18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Analysing social data on adoption of conservation practices: Exploring Bayesian networks

Ticehurst, J.L. [1] **Curtis, A.** [2] **and Merritt, W.S.** [1]

[1] *Integrated Catchment Assessment and Management (iCAM) Centre, The Fenner School of Environment and Society, The Australian National University, Australian Capital Territory*
Email: jenifer.ticehurst@anu.edu.au
[2] *Institute for Land, Water and Society, Charles Sturt University, Albury, New South Wales*

Abstract: Australia's Natural Resource Management (NRM) regions are required to report to the Australian Government on the impact their investments have on natural resource condition. Reporting to date has typically been limited to describing the nature of investment and on-ground activities that have taken place. In some cases activities are linked by crude assumptions to an expected change in condition. With the Australian Government pushing for regional reporting that focuses on outcomes (change in condition) rather than outputs (activity and dollars spent), the regions are seeking ways to improve their current investment planning and resource condition reporting. This requires an improved understanding of biophysical systems to establish causal links between management actions and change in resource condition, and social research to better understand who in the community is likely to respond to which type of environmental programs and why.

Techniques exist to analyse large complex social data sets including statistical and social-psychology models. Bayesian Networks (BNs), which are increasingly being used to model environmental systems, have not often been used in the analysis of social data. A BN (Figure 1) is a dynamic way to analyse complex cause and effect relationships. This paper explores the utility of BNs for analysis of social data sets and compares this approach to other more commonly applied techniques.

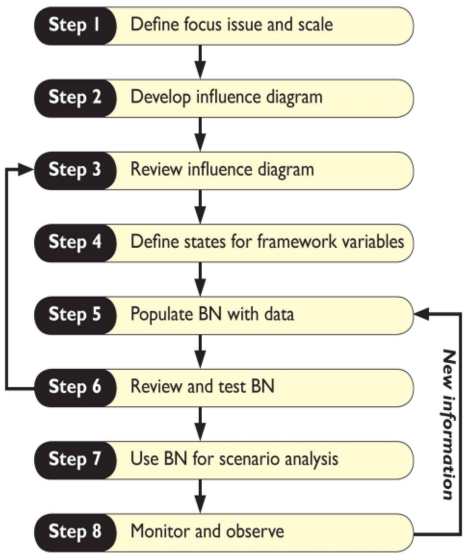

Figure 1. Typical steps used to develop a BN (Source: Ticehurst et al. [in preparation]). Here an influence diagram is a conceptual map of the perceived causality of a system, which is reviewed by relevant stakeholders and experts..

Survey data from the Wimmera Catchment Management Authority was used to develop a BN model of the social drivers of conservation activity adopted by landholders to protect native vegetation. The resulting BN shows relationships between a landholder's likelihood of fencing native vegetation and their values, knowledge, attitudes, income and access to government support. It clearly illustrated the importance of government funding in the uptake of conservation, but also showed that a significant amount of fencing was carried out in the absence of government programs. In the absence of government funding, on-farm income was found to be critical to the uptake this activity, illustrating that whilst landholders may have been willing to adopt recommended practice, the behaviour depended on their 'capacity to change' (in this case, financial capacity).

This paper suggests that BNs based on social research could be used by managers to support their decision-making and reporting, and has value for researchers as a tool for analysing, interpreting and communicating social data.

Keywords: *social data, Bayesian networks, landholder adoption*

1. INTRODUCTION

The Natural Resource Management (NRM) Regions of Australia are required to report to the Australian Government on the impact their management has upon natural resource condition. Reporting to date has typically described the investment and on-ground activity that have taken place in the region. With increased focus on reporting resource condition outcomes rather than the input activity and dollars spent (Hajkowicz, 2008) the NRM Regions are seeking ways to tighten their investment planning and resource condition reporting processes. This requires a better understanding of the likely change in biophysical condition with a change in on-ground activity, while accounting for the expected implementation of conservation activities by private landholders.

