Development And Application Of A Bayesian Decision Support Tool To Assist In The Management Of An Endangered Species

Pollino, C.A. and **A.K. White**

Water Studies Centre, Monash University, Clayton, Email: carmel.pollino@sci.monash.edu.au

Keywords: *Ecological risk assessment; Endangered species, Bayesian Networks,*

EXTENDED ABSTRACT

Bayesian decision support tools are becoming increasingly popular as a modelling framework that can analyse complex problems, resolve controversies, and support future decision-making in an adaptive management framework. This paper introduces a model designed to assist the management of the endangered Camphora swamp eucalypt (*Eucalyptus camphora*). This tree species is found in the Yellingbo Nature Conservation Reserve (YNCR), an isolated patch of forest in the Yarra Valley (Victoria, Australia).

The eucalypt community provides both habitat and food for a variety of threatened and endangered flora and fauna. Over the last 20 years *the E. camphora* has become increasingly threatened by dieback. In order to maintain and rehabilitate existing trees and encourage regeneration, management strategies and action plans have concentrated on restoring the hydrological regime, which has been altered due to agricultural activities within the catchment. However, research suggests that nutrient enrichment from surrounding horticulture and livestock is having a greater impact on the health of the trees. The Bayesian decision support tool has been used to examine the differences between these two hypotheses. The tool will also promote future integrative and iterative monitoring and research in the YNCR.

This project was undertaken as part of a larger piece of work developing risk-based assessment guidelines for natural resource management, a process known as Ecological Risk Assessment. The Woori Yallock Creek Catchment, of which YNCR is a part, was chosen for the case study due to the high ecological assets identified in the catchment, the diverse land use and activities in the region, and the large body of knowledge and data available. The case study aims to identify environmental assets at greatest risk from ecological degradation in the Woori Yallock Creek Catchment and subsequently identify options for managing these risks.

Phase One of the risk assessment process is problem formulation. Input was sought from a wide range of interest groups via one-on-one interviews and a workshop. The priority environmental assets in the catchment were subsequently identified, along with the hazards that threatened these assets. The Sedge-rich *E. camphora* community within YNCR was identified as one of the priority threatened environmental assets within the Woori Yallock Creek Catchment. For the purposes of this case study the condition of *E. camphora* was identified as the management end point for which YNCR will be managed.

Phase Two (risk analysis) utilised Bayesian networks (BN) to quantify the risks to *E. camphora*. BNs are probabilistic networks that support reasoning under uncertainty. BNs are used to establish causal relationships between key factors and final outcomes, and maintain clarity by making causal assumptions explicit (Stow and Borsuk 2003). They are particularly useful for uncertainty analysis as they have the ability to consider inadequate knowledge or understanding of system processes, inherent randomness, subjective judgment and vagueness in parameter estimation, disagreement, measurement error and sampling error (Morgan and Henrion 1990).

Risks to the condition of *E. camphora* have been prioritised and key knowledge gaps identified, while accounting for predictive uncertainties. Using the networks, the outcomes of a range of management scenarios have also been tested.

The parameters that are most influential in determining *E. camphora* condition according to the model are generally soil nutrients, soil cations, pests or disease and inundation patterns (duration and frequency). The findings show *E. camphora* responses are different for each region in the model, and findings are specific for each survey.

Much of the data used to parameterize the model was patchy and qualitative. This has contributed to significant knowledge gaps. The results of this study should be viewed as a guide to further work.

1. INTRODUCTION

The Woori Yallock Creek Catchment was chosen as a case study for an Ecological Risk Assessment project investigating the risks that contaminants pose to environmental sustainability in catchments.

The objectives of the study were to:

- Determine priority threatened environmental assets within the region;
- Identify associated hazards to environmental assets;
- Quantitatively assess the hazards to assets, and prioritise according to their degree of importance;
- Recommend priority hazards that should be targeted by management agencies;
- Advise where current monitoring programs can be improved and recommend additional studies that should be conducted to address key knowledge gaps.

2. ECOLOGICAL RISK ASSESMENT

Ecological Risk Assessment (ERA) is a methodology for determining the level of risk posed by hazards (e.g. biocides, nutrients and sediment) to identified ecological assets. ERAs evolved from the need to develop a formalised approach to environmental risk management. They are particularly useful for assessing the effects of multiple hazards to flora and fauna within inherently complex ecosystems.

The initial phase of a risk assessment is **problem formulation**, which involves identifying key/threatened environmental assets and identifying existing or potential hazards to environmental assets.

