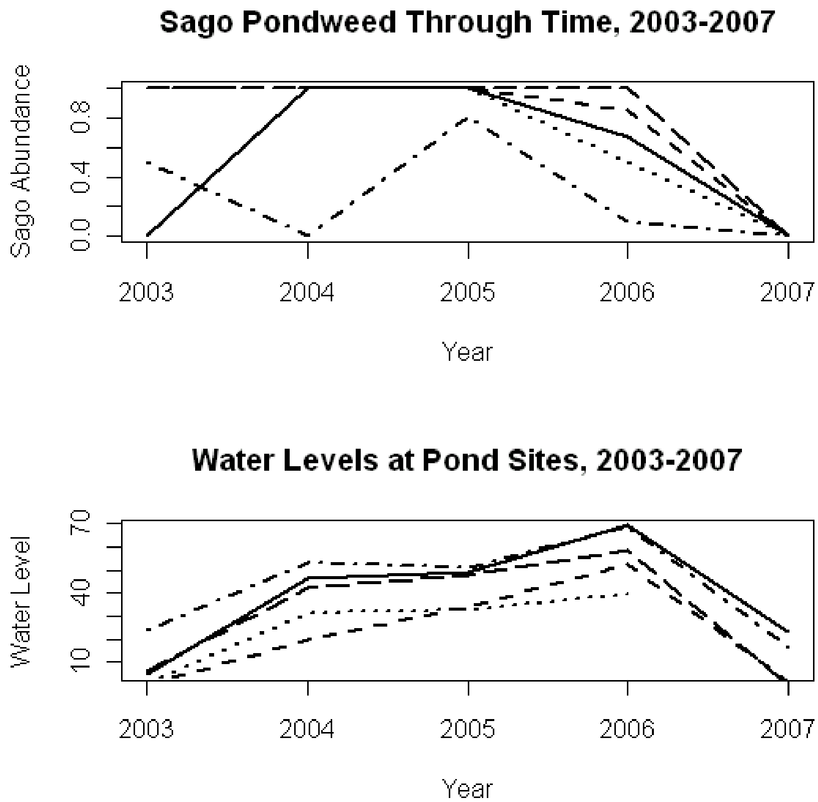

Fig. 4. Water levels and sago pondweed abundance measurements at selected sites through the study period. A spike in water levels in 2006 was associated with a severe decline in sago pondweed production the following year.

In the year following the highest water levels recorded during the study, sago pondweed abundance dropped drastically. We did not have enough temporal data to see this pattern repeated, but it could help managers evaluate the tradeoffs of different water level scenarios. High water levels seemed to favor northern water milfoil, and also may have inhibited sago pondweed growth and possibly turion production. If sago pondweed is preferred by key water birds at Red Rock Lakes NWR, it might be best to allow levels to fluctuate towards low levels as conditions permit. Our results are consistent with additional analyses at Red Rock Lakes NWR (Sojda et al. *in review*.), and in Manitoba (Anderson, 1978).

There were several clear benefits of verifying the linkage structure of the BN in this case study. First, it allowed us to include explanatory variables whose effects on vegetation abundance were initially unclear. Having four or five parents for the terminal variables in a BN is atypical, due to the high number of combinations that must be accounted for in the estimation step. Second, the conditional probability tables for our two submergent vegetation species showed interactions between covariates which likely would not have been predicted by expert opinion alone. As an example, percent clay soil composition in open water sites seemed to have a much different associations with the presence of high cation exchange capacities than with low capacities (Table 5). This type of interaction among specific soil characteristics is not well known for many plant species.

5. Conclusions

The method presented here for verifying linkage structures of BNs provides researchers and managers another tool for making sound decisions regarding ecological processes. It is not always possible to convert a conceptual model into a functioning BN using expert opinion alone, as it requires a tremendous level of detailed information. With a mix of empirical data and an *a priori* linkage structure in mind, users can develop a BN whose dependency structure can be tested with available empirical data. Ensuring that the graphical structure of the BN is adequate provides for greater insights into an ecosystem's dynamics and allows for a clearer picture concerning the complicated interrelationships amongst variables.

It is important to note that Bayesian networks are often interpreted as causal models, despite the philosophical issues that arise in attempting to demonstrate causation in an ecological system. The goodness-of-fit test presented here is not capable of confirming “true” causal relationships. Instead, it provides a measure of credibility concerning the causal model based on testing the implied statistical independences using observed data. The d-separation statements implied by a graph represent a specific set of logical statements, but do not paint a complete causal picture. The directions of the arrows in the linkage structure are not uniquely determined; several different causal graphs are capable of having the same basis set of d-separation statements.