Pannell et al. (2006) suggest that the adoption of conservation practices occurs as part of a process where landholders become aware of a practice and then assess its relevance to their situation. Decision making is affected by a complex web of factors that is specific to the individual, particular technologies and to the social context of the landholder at that time. In order to identify causal relationships and trends in such complex systems a suite of tools exist for social data analysis, most noticeably social – psychological models (e.g. the Theory of Planned Behaviour – Azjen, 1991) and statistical analyses such as pairwise and multi-variate statistics, including regression models (e.g. Curtis and Robertson, 2003).

This suite of techniques is widely accepted and effective in assisting those attempting to identify differences between groups, and relationships between variables. However, not all statistical analyses can be applied to all data types (e.g. categorical, scalar, continuous), and some methods require the data to be suitably distributed (e.g. normally distributed or linearly related).

Bayesian Networks (BNs) are one of many techniques available to integrate data, knowledge and information from different sources, disciplines and knowledge types. BNs are developed to represent a system through a series of variables joined by causal links (termed the influence or conceptual diagram) where each link is described using probability distributions. There is no set structure or theory for the conceptual diagram allowing model developers to represent a wide range of variables considered to influence an outcome (e.g. economic, environmental, and social factors). The probability distributions can be determined using both quantitative (observed data, mathematical relationships or model simulation results) and qualitative (expert and local knowledge) information. The advantages and disadvantages of BNs for integration in Natural Resource Management (NRM) are further explored in Ticehurst et al. (2008).

In the scientific literature, BNs have been increasingly used to model the biophysical, social, and/or economic impacts of implementing NRM strategies or technologies (e.g. Cain et al. 1999) but have not often been used to explore the influence of social factors on technology adoption by rural landholders. Castelletti and Soncini-Sessa (2007) developed an integrated model of a water reservoir network by coupling a BN to hydrological models. The simple BN in the model represented farmers' behaviour in the irrigation districts. They found BNs useful in organising farmers responses to considered actions (extension and incentives) under different levels of expectations (low, medium and high), given that there 'were no physical laws describing farmers' response' (Castelletti and Soncini-Sessa, 2007).

Sebastiani and Ramoni (2001) developed BNs using data from the British Office of National Statistics to demonstrate the potential of the technique in analysing survey data. This annual survey provides information on population, housing, education, employment health and income. The BN developed from these data was able to show directed and conditional dependencies between network variables and, by propagating the network, undirected associations (Sebastiani and Ramoni, 2001).

This paper describes a small pilot study, funded by the Landscape Logic project (www.landscapelogic.org.au) to explore the usefulness of BNs for the analysis of social data sets compared to other techniques. To do this we have focused on exploring the uptake of fencing to manage stock access to native bushland and grasslands in the Wimmera Catchment Management Authority (CMA) region in western Victoria, Australia, using existing quantitative data and one expert's opinion.

2. CASE STUDY BACKGROUND

The Wimmera CMA comprises a large range of landscapes across 23,500 km^2 of north western Victoria including the catchment of the Wimmera River. Widespread clearing of native vegetation occurred over the last century for crop and livestock production. The region has a population of approximately 50,000 people, about one-third of who live on family farms or in small towns.

In 2007, Curtis et al. (2008) sent a survey to 1,000 rural property owners (>10 ha) selected at random from local government ratepayer data bases. A final response rate of 50% was achieved (N=503). The survey collected data to improve understanding of trends in social/ farming structure (property size, property turnover, property subdivision/ amalgamation), landholder adoption of recommended practices identified by the CMA's Regional Catchment Strategy, progress towards achievement of resource condition objectives in the region, landholder acceptance of a range of NRM policy instruments, and landholders' preferred sources of NRM information. For a more detailed description of the region and the survey method see Curtis et al. (2008).

3. DATA ANALYSIS METHODS

The survey responses were analysed using conventional techniques, briefly described in Section 3.1. The same data were also used to develop a BN (Section 3.2) to further investigate the causal relationships between social drivers and the adoption of conservation practices.

3.1. Conventional Analysis

Curtis et al. (2008) undertook statistical analyses (Table 1) to explore factors affecting the adoption of each of the 10 Current Recommended Practices (CRP) included in the survey and also identified in the Wimmera CMA Regional Catchment Strategy. These practices included the area of gully erosion addressed, fencing to manage stock access to waterways, and planting trees and shrubs, including through direct seeding.