The **risk analysis** stage further investigates the hazards by quantitatively determining the likelihood of the adverse effect occurring and the consequences to priority environmental assets if such an event did occur.

3. CASE STUDY AREA

The Woori Yallock Creek catchment (Figure 1) is a sub catchment of the Yarra River basin, and has an area of 272 km^2. The headwaters of the Woori Yallock Creek Catchment originate in the valleys of the Dandenong and Yarra Ranges, with the lower reaches forming a broad floodplain (Woori Yallock Creek Sub Catchment Working Group, 1999).

Approximately 80% of the catchment has been cleared for agriculture and is dedicated to rural landuse.

Figure 1: Map of Woori Yallock Creek Catchment

Flows in the sub-catchment have been altered, with demand for water being year round (Zampatti et al. 1999).

Water quality monitoring data show that turbidity, nutrients (phosphorous and nitrogen), salinity (fluctuations only), and dissolved oxygen (percent saturation only) are currently not meeting EPA water quality criteria (assessment against Schedule F7 of Waters of Victoria). Septic tanks and sewage systems also discharge into the catchment, contributing to the high nutrient levels in surface water.

This paper describes the 2 phases of this study:

Phase One - Problem formulation;
Phase Two - Risk analysis.

4. PROBLEM FORMULATION

Problem formulation was undertaken via two approaches: one-on-one interviews and a stakeholder workshop.

4.1. Identification of Stakeholder Values

Elicitation of stakeholder catchment values took the form of interviews and a workshop. The purpose was to elicit environmental assets, the associated threats, and conceptual models. Stakeholders included community groups, resource mangers, scientists, local farmers, horticulturalists and Landcare groups.

According to stakeholders, an area of high importance in the catchment was the Yellingbo Nature Conservation Reserve (YNCR). In particular, the Sedge-rich *E. camphora* community within the YNCR was identified as one of the priority threatened environmental assets within the Woori Yallock Creek Catchment.

4.2. Yellingbo Nature Conservation Reserve (YNCR)

The YNCR is of high botanical and zoological significance. It contains approximately 285 native flora species and 230 native vertebrate species. Resident threatened/endangered species include the Sedge-rich *E. camphora* Swamp Community, the Helmeted Honeyeater (*Lichenostomus melanops cassidix*) and the Leadbeater's Possum (*Gymnobelideus leadbeateri*).

A number of serious hazards impact the reserve vegetation communities. These include dieback, insect attack, weed invasion, changed hydrology, sedimentation and eutrophication (Parks Victoria, 2004). The linear nature of the reserve makes it susceptible to edge-effect disturbances, including weed invasion, wind throw, clearing, dieback and use of fertilizers and herbicides (Parks Victoria, 2004). High turbidity and nutrient levels are believed to be contributing to the decline in quality of the riparian zone (Parks Victoria, 2004).

4.3. Eucalytus camphora

Dieback of *E. camphora* has been identified as one of the main threats to the health to the YNCR. *E. camphora* is predominately found in the YNCR, an isolated patch of forest in the Yarra Valley (Victoria, Australia). The eucalypt community provides both habitat and food for a variety of threatened and endangered flora and fauna.

E. camphora has suffered major dieback since the 1970s, which is ongoing. The sedge-rich *E. camphora* Swamp Community is restricted to the reserve and listed as endangered. The reserve has been actively managed since about the 1950s, and dieback has been actively managed since the 1990s.

In order to maintain and rehabilitate existing trees and encourage regeneration, management strategies and action plans have concentrated on restoring the hydrological regime, which has been altered due to agricultural activities within the catchment. However, other research suggests that nutrient enrichment from surrounding horticulture and livestock is also having an impact on the health of the trees. To date, no management strategies have been effective in slowing dieback in the area.

5. RISK ANALYSIS

Bayesian networks were used to further describe risk factors, and to predict responses in *E. camphora* condition to changing environmental variables.

The Bayesian decision support tool was used to examine the differences between these conflicting hypotheses regarding the cause of *E. camphora* dieback, and to further inform management actions and identify key research areas. The tool also aims to promote future integrative and iterative monitoring and research in the YNCR.

Background on Bayesian Network (BN) Model:

BNs are graphical models consisting of a directed acyclic graph with an associated set of conditional probability distributions used to represent linkages (Reckhow 2002). Probabilities can be estimated using expert knowledge, empirical data, or both (Marcot et al. 2001, Rieman et al. 2001). Network probabilities are updated using Bayes' rule. Networks can be updated as new scientific knowledge is made available. This type of iterative improvement fits into the ecological risk assessment paradigm (Pollino et al. in press).

