Observations from field trials with several elicitation techniques in an ecological domain

Colette R. Thomas*
Water Studies Centre
Monash University,
VIC, 3800, Australia

Ann E. Nicholson
Faculty of Information Technology
Monash University
VIC, 3800, Australia

Barry T. Hart
Water Studies Centre
Monash University,
VIC, 3800, Australia

Abstract

Quantitative ecologists use Bayesian networks (BNs) to integrate their collective understanding of system processes, and to adaptively investigate management alternatives. Consequently, subjective probability assessments are often critical for ecological BNs. Several published probability elicitation techniques were trialled in development of a prototype ecological BN. These included verbal, numeric, text and matrix formats. Observations of the participant's preferences for and performances under the different formats are described and discussed.

1 INTRODUCTION

We wanted to construct a BN collaboratively with the key end-user group for the domain, namely tropical seagrass managers and scientists in the Great Barrier Reef World Heritage Area (GBRWHA), in northeastern Australia. In this region elevated nutrient and sediments entering the GBRWHA from river flows are considered one of the most important land-based influences on the system (Brodie et al. 2007), although the issue has been contentious (e.g. Starck 2005). A risk-based approach using BNs was considered way to tackle these problems in the GBRWHA, however data scarcity meant that experts were required to provide some of the probability estimations for the BN.

Given the complexity and data scarcity of most ecological systems, significant effort is required to maximise the extraction of information from available data. Most data types can be adapted to BN analysis this is one of the reasons why BNs are so appealing to ecological risk practitioners. However, rarely in an ecological application are all pertinent relationships represented adequately, if at all, by empirical data. In these instances machine learning and expert knowledge can be used to quantify these system relationships with probabilities. Many elicitation methods are available, but little guidance exists about how to choose between them or the biases they may introduce. We used our need for expert probabilities as an opportunity to informally trial several extant techniques. After introducing our domain, we describe the methods used, and our observations of participant responses.

*Current address: CSIRO Sustainable Ecosystems Davies Laboratory, Townsville, QLD, 4814, Australia.

2 THE ECOLOGICAL DOMAIN

The effect of land-based activities on marine ecosystems is a matter of global concern (GESAMP 2001). With the recognition of these persistent problems also comes acknowledgement that they cannot be properly managed without understanding the interdependencies that exist between marine and land-based systems (GESAMP 2001). This is equally true for coastal lands draining to the GBRWHA, which extends 2,000 km along the coast (Brodie et al. 2001a). The GBRWHA contains approximately 3,000 reefs, large areas of seagrass and inshore mangrove forests (Brodie et al. 2001a).

The region shown in Figure 1 is primarily agricultural, covering approximately 410,000 km^2 of land (Rayment 2005) draining directly into the Great Barrier Reef lagoon. Agricultural runoff containing soil, nutrients and chemicals drains from catchments into rivers which discharge into the GBRWHA. Elevated turbidity and nutrients levels have been measured in river plumes extending from many river mouths into the lagoon (Devlin et al. 2001b, Brodie et al. 2001b, Furnas 2003), however direct linkages between river water quality and the health of GBR ecosystems remain difficult to establish (Crossland et al. 1997).

Seagrasses are among the most productive ecosystems in the world (Duarte & Chiscano 1999). The global

Figure 1: Catchments of the Great Barrier Reef World Heritage Area, indicating the study catchment.

ecosystem services provided by seagrasses have been valued at US$3.8 trillion per year (Costanza et al. 1997). Seagrasses provide connectivity between mangroves and reefs (Mumby et al. 2004), habitat and nursery areas for algae, invertebrates and fish (Heck Jr. et al. 2003), and are the primary food source of sea turtles and dugong (Marsh et al. 1999, Aragones et al. 2006). Dugong and sea turtles are vulnerable to extinction globally (IUCN 2000) and their protection in the GBRWHA is a condition that must be met to maintain World Heritage listing.

Threatened species can be conserved if the ecosystems they use for food and shelter are protected. Ecological risk analysis can help identify the biophysical factors and processes that maintain or threaten the health of these ecosystems (Hart et al. 2006). However, ecological knowledge is notoriously insufficient for most ecological risk analysis applications. This is particularly true in Australia, where landscapes are vast relative to the resources available to observe them. Subjective probability assessments are a critical data source to fill these gaps.