Table 1. Statistical analyses of the Wimmera survey data (Source: Curtis et al. 2008)

Analysis / Test	Description
Spearman rank order correlations	Used as an exploratory tool to search for relationships between variables as well as natural groupings (e.g. Why is the property important to respondent?)
Pearson's Chi-square	Used to compare categorical data against each other (e.g. Is there a relationship between completing a Property Management Plan (PMP) short course and how important the respondent believes native vegetation is on the property for providing habitat?)
Kruskal-Wallis Rank Sum	Used to determine the significant difference of a continuous variable based on a second grouping of variables (e.g. Is there any significant difference in property size between those who consider their occupation to be a farmer and those who don't (e.g. professional, trade, retiree, etc)?)
Multiple linear regression	Used to identify relationships between all continuous variables. (e.g. How long the respondent has lived in the local district and the size of their property)
Multiple logistic regression	Used to identify relationships when the dependent variable is categorical (i.e. yes/no) (e.g. Is there a relationship between the area of bushland fenced and if the respondent has a long-term vision for the property?)

The process for the statistical analyses was to consider each CRP and do pairwise comparisons with all variables to see which are significantly related. As the survey data are of different data types - categorical, ordinal, scale, etc – the different tests in Table 1 were applied as appropriate. Any variables that had a small response rate and/or had no plausible explanation were removed from further analyses. Automated stepwise modelling was then conducted using the Akaike Information Criterion (AIC) to select variables: linear models were constructed when the dependant variable (the Current Recommended Practice) was continuous or an ordinal scale and logistic models were constructed when the dependant variable was categorical (i.e. yes/no). This analysis was used to identify which variables most influenced the adoption of each CRP.

3.2. Bayesian Network Analysis

The process typically used to develop a BN was shown in Figure 1.

Step 1: Of the ten CRPs, consultation with the Victorian NRM partners of Landscape Logic project narrowed down which management actions would be most relevant to them, and thus would be used to develop the Wimmera BN. A main investment focus for the NRM regions involved in Landscape Logic was the fencing of native bushland and grasslands to manage stock access (*now referred to as fencing bushland*), which is the CRP that is the focus of this paper.

Steps 2 to 6: As part of the Landscape Logic project, work had commenced to document the drivers that influence native vegetation condition and develop a conceptual diagram. The resulting social framework was highly complicated with a large number of nodes representing the technological, biophysical, demographic, economic and policy variables that affect landholders' willingness and capacity to change. The complexity of the network (40+ variables) limited the practicality and ease of developing a BN that would be used to support decision making by NRM regions.

Results from a conventional statistical analysis of the Wimmera survey data, together with an expert opinion, were used to identify key variables and likely causal relationships, resulting in a less complex BN (13 variables). The expert was the social scientist familiar with the data and the Wimmera region as a result of studies over a 7 year period. Expert opinion was also used to assist defining the states for each variables within the BN. The data tables in the BN were constructed from the responses to the survey questions relevant to the fencing of bushland. A spreadsheet was developed that listed, for each of the 503 survey respondents, the state of each variable in the network. This data was imported into Netica (http://www.norsys.com/netica.html) and used to populate the BN model.

The BN was used to further investigate causal relationships by performing a sensitivity analysis for each variable in the BN. The sensitivity analysis ranks the influence of each variable in the BN on the variable of interest using the Mutual Information measure (for categorical variables) or the Variance of Beliefs measure (for continuous variables) described in Marcot et al. (2006). A value of 0 indicates no influence whilst a measure of 1 indicates a perfect causal relationship between two variables. The BN was then used to investigate causal relationships between the most sensitive variables.

4. KEY FINDINGS

4.1. Key factors driving adoption

Analysis of the Wimmera survey data using conventional techniques outlined in Section 3.1 identified a number of properties related to the adoption of fencing bushland (Table 2). These were used to identify important BN variables, while expert opinion was used to determine the relationships between them. Three variables, identified in Table 3, were added to the BN given their perceived importance (Figure 2).

Table 2. Variables identified from the conventional analysis to have a significant relationship with the fencing of bushland.