The BN approach was considered to be the optimal modelling option for this study given that much of the information available was qualitative, and that the final product had to form part of an iterative process of adaptive management.

The steps involved in the development of a Bayesian network model are outlined in Pollino et al. (in press). The modelling software Netica (Norsys, Inc.) was the software platform used for model building. In developing relationships between variables, both qualitative and quantitative information was used.

5.1. Bayesian network for *Eucalyptus camphora* condition

The *E. camphora* swamp community is a complex system with multiple factors (hazards) interacting and influencing eucalypt condition.

BN structure

The BN modelling approach was used to graphically represent the independent and interacting variables influencing *E. camphora* condition. The BN structure was based on the conceptual model constructed from Phase 1, and developed using input from expert stakeholders.

Using the approach of Borsuk et al. (2004), the criteria for inclusion of variables in the BN was to determine whether the variable was 1) manageable, 2) predictable, or 3) observable at the scale of the management problem. If the variable did not meet one of these criteria, it was not included.

In the final model, the variables represent the key factors (hazards) in the YNCR that are recognised in the scientific literature and by stakeholders as being important in determining the condition of *E. camphora*. A schematic of the BN is shown below.

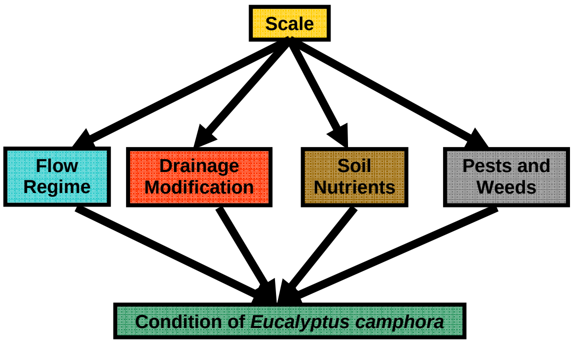

Figure 2: Schematic of the *E. camphora* Bayesian network.

Data and Information Case File

A set of findings can be entered into nodes of a BN as a case. A library case file was generated using the various data and knowledge sources. Sources included Parks Victoria, the Department of Sustainability and Environment, the University of Melbourne, and consultant studies.

Conditional Probability Distributions

In order to represent continuous distributions in the BN, nodes were discretised into sub-ranges. Where possible, nodes are discretised according to existing guidelines and classifications. The remaining nodes are discretised based on expert opinion or using percentiles of data.

Conditional probability distributions can be obtained three ways; they can be judged directly from experts, obtained from scientific literature, or obtained by fitting a network to a set of observed cases (Henrion et al. 1996). In this study all three information sources were used.

Qualitative and quantitative information was set up as a set of cases. Netica was used to perform updating using the EM algorithm (solving the network by finding the posterior probability for each node based on information in the cases) or specified using equations. The EM algorithm searches over the Bayesian network conditional probability tables in an attempt to maximize the probability of the data given the Bayesian network (i.e. minimize negative log likelihood) (Norsys, 2005). Conditional probabilities are assumed to be independent.

5.2. Model scales

A database was compiled. Data was incorporated into the model as case files, and represented in the network as probability bars. Data was kept explicit for temporal (where possible) and spatial scales. The source of the data was also kept explicit. In the absence of monitoring data, probabilities are equally distributed.

Model scales are described below.

Temporal scale

The model is a static representation of a scenario, and is not dynamic. Apart from flow and rainfall, no temporal scales were added to the model.

Spatial scale

Spatially the model considers different creeks (Woori Yallock, Cockatoo and Macclesfield Creeks). These are further divided into Healthy/Unhealthy portions of creeks according to the data obtained from information sources used.

Report/Survey

Many surveys/studies have been conducted looking at the health of the tree community and researching the role of potential stressors in the YNCR. In each report, key stressors to *E. camphora* were identified, and associated with these were a range of management actions. The findings and recommendations of the reports/surveys have not been consistent.

In order to maintain and rehabilitate existing swamp gum and encourage regeneration, recent management strategies (Parks Victoria 2004) have concentrated on restoring the hydrological regime. However, research from the same time period and at overlapping sites suggests that nutrient enrichment from surrounding horticulture and livestock is having a greater impact on the health of the trees (Kasel 1999).