3 PREPARATION

The difficulties of BN graph-building in the absence of substantial practical guidance has been acknowledged in the literature (Neil et al. 2000). However, valuable contributions to the development and communication of a coherent ecological BN methodology are increasing (e.g. Cain 2001, Ticehurst et al. 2007). In particular the Quantitative Knowledge Engineering of Bayesian Networks (Q-KEBN) methodology (Woodberry et al. 2004, Pollino et al. 2005) provides a broad framework for parameterising and evaluating BNs. Recent research has seen the development of a new framework for structural elicitation, and the extension of the parameter estimation and evaluation phases of the Q-KEBN method (Thomas et al. 2005, Thomas 2008).

The new framework was applied as follows. A tiered bottom-up approach was used to simplify a complex descriptive model to a smaller, more focused model of roughly half the size. The process worked through a rough hierarchy of system specificity (primary, secondary and tertiary factors controlling seagrass ecology) to create a graphical model of the system. Graphical modelling was followed by a phase of explicit simplification, then a phase of critical review and verification. The simplified model provided a starting point for parameterisation and refinement tasks. Automated methods were not used to learn the network structure. Once the qualitative structural characteristics were identified, relationships were quantified and parameterised (Thomas 2008).

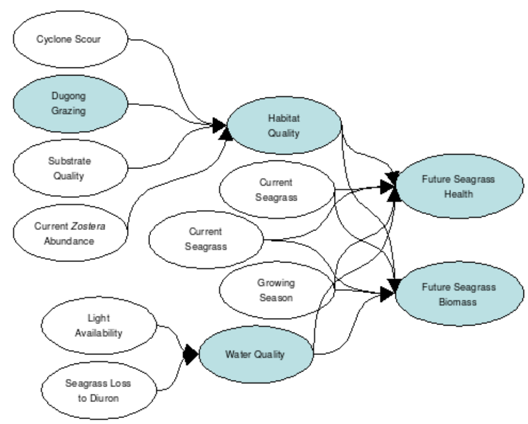

Figure 2: Seagrass health and abundance submodel

Six seagrass experts were invited to provide subjective probabilities over nodes relating directly to their area of expertise (seagrass ecology). These experts had participated in structural elicitation workshops and were familiar with the BN domain. Details of the interviewing process and examples of the Verbal Elicitor software (Hope et al. 2002) and probability assessment worksheets supplemented the invitation to participate, and three experts accepted and participated in the interviews. Research shows that three to five good quality experts are often sufficient for similar, forecasting, tasks (Clemen & Winkler 1999).

Research also shows that if experts are made aware of potential biases, and are provided with training and feedback, the incidence of bias is likely to be reduced (Kahneman et al. 1982, Merkhofer 1987, List 2001). Accordingly, seagrass experts were provided with background material describing how heuristics and biases can influence subjective judgment. Materials were also provided that described and placed nodes in the context of the wider BN, and explained concepts of causal interaction and independence that were relevant to later CPT partitioning tasks. Experts were allowed approximately two weeks to digest and, if required, clarify the material before committing themselves to the elicitation process. Experts were interviewed once, individually, in private meeting rooms at or near their workplaces. All experts were interviewed by the same person.

Training sessions were used at the beginning of each interview to familiarise experts with BN concepts and allow them to experiment with all response formats. Training began with a brief explanation of BN concepts and components. The Animals BN (Norsys 2007) and a domain-relevant BN called Simple Eutrophication (Webb unpubl.) were used to demonstrate how BNs work. The Animals BN is a simple, qualitative school-level animal classification network and the Simple Eutrophication BN is a scientific algal bloom generation network – a context familiar to the experts. Experts used these BNs to test run all formats except the freehand sketch.

'Test runs' started with a two-parent node from Animals, but the CPTs became progressively more complex as the training continued, moving to the Simple Eutrophication BN. Experts were provided with an example of each elicitation format. Parent and child node details on these examples had been completed by the knowledge engineer prior to training, ensuring that all experts were trained on the same information. The response areas on the forms had been left blank. The expert was given a copy of each format and for the first test run they completed each form with as much assistance as they requested. Subsequent forms were provided for remaining examples, and the amount of assistance was progressively reduced until the experts were confident enough to use each format unassisted.