Variable
• A higher rating to 'native vegetation on my property provides habitat for native animals'
• A higher rating to 'being able to pass the property on in better condition'
• Higher self-assessed knowledge of the ability of perennial vegetation to prevent water tables rising
• Higher self-assessed knowledge of how to protect and improve the health of native bush areas
• Involvement in whole farm planning
• Having a long-term plan or vision
• Landcare membership or involvement
• Membership of a local commodity group
• Larger property size
• Identifying as a farmer by occupation
• Support from government
• Beef cattle producers
• Sheep meat producers
• Have patches of native bush

The final outcome variable in the BN is *Fence_bushland* (Figure 2). This variable is most sensitive to *Govt_support*, then *PMP* (Property Management Planning) and then *Local_organizations* (Figure 3). This reflects the strong link established from the survey between fencing of bushland and government support, local organisations and attendance at a PMP course. The BN showed that in the absence of government support, 52% of respondents fenced bushland while with government support this increased to 87% (Figure 4). The variables *Values*, *Residence_time*, *Longterm_vision* and *Knowledge* had negligible impact on the area of bushland fenced. Of the network variables that were not identified as being significant by the conventional analysis (*Labour_time*, *Onfarm_income*, and *Residence_time*), only *Labour_time* was identified as having a relatively large impact on *Fence_bushland* although *Onfarm_income* did influence the variable to a lesser degree.

In this study the importance of government funding on the uptake of fencing activities was important, however, it was apparent that considerable work was undertaken by landholders in the absence of funding. BN analysis also suggests that the level of on-farm income was critical to uptake, especially in the absence of government funding. With no support from the government, 50% of landholders with an on-farm income of **less** than $50,000 fenced bushland, compared to 66% of landholders with an on-farm income of **greater** than $50,000.

The BN showed that after government support, fencing of bushland was next most sensitive to landholders completing a Property Management Planning (PMP) course which increased the likelihood of fencing bushland from 55% to 69% (Figure 3). Similarly being a member of both Landcare and a commodity group increased the likelihood of fencing bushland from 56% to 71%. At least part of the contribution of these groups is their involvement in trials and field days where landholders can test the efficacy and explore the relevance of new practices with their peers under local conditions. These groups are also important in establishing norms about what "good farming in this district" involves. In order to promote the fencing of bushland, money and technical support provided by government, and the knowledge, skills and confidence developed through participation in PMP courses and local organisations are important.

Table 3. Description of the BN variables

Variables	Description
Fence_bushland	Whether the respondent has fenced native vegetation or grassland on their property
Govt_support	Whether the respondent has received external funding to fence existing native vegetation
Industry	The current enterprise(s) on the land: sheep for meat (*lamb*), beef cattle (*beef*), and/or cropping *(crop*).
Knowledge	The level of knowledge the respondent has on topics relevant to the fencing of native bushland and grasslands (i.e. the use of perennial vegetation to prevent watertables rising [*perennial vegetation*]; how to protect and improve the health of native bush on properties [*vegetation health*]).
Labour_time ^	How many hours per week the respondent worked on the farm in the last 12 months
Local_organisations	Whether the respondent belonged to landcare, and/or another commodity support group such as *Topcrop*
Longterm_vision	Whether the respondent had any long term vision about how they would like to improve their property
Occupation	The respondents nominated occupation
Onfarm_income ^	The approximate on-farm profit the respondent received in the previous financial year
PMP	Whether the respondent had completed a Property Management Plan (PMP) short course
Property_size	The size of the property the respondent owns/manages, where the median size is 630ha.
Residence_time ^	Average of years respondent had owned/managed their land and years they had lived locally district
Values	How important the respondent thought native vegetation on their property was to provide habitat for native animals (*habitat*), and that they were able to pass the property on to others in better condition (*better condition*)

^ variables that were included based on expert opinion.