Using our case study as an example, the relationship between percent clay, northern water milfoil and sago pondweed (Fig. 3) suggests that percent clay directly influences the abundance of both submergent vegetation species. The arrows could have been redrawn, however, so that water milfoil affects percent clay, which in turn affects sago pondweed. The d-separation statements associated with these two relationships are identical (sago pondweed and water milfoil are independent, after accounting for the amount of clay in the soil), but biology and common sense lead us to choose the first relationship over the second. If a causal relationship is to be deduced, it must rely strongly on relevant biological principles as well as common sense. And since the *de facto* interpretation of a BN involves making causal statements, verifying the conditional independencies implied by a graph using appropriate statistical tests based on observable data is a necessary step in the model development process, as we did with the Red Rock Lakes data.

A potential drawback with using a BN is that users inherently lose valuable information by binning and categorizing continuous variables. For example, Irvine and Rodhouse (2010) found that condensing Daubenmire cover class data into just 2 cover classes, as we did here, resulted in less power to detect temporal trend in plant abundance monitoring data. However, the broader class of graphical models (e.g., path analysis, structural equation models) may provide deeper insights into complicated wetland ecosystem dynamics (Ahronditsis et al., 2007; Grace and Pugesek, 1997) and could be used to investigate multi-scale spatial processes (Irvine and Gitelman, *in press*). We presented both the BN (Table 5) and path analysis estimates (Table 3) since we had data available for every node in our final graph (Fig. 3). The presentation of our results in multiple ways hopefully enhances understanding by a variety of stakeholders. But, we feel, in terms of conveying information to managers the BN table is easier to explain and a more concise way to portray the relevant information needed for sound decision-making.

Clearly, researchers do not always have large datasets available to use for testing the d-separation statements implied by a proposed graph. Indeed, the ability to incorporate expert opinion into BNs is one benefit (McCann et al., 2006). For researchers in this situation, we suggest that if data exists for some pieces of the model and not others, the data can be used to test relevant conditional independencies for certain subsets of the graph. At the very least, it will provide insight into whether a few of the graph's implied conditional independencies are supported by empirical data. The dependency structure can be adjusted based on these results, if necessary. In addition, researchers and managers who rely on expert knowledge driven BNs can verify whether the dependency structure of their model is supported by field data gathered in subsequent years.

Bayesian networks are an important ecological modelling methodology. All such models need empirical verification and validation (Cohen and Howe, 1989; Johnson, 2001; Sojda, 2007), and we believe that the methodology we have described is one useful way to incorporate field based measurements to determine the uncertainty in the linkage structure. Other goodness-of-fit measurements are available for structural equation models and path analysis (e.g., Saris et al., 2009; Stage et al., 2004). Relatively few ecological applications have considered assessing model uncertainty in Bayesian networks, although prediction uncertainty and model sensitivity are generally recommended (Marcot et al., 2006). We feel if a BN is used for interpretation about ecological processes, assessing the uncertainty concerning the linkage structure is imperative.

Acknowledgements

R. H. Buxton was instrumental in the hard work of installing wells and piezometers, as well as downloading and processing data. Without him, this project could not have happened in harsh, wilderness conditions. The scientific encouragement of L. Fredrickson to study the role of groundwater in this wetland complex was most welcome. Likewise, statistical analysis and guidance regarding water levels from M. Greenwood was most helpful. Staff at Red Rock Lakes NWR are recognized for their logistic support and technical help, especially J. Warren. D. Paullin's original random sampling scheme in Lower Red Rock Lake is also recognized. In addition, the authors sincerely thank A. I. Gitelman and J.B. Grace for their help in reviewing a draft of this paper, as well as the anonymous reviewers for their suggestions.

The authors would like to recognize the United States Geological Survey's Northern Rocky Mountain Science Center and Montana State University for their support. This project was funded by the USDI-Geological Survey and Fish and Wildlife Service. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Aalders, Inge., 2008. Modeling land-use decision behavior with Bayesian belief networks. Ecology and Society. 13, 16 [online].

Anderson, M.G., 1978. Distribution of sago pondweed (Potamogeton pectinatus L.) on a Northern prairie marsh. Ecology 59,154-160.

Arhonditsis, G.B., Stow, C.A., Steinberg, L.J., Kenney, M.A., Lathrop, L.C., McBride, S.J., Reckhow, K.H., 2006. Exploring ecological patterns with structural equation modeling and Bayesian analysis. Ecological Modelling 192, 385-409.

Barrows, C.W., Allen M.F., 2007. Biological monitoring and bridging the gap between land management and science. Natural Areas Journal 27, 194-197.

Daubenmire, R., 1959. A canopy-coverage method of vegetational analysis. Northwest Sd. 33,43-64.

Dlamini, W. M., 2010. A Bayesian belief network analysis of factors influencing wildlife occurrence in Swaziland. Environmental Modeling and Software 25, 199-208.

Cohen, P. R., Howe, A.E., 1989. Toward AI research methodology, three case studies in evaluation. IEEE Transactions on Systems, Man, and Cybernetics 19, 634e646.

Fisher, R.A., 1932. Statistical Methods for Research Workers 4th edition. London: Oliver and Boyd.

Grace, J.B., Pugesek, B.H. , 1997. A Structural Equation Model of Plant Species Richness and Its Application to a Coastal Wetland. The American Naturalist, 49, 436-460.