In the model, the survey/report, was made a causal factor for data nodes. This was done to determine if there are synergies between surveys, and to determine whether there was a strong causal link between survey and *E. camphora* condition (i.e. whether the findings of each survey were biased according to variables measured).

5.3. Assessment and prioritization of the hazards to *Eucalyptus camphora*

Sensitivity analysis was used to determine the variables that have the most influence on the condition of *E. camphora*. Sensitivity analysis results are represented by plotting the variation in a target node, given the change in other network nodes over a probability range of 0 to 1. These outcomes can be used as a guide to identifying influential variables of *E. camphora* condition.

Sensitive analysis can also be used to identify key hazards, and parameters that need to be determined accurately. In a management context, it is these variables that may represent key management actions (Pollino et al. in press).

Here we discuss two models with two types of outputs. The first type links the findings of each variable to the report/study where that data/information was obtained. This was done to identify whether a common cause for the decline in *E. camphora* has been identified. The second model type integrated all report information from the YNCR, with report/survey removed from being an explicit factor in analyses, enabling an integrated assessment of all study findings.

Report/Survey Specific Model

The temporal and spatial scales of the reports and surveys were mapped using GIS.

- All sites, all surveys

When considering all sites and all surveys, the most influential link to the endpoint was region (Figure 3).

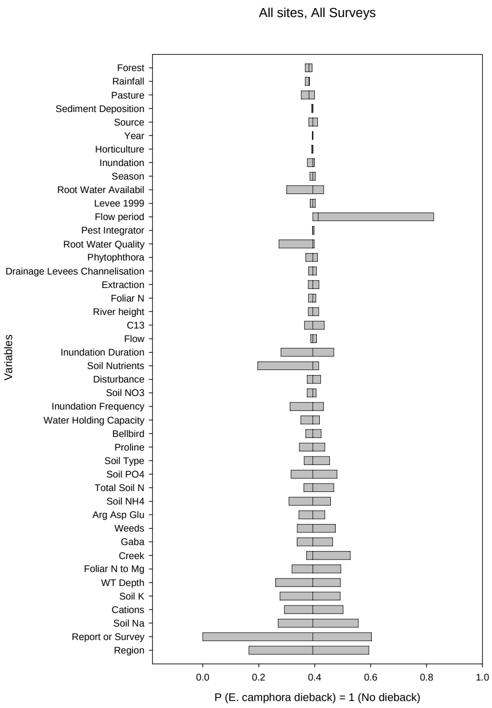

Figure 3: Sensitivity analysis results for *E. camphora* model.

Findings indicate that the response if *E. camphora* is region specific. This is consistent with report findings and advice from experts stakeholders consulted in this study.

The next sensitive variable is report/survey author. This indicates that the findings of each report are specific to the study, not to the problem being investigated. Sensitivity analyses also show that soil cations are influential in the network.

- Region specific, all surveys

Sensitivity analyses were conducted for each region, and findings consistently show the author of the report as being the most influential variable to *E. camphora*.

Sensitivity analyses were also conducted on networks specific to Cockatoo Creek, and specific to reports. This was done to demonstrate the potential for different studies to identify different sets of priority stressors, despite being conducted in similar areas over similar time scales.

- Cockatoo Creek, Craigie et al. (1998)

In figure 4, the influential variables in the BN are ranked.

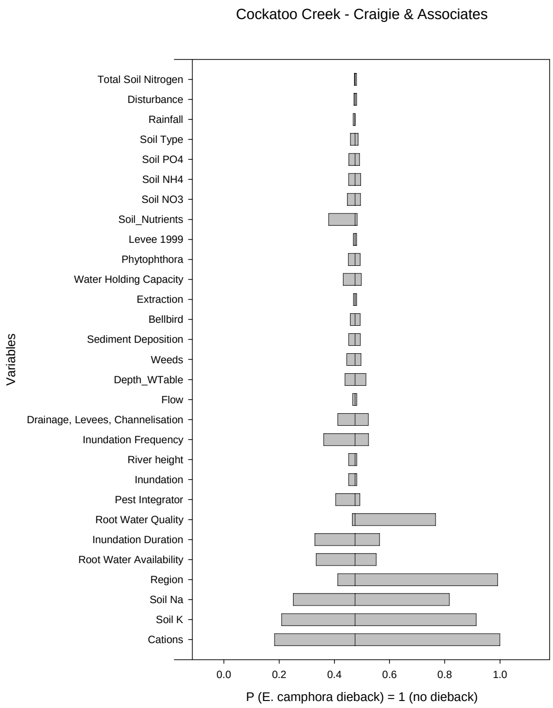

Figure 4: Sensitivity analysis results for *E. camphora* model, relevant for Cockatoo Creek and the Craigie et al. (1998) report.