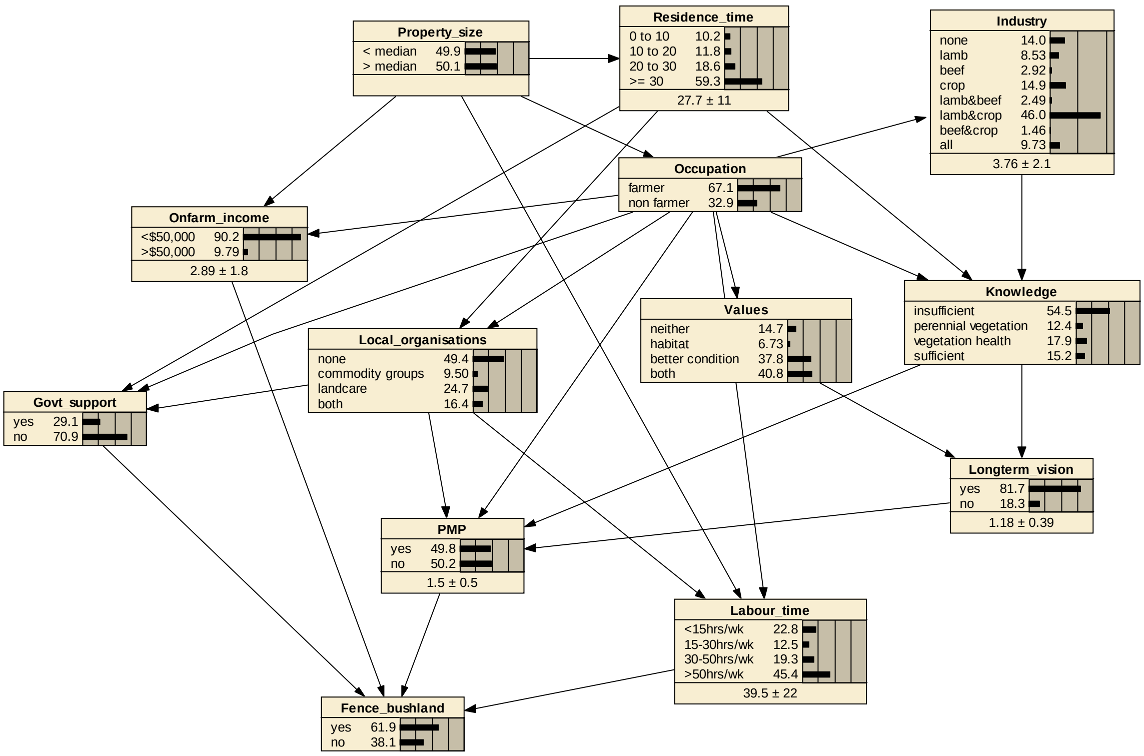

Figure 2. Wimmera BN showing the structure and states

Occupation was one of the top three factors influencing attendance at a PMP course and membership of a local organization. The BN can consider the responses of farmers only by assigning the occupation for farmers to be 100% (Figure 5, left side), and similarly for non-farmers (Figure 5, right side). Survey respondents identifying as farmers were more likely to fence their bushland than property owners who identified as non-farmers. Of those identifying as non-farmers, only 29% were likely to belong to Landcare or a commodity group, 28% to complete a PMP course, and 57% to fence bushland. By comparison, those identifying as farmers were 61% likely to belong to Landcare or a commodity group, 60% likely to complete a PMP course, and 64% likely to fence bushland. Farmers also spent considerably more time working on the property than non-farmers. Combined, these factors apparently off-set the lower responses for conservation attitudes and values of farmers which would otherwise be expected to lead to lower implementation of conservation practices than for non-farmers (Curtis et al. 2008).

5. DISCUSSION AND CONCLUSIONS

To better target the investment of NRM funds to achieve change in natural resource condition, it is important to improve understanding of the impact proposed management practice. At the same time, it is also important to improve understanding of the factors that influence landholder uptake of different conservation practices. This paper explored the potential for BNs to enhance understanding of the influence of social and economic factors on landholder decision making.

The development of this BN was facilitated by pre-existing survey data which empirically defined a set of factors that influence adoption; knowledge gained from the statistical analyses of that survey data described in Curtis et al. (2008); and the interaction of the BN team with the social scientist who carried out that survey. Without the prior learning and expert opinion, the conceptual diagram would have been more difficult to develop and potentially more complex. That said, BNs are well-suited to adaptive development and can be progressively refined or simplified through sensitivity analyses and through the integration of additional expert opinion and other qualitative information.