Irvine, K.M., and Gitelman, A. I., in press. Graphical spatial models: a new view on interpreting spatial pattern. Environmental and Ecological Statistics, DOI: 10.1007/s10651-010-0146-8 [online]

Irvine, K.M., and Rodhouse, T.R., 2010. Power analysis for trend in ordinal cover classes: implications for long-term vegetation monitoring. Journal of Vegetation Science, 21, 1152–1161.

Jensen, F.V., 2001. Bayesian networks and decision graphs. Springer.

Johnson, D.H., 2001. Validating and evaluating models. In, Shenk, T.M., Franklin, A.B. (Eds.), Modelling in Natural Resource Management, Development, Interpretation, and Application. Island Press, Washington, DCJohnson, S., Mengersen, K., de Waal, A., Marnewick, K., Cilliers, D., Houser, A.M., Boast, L., 2010. Modelling cheetah relocation success in southern Africa using an iterative Bayesian network development cycle. Ecological Modelling 221, 641-651.

Kantrud, H. A., 1990. Sago pondweed (Potamogeton pectinatus L.), A literature review. U.S. Fish and Wildlife Service, Fish and Wildlife Resource Publication 176. Jamestown, ND, Northern Prairie Wildlife Research Center,. Online. <http,//www.npwrc.usgs.gov/resource/plants/pondweed/index.htm> (Version 16JUL97).

King, E. G., Hobbs, R. J., 2006. Identifying linkages among conceptual models of ecosystem degradation and restoration, towards an integrative framework. Restoration Ecology 14, 369-378.

Korb, K.B., Nicholson, A.E., 2004. Bayesian artificial intelligence. Chapman and Hall. Boca Raton, Florida. 364 pages.

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

McCann, R.K., Marcot, B.G., Ellis, R., 2006. Bayesian belief networks, applications in ecology and natural resource management. Canadian Journal of Forest Research 36, 3053-3062.

McNay, R.S., Marcot, B.G., Brumovsky, V., Ellis, R., 2006. A Bayesian approach to evaluating habitat for woodland caribou in north-central British Columbia. Canadian Journal of Forest Research 36, 3117-3133.

Mead, R., Paxton, J., Sojda, R.S., 2006. Applications of Bayesian Networks in Ecological Modeling. Proceedings of Environmental Modelling and Simulation.

Meek, C. 1997. Graphical Models: Selecting causal and statistical models. PhD thesis, Carnegie Mellon University.

Ogden, J.C., Davis, S.M., Jacobs, K.J., Barnes T., Fling, H.E., 2005. The use of conceptual ecological models to guide ecosystem restoration in south Florida. Wetlands 25, 795-809.Peterson, D.P., Rieman, B.E., Dunham, J.B., Fausch, K.D., Young, M.K., 2008. Analysis of trade-offs by non- native brook trout and intentional isolation for native westslope cutthroat trout. Canadian Journal of Fisheries and Aquatic Sciences 65,557-573.

Renken, H., Mumby, P.J., 2009. Modelling the dynamics of coral reef marcoalgae using a Bayesian belief network approach. Ecological Modelling 220, 1305-1314.

Rudnick, D.T., Ortner, P.B., Browder, J.A., Davis, S.M., 2005. A conceptual model of Florida Bay. Wetlands 25, 870-883.

Russell, S., Norvig, P., 2002. Artificial Intelligence, A Modern Approach. Prentice Hall.

Saris, W.E., Satorra, A., van derVeld, W.M., 2009. Testing structural equation models or detection of misspecifications?. Structural Equation Modeling: A Multidisciplinary Journal 16, 561-582.

Shipley, B., 2002. Cause and correlation in biology, A user's guide to path analysis, structural equations, and causal inference. Cambridge.

Shipley, B., 2009. Confirmatory path analysis in a generalized multilevel context. Ecology 90, 363-368.

Smith, C.S., Howes, A.L. , Price, B., McAlpine, C.A., 2007. Using a Bayesian belief network to predict suitable habitat of an endangered animal, the Julia Creek dunnart. Biological Conservation 139, 333-347.

Sojda, R. S., 2007. Empirical evaluation of decision support systems, needs, definitions, potential methods, and an example pertaining to waterfowl management. Environmental Modelling and Software 22, 269-277. http,//www.sciencedirect.com/science/article/B6VHC-4JB9W57-1/2/a281828776a7fd1209ca763cad415f7c.

Sojda, R. S., Sharp, J. L., Greenwood, M., Rosenberry, D.O., Warren, J.M., in review. Statistical classification of vegetation and water depths in montane wetlands.

Spirtes, P., Glymour, C., Scheines, R., 2001. Causation, prediction and search. MIT Press.

Spirtes, P., Glymour, C., Scheines, R. The TETRAD Project, http,//www.phil.cmu.edu/projects/TETRAD/index.html>

Stage, F.K., Carter, H.C., Nora, A., 2004. Path analysis: an introduction and analysis of a decade of research. The Journal of Educational Research 98, 5-12.

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