The efficacy of the training in reducing bias could not be measured for practical reasons. Similarly cost and practicality issues prevented feedback being provided to experts about the accuracy of their subjective judgments. Each interview took up to eight hours, with breaks provided every two hours.

4 TOOLS

Five nodes required subjective probabilities to be provided by experts. Three nodes (Future Seagrass Biomass, Future Seagrass Health, Dugong Grazing) had a parent node that also required subjective assessment. Probabilities for these three nodes were elicited last, ensuring that experts thoroughly understood the parent variables prior to specifying associated child node probabilities.

Experts were encouraged to complete as many probability assessments as possible. To facilitate this, the coding effort required from experts was reduced using four strategies, presented below.

1. Start with simpler nodes and work up to more complex assessments. Effort was made to simplify the range of state combinations (i.e. the size of the CPT) of the first nodes that were elicited, so experts did not become overwhelmed by the time they came to assess the two critical, and complex, endpoint nodes late in the day. To further insure that endpoint nodes received sufficient assessment, a rough guide to the amount of time that could be spent on each node was provided.

2. Reduce the number of assessments to those lying in critical areas of the distribution. This was achieved either by eliciting the best, worst and moderate cases or the 10th, 50th and 90th percentile regions of the distribution before gathering everything else in between, or by directing the experts to complete assessments for the cases they felt most confident about before contemplating more difficult assessments.

3. If it became clear that an expert could not complete the task within the session, the most difficult parent state combinations were set aside entirely and one child state was omitted from the remaining assessments. Omitted child states were later completed using a simple default rule requiring that the probabilities of the child states sum to one.

4. To provide flexibility in probability coding and response tasks, five different response formats were provided. Prior to training, each format was first explained. Training began with small and conceptually simple nodes.

   Experts could use any of the five response formats provided. The formats used were:

- graph paper for sketching probability distributions and associated parameters. Domain experts sketch the distribution they believe best represents the parent-child relationship, indicating pertinent parameter values where appropriate (e.g. mean, maxima).

- the Verbal Elicitor software (Hope et al. 2002; Figure 3). This software, based on work in van der Gaag et al. (1999), allows entry of probability values in ordinary English. The domain expert makes qualitative assessments using a scale with numerical and verbal anchors, by selecting a verbal cue such as 'unlikely' or 'almost certain'. The associated numerical probabilities are either set manually or optimised to minimise probabilistic incoherency.

- text-scale worksheet (Figure 4). This method is adapted from van der Gaag et al. (1999). The knowledge engineer reads aloud the description of the parent-child state combination. The expert circles the preferred verbal or numeric anchor, or slashes the scale axis at a preferred point along it.

- matrix worksheet (Figure 6). This method is adapted from Laskey & Mahoney (2000). Domain experts complete the full series of dependent variable state responses, given information provided on the conditioning variables. A copy of the same verbal-numeric scale used in the above format was provided to enable choice between verbal and numeric responses.

- partitioned conditional probability table matrices (Figure 5). Domain experts identified conditioning variable states that would not change the value of the dependent variable response. A copy of the same verbal-numeric scale used in the above format was provided to enable choice between verbal and numeric responses.

Experts could change response formats between nodes but were discouraged from changing from one format to another in the middle of a node assessment. Experts were given the choice of response format only for nodes with less than three parents. For nodes with more than three parents CPT partitioning was used. Experts were encouraged to use verbal-numeric responses across all formats but were never constrained to do so.

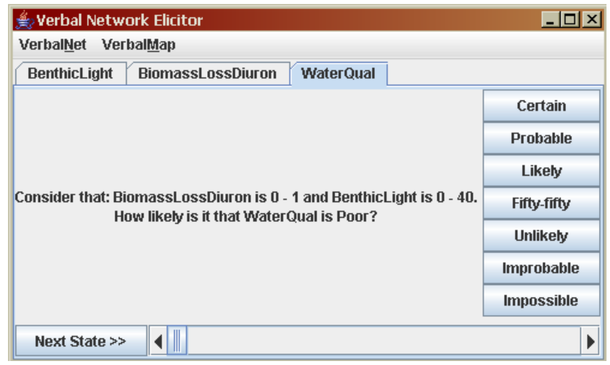

Figure 3: Screenshot from the Verbal Elicitor software (Hope et al. 2000)