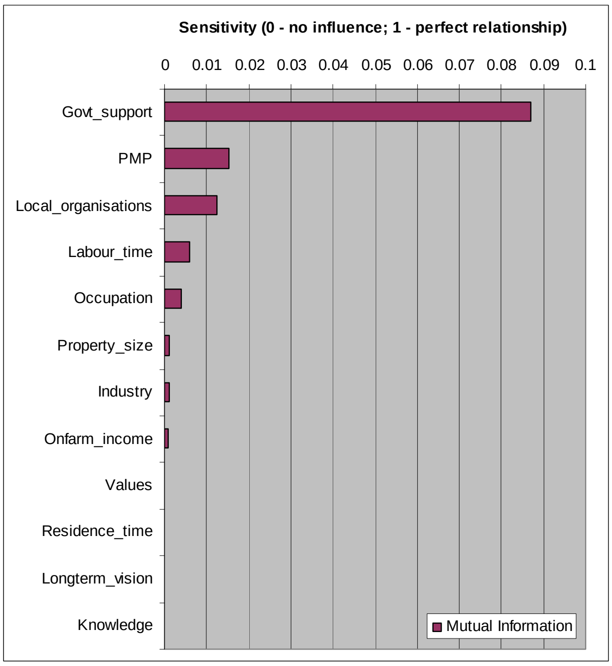

Figure 3. The sensitivity of *Fence_bushland* to all other variables in the BN.

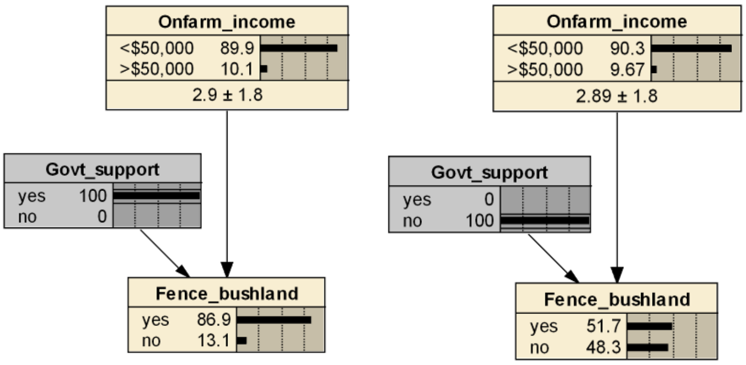

Figure 4. The impact of government support on the fencing of bushland (*Left:* with support; *Right:* no support). Note that for clarity only the outcome variable (*Fence_bushland*) and the *Govt_support* and *Onfarm_income* variables are shown.

Government support has the greatest influence on the fencing of native bushland and grassland to manage stock access, increasing the likelihood from 51.7% to 86.9%. Landholders identifying as farmers are more likely to belong to relevant local organisations (e.g. Landcare, commodity groups), complete a property management plan, have higher on-farm profit and labour levels, and access government support. Although they place less importance than non-farming landholders on values such as habitat provision (BN results 40% farmers compared to 61% non-farmers not shown here), the combined effect of these characteristics meant farmers were more likely to fence bushland, than non-farmers.

Given that the same data was used in the conventional and BN analyses, the key findings are, not surprisingly, similar. The advantage of the BN approach is that the relationships derived from the survey data and inferred by the social scientist involved are made more explicit and are more open to scrutiny and analysis by investors and policy makers. Anecdotal evidence found that BNs were particularly useful in the interpretation of survey data and communication of analyses (Curtis, 2008 pers.comm.). The graphical representations of the results (e.g. Figure 4) were used to communicate the results to the relevant stakeholders from the Wimmera CMA in a clear, simple manner. BNs were also useful to social scientists as they enable them to structure their understanding of the interactions and strength of the many different factors that influence the adoption of conservation practices or other technologies response.