Soil cations were the most influential node in the network. Root water availability was also important, primarily due to prolonged inundation duration and the high frequency of inundation. Water quality at the roots is predicted to be poor. Disturbance due to drainage, levees and channelisation was a factor contributing to prolonged inundation duration.

- Cockatoo Creek, University of Melbourne

Multiple factors were found to be influential in the University of Melbourne network (Figure 5). This reflects the greater number of variables measured. Soil nitrate and pests, (particularly *Phytophthora*) were important variables. Soil cations, soil type (due to water holding capacity), soil phosphate and water-table depth were also influential variables.

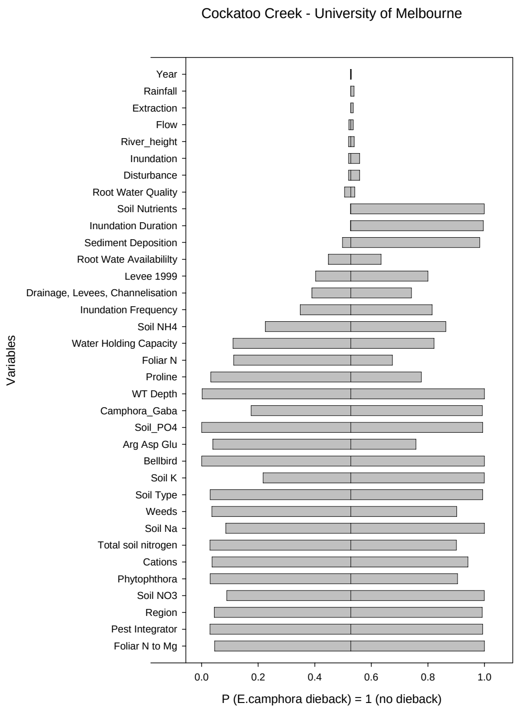

Figure 5: Sensitivity analysis results for *E. camphora* model, relevant for Cockatoo Creek and the University of Melbourne (Granger et al. 1994, Kasel 1999) reports.

<u>**Integrated Model**</u>

The report source node was removed in order to enable an integrative assessment of system stressors.

The major variables important to maintaining *E. camphora* condition were again ranked using sensitivity analyses.

- All creeks

Sensitivity analysis indicates pests (weeds and bellbirds), soil nitrate and ammonium, and soil type are important variables. The model again indicates that *E. camphora* condition is specific to each region.

- Woori Yallock Creek

Sensitivity analysis indicates water availability at the roots is influential. Pests (in the form of weeds) are also important. *Phytophthora* requires further investigation, as concluded by Limongiello and Keane (1995). Another important variable is soil nitrate concentration, which are elevated in this area.

- Macclesfield Creek

Sensitivity analysis indicates several variables are of high importance. Cations in the soil, *Phytophthora*, and soil nutrients are important variables measuring health of *E. camphora* within the Macclesfield Creek area. Soil cations are greatly increased in 'unhealthy' communities of trees, as compared to 'healthy' communities. Soil nutrients are also reduced in 'unhealthy' communities.

- Cockatoo Creek

Several variables are influential at Cockatoo Creek. Important variables include pests (weeds, bellbirds, and *Phytophthora*), soil nitrate and phosphate, soil type (and water-holding capacity), water table depth, and soil cations.

5.4. Model Validation Tests

Predictive accuracy was used to evaluate the *E. camphora* model, where 80% of data was used for training, and 20% used for testing. The error rate of the model was 27%. This finding gives us confidence in model predictions. Nonetheless, continued validation of the model is essential.

5.5. Model Limitations

The data used to parameterise the model was very patchy and much of the data was qualitative. Often data (for example, hydrological and nutrient data) was not available across the spatial or temporal scales examined in the network. Hence, the results of this study should be viewed simply as a guide to further work.

Despite these limitations, the development of this model enabled a complex system to be conceptualized, complex processes to be integrated into a single endpoint, and disparate studies and monitoring data to be aggregated. The end product is a model which has performed much better in validation tests than anticipated. The model also readily identifies key variables and knowledge /information gaps, specific to a creek or region of creek in the YNCR.