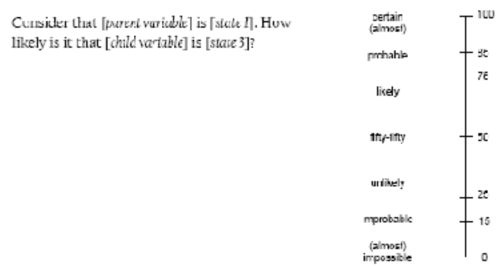

Figure 4: Extract from the text-scale worksheet (adapted from van der Gaag et al. 1999).

Cyclone Scour?	Substrate Quality	Zostera Biomass	% Lost to Dugong	Habitat Quality Poor	Moderate	Good
yes	low	0 to 5	0 to 10	probable	improbable	impossible
yes	low	0 to 5	10 to 90	probable	improbable	impossible
yes	low	0 to 5	90 to 100	probable	improbable	impossible
yes	low	5 to 15	0 to 10	probable	improbable	impossible
yes	low	5 to 15	10 to 90	probable	improbable	impossible
yes	low	5 to 15	90 to 100	probable	improbable	impossible
yes	low	15 to 100	0 to 10	probable	improbable	impossible
yes	low	15 to 100	10 to 90	probable	improbable	impossible
yes	low	15 to 100	90 to 100	probable	improbable	impossible
yes	high	0 to 5	0 to 10	probable	improbable	impossible
yes	high	0 to 5	10 to 90	probable	improbable	impossible
yes	high	0 to 5	90 to 100	probable	improbable	impossible
yes	high	5 to 15	0 to 10	probable	improbable	impossible
yes	high	5 to 15	10 to 90	probable	improbable	impossible
yes	high	5 to 15	90 to 100	probable	improbable	impossible
yes	high	15 to 100	0 to 10	probable	improbable	impossible
yes	high	15 to 100	10 to 90	probable	improbable	impossible
yes	high	15 to 100	90 to 100	probable	improbable	impossible
no	low	0 to 5	0 to 10			
no	low	0 to 5	10 to 90			
no	low	0 to 5	90 to 100			
no	low	5 to 15	0 to 10			
no	low	5 to 15	10 to 90			
no	low	5 to 15	90 to 100			
no	low	15 to 100	0 to 10			
no	low	15 to 100	10 to 90			
no	low	15 to 100	90 to 100			
no	high	0 to 5	0 to 10			
no	high	0 to 5	10 to 90			
no	high	0 to 5	90 to 100			
no	high	5 to 15	0 to 10			
no	high	5 to 15	10 to 90			
no	high	5 to 15	90 to 100			
no	high	15 to 100	0 to 10			
no	high	15 to 100	10 to 90			
no	high	15 to 100	90 to 100			

Figure 5: The CPT for the Habitat Quality node. A partition over the Cyclone Scour node is indicated with double lines

5 OBSERVATIONS ON ELICITATION PROCESSES

5.1 Biases in approach selection

Zimmer (1984) claims that different presentation modes put different emphasis on different areas of the problem-space, and Windschitl and Wells (1995) show that verbal expressions of uncertainty are more affected by presentation format than are numeric expressions. Our observations appear to support this, because in our study the response format appeared to play a role in probability elicitation results. No single format was collectively favoured by the experts over the others. Interestingly, the option to sketch the node's probability distribution was never taken up by experts during elicitation. This might indicate that familiarising experts with training materials before elicitations begin has some benefit. However, the effect of training on format preference was not tested in this study so we cannot be sure.

Text-scale worksheets and the VE software were generally preferred in both the training sessions and during elicitations of simple BN nodes. As node relationships became more complex, experts tended to prefer matrix-style formats and were eventually constrained to CPT partitioning formats for the final two, complex, nodes (Future Seagrass Biomass and Future Seagrass Health). Overall, one expert preferred the VE format and one preferred the text-scale format using verbal responses. The third expert preferred the matrix worksheet using numeric responses, stating that scientists were more accustomed to receiving information in numeric/matrix rather than verbal/text formats.