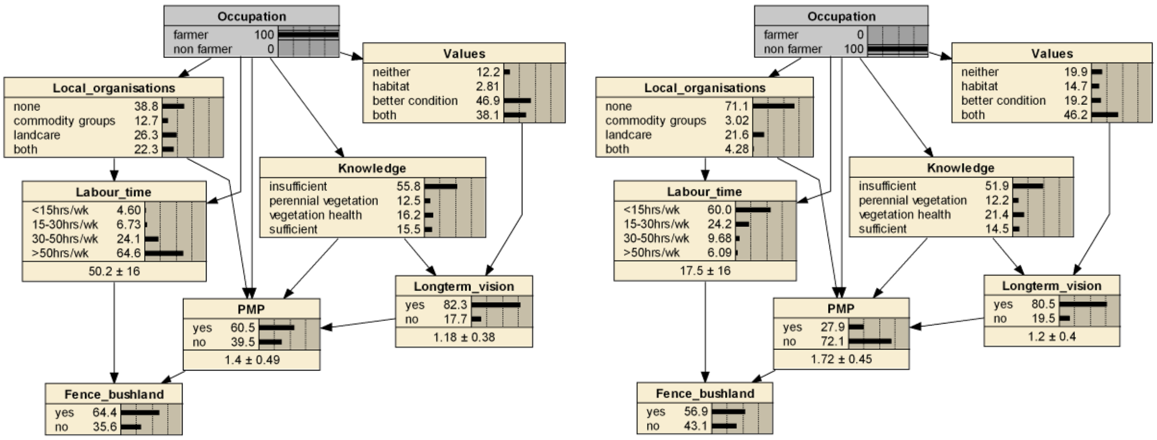

Figure 5. The impact of occupation on values, self assessed knowledge, labour time, involvement with local organisations or PMP courses and the fencing of bushland (Left: farmers, Right: non- farmers).

This paper has demonstrated the value of BNs in the representing the social drivers of the adoption of conservation practices that influence the condition and extent of native vegetation. For the Landscape Logic project, researchers from the Charles Sturt University are currently undertaking social research into landholders adoption of vegetation management practices in three NRM regions in Victoria. BNs will be developed based on this research and linked to biophysical BNs that model the impact of management practices on native vegetation in the regions.

ACKNOWLEDGEMENTS

We gratefully acknowledge the Wimmera CMA for granting permission to use the Wimmera survey data for this pilot study. This work was undertaken as part of the Landscape Logic Commonwealth Environment Research Facility (CERF) project which is funded by the Australian Government Department of Environment, Water, Heritage and the Arts.

REFERENCES

Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behavior and Human Decision Processes*, **50**, 179-211.

Cain, J., Batchelor, C., and Waughray, D. (1999). Belief networks: a framework for the participatory development of natural resource management strategies. *Environment, Development and Sustainability*, **1**, 123-133.

Castelletti, A., and Soncini-Sessa, R. (2007). Coupling real-time control and socio-economic issues in participatory river basin planning, *Environmental Modelling and Software*, **22**(8), 1114-1128.

Curtis, A., and Robertson, A. (2003). Understanding landholder management of river frontages: the Goulburn Broken. *Ecological Management and Restoration*, **4**(1), 45-54.

Curtis, A., S. McDonald, E. Mendham and R. Sample, (2008), Understanding the social drivers for natural resource management in the Wimmera region, Client Report for Wimmera Catchment Management Authority, Victoria, Australia, Institute for Land and Water, Charles Sturt University.

Hajkowicz, S. (2008). The evolution of Australia's natural resource management programs: Towards improved targeting and evaluation of investments, *Land Use Policy*, doi:10.1016/j.landusepol.2008.06.004.

Karppinen, H. (2005). Forest owners' choice of reforestation method: an application of the theory of planned behaviour. *Forest Policy and Economics*, **7**, 393-409.

Marcot, B.G., Steventon, J.D., Sutherland, G.D. and McCann, R.K. (2006). Guidelines for developing and updating Bayesian belief networks for ecological modelling. *Canadian Journal of Forest Research*, **36**, 3063-3074.

Sebastiani, P., and Ramoni, M. (2001). On the use of bayesian networks to analyze survey data. *Research in Official Statistics*, **4**(1), 53-64.

Ticehurst, J.L., R.A. Letcher, and D. Rissik, (2008), Integration modelling and decision support: a case study of the Coastal Lake Assessment and Management (CLAM) tool. Mathematics and Computers in Simulation, 78, 435-449.

Ticehurst, J.L., Pollino, C.A., and Lefroy, E.C. (*in preparation*). Overcoming barriers to participatory modeling in environmental management: lessons from two Australian case studies.