6. RECOMMENDATIONS AND KNOWLEDGE GAPS

The model has been used to identify future field studies and sampling strategies that can be implemented to improve our understanding of the hazards to *E. camphora*. With future monitoring and research, the BN can be iteratively updated. This will inform and assist the development of future management strategies.

6.1. Future Reports/Surveys

As *E. camphora* condition in the YNCR is more dependent on the author of the report/survey, rather than the results generated, it is essential that the contrasting findings be reconciled. Clearly what is required in the YNCR is a well structured monitoring program investigating knowledge gaps.

6.2. Knowledge Gaps

Key knowledge gaps include:

Pests: Role of *Phytophthora*?

Nutrient inputs: Major sources of nutrients need to be identified.

Soil nutrients: Is the area saturated with nutrients, having a major impact on tree health, and increasing vulnerability to pests and diseases?

Soil cations: Cations are clearly elevated in parts of the swamp that are declining. Does this indicate eutrophication?

Inundation duration and frequency: A more complete understanding of the patterns of inundation in the YNCR is essential

7. CONCLUSIONS

The graphical nature of BNs aids in the representation and communication of complex systems. BNs are valuable tools to take to workshops and expert interviews in order to elicit further information and to interrogate conceptual models. BNs provide a starting point for people to work with, whether they agree or disagree with the representation of the system structure, the relationships within that structure, the states of the variables, or the data informing these variables. Another advantage to using BNs is the ability to update the model when new information comes to light, or our understanding of the system changes. This model should feed into a broader program aimed at determining the major factors leading to the decline of the *E. camphora* community.

8. REFERENCES

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

Craigie, N. M., S. O. Brizga, and P. Condina. 1998. Assessment of proposed works to ameriolate the effect of hydrological processes on vegetation dieback at Cockatoo Creek Swamp, Yellingbo Reserve. Neil M. Craigie Pty. Ltd., Croydon.

Granger, L., S. Kassel, and M. A. Adams. 1994. Tree decline in southeastern Australia: Nitrate reductase activity and indications of unbalanced nutrition in *Eucalyptus ovata* (Labill.) and *E. camphora* (R.T. Baker) communities in Yellingbo, Victoria. *Oecologia* **98**:221-228.

Henrion, M., M. Pradhan, B. Del Favero, K. Huang, G. Provan, and O'Rorke. 1996. Why is diagnosis using belief networks insenstitive to imprecision in probabilities? Pages 307-314 in *Proc. 12th Conf. Uncertainty in Artificial Intelligence.*

Kasel, S. 1999. The decline of *Eucalyptus camphora* and *E. ovata* within the Yellingbo State Nature Conservation Reserve, Victoria. PhD. University of Western Australia, Melbourne.

Limongiello, N., and P. Keane. 1995. Progress report on an investigation into the cause of dieback of *Eucalyptus camphora* in Yellingbo Nature Reserve. Unpublished report to the Helmeted honeyeater Recovery Team, School of Botany, Latrobe University, Melbourne.

Marcot, B. G., R. S. Holthausen, M. G. Raphael, M. Rowland, and M. Wisdom. 2001. Using Bayesian Belief Networks to Evaluate Fish and Wildlife Population Viability Under Land Management Alternatives from an Environmental Impact Statement. *Forest Ecology and Management.* **153**:29-42.

Morgan, M. G., and M. Henrion. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge, UK.

Parks Victoria. 2004. Yellingbo Nature Conservation Reserve Management Plan. Parks Victoria, Melbourne.

Pollino, C. A., O. Woodberry, A. E. Nicholson, K. B. Korb, and B. T. Hart. in press. Parameterisation of a Bayesian network for use in an ecological risk management case study. *Environmental Modelling & Software.*

Reckhow, K. 2002. Bayesian Approaches in Ecological Analysis and Modeling. *in*, http://www2.ncsu.edu/ncsu/CIL/WRRI/about/reckhow/Bayesapproach.pdf.

Rieman, B. E., J. T. Peterson, J. Clayton, P. Howell, R. Thurow, W. Thompson, and D. C. Lee. 2001. Evaluation of potential effects of federal land management alternatives on trends of salmonids and their habitats in the interior Columbia River basin. *Forest Ecology and Management*. **153**:43-62.

Stow, C. A., and M. E. Borsuk. 2003. Enhancing causal assessment of fishkills using graphical models. *Ecosystems* **6**:11-19.

Zampatti, B., J. Lieschke, L. Grgat, and T. Raadik. 1999. An Assessment of Environmental Flow Requirements for the Woori Yallock Creek Catchment. Natural Resources and Environment, Victoria.