During training assessments with non-matrix formats (using VE and the text-scale worksheet) some experts showed a tendency to prefer positive cues. Answers were often bunched at the upper end of the verbal scale, with experts showing preferences for cues such as 'likely' and 'probable', and avoiding cues such as 'unlikely' or 'improbable', even though they were attempting to represent small probabilities.

The pattern was not clearly observed during assessments using probability matrix formats, indicating that the assessment format may influence experts' probability allocations. However, when reminded that parent state combinations presented to them were just scenarios of possible system responses, experts were able to refocus their assessment on the child state again, usually resulting in a different assessment value.

If during an elicitation we noticed the expert having difficulty allocating probabilities coherently, we tried using a budget metaphor to explain how probabilities needed to be distributed in the CPTs. Participants were told they effectively had 100 probability units for every parent instantiation. It was explained that this was like a budget that needed to be completely allocated into all available child states, with the largest number of units going to the best (most likely) child state choice for that parent instantiation. This appeared to clarify for the expert the problems that ensue under/over-specification, if the probability budget is not balanced appropriately. This usually happened during elicitation of the larger CPTs.

Subsequent to these discussions, we observed two things; 1) probability assessments were completed faster and with reduced under/over-specification error, and 2) experts became more inclined to use matrix formats. When matrix formats were adopted in this way, the expert's mode of assessment changed from one of sequential consideration of individual parent-child instantiations to a system where they considered *sets* of conditioning parent states to contextualise and iteratively re-calibrate their child node assessments on the fly. The experts appeared to first roughly rank instantiations against the available child states then allocate or calibrate individual probability allocations accordingly. In this sense the experts appeared to be mentally creating their own CPT partitioning systems to reduce the cognitive burden of large elicitation tasks.

This change of approach resulted in substantially fewer instances of what we suspect to be a positivity bias (described in following sections). These reductions were observed in both verbal and numeric response types. It is interesting to note that although one expert initially continued to use verbal responses when switching from a text-scale to a matrix format, once s/he started ranking responses as relative probabilities across the child node, numeric responses were preferred for the remainder of the interview. These observations indicate that provision of greater context may improve probability estimation. Development of interactive online tools or better utilisation of the BN GUI itself may help participants actively reorganise/rank CPTs and may be a good first step towards testing these observations more closely.

5.2 Quantifier effect

Verbal and numeric expressions of quantifiers (*few, not all, some*) and probabilities contain rhetorical and perspectival information (Moxey & Sanford 2000). Consequently, subtle but powerful information can be communicated and so can influence the inferences and responses of readers.

Moxey & Sanford go on to propose that negative quantifiers like *not all* put a different perspective on the interpretation of events, which can affect the value judgment placed on the outcome.

They give the following example:

"**(10)** There is a small probability of death, which is a good*/bad thing.

(10') It is improbable that anyone will die, which is a good/bad* thing.

(11) There is a small risk of death, which is a good*/bad thing.

(11') There is an insignificant risk of death, which is a good/bad* thing."

In this example an asterisk denotes an unacceptable response. Moxey & Sanford (2000) suggest that negative quantifiers invite the reader or listener to presuppose that things are more probable or risky than they actually are. This pattern is consistent regardless of how much confidence is being expressed (e.g. *not quite certain* vs. *small probability*; Moxey & Sanford (2000)).

If the phrasing of conditioning statements can affect the perspectives and inferences of the participants, then the reasoning processes they use to generate probability assessments are also likely to be affected. This may have implications for BN knowledge engineers, because the example above is an inverse representation of the kind of conditioning statements used in probability elicitation for BNs. Rephrased as a BN elicitation query of the type used in recent research, the statement could read something like;

> "If blood alcohol level is low and the speed of the car is low, the probability of death is ____."

with participants required to complete the statement with the most accurate of the probability expressions offered. It is difficult to differentiate instances of the quantifier effect from positivity bias, which is discussed in the following section. Examples of possible instances of the quantifier effect are described in the following section on positivity bias.

5.3 Positivity bias

Teigen & Brun (2003) have shown that participants choose verbal probability phrases to correspond with the linguistic rather than the numeric content of presented information. Their experiments show that sentences containing positive quantifiers – phrases with positive directionality – will tend to receive positively framed responses, indicating that probability estimates are influenced by the way the conditioning information is presented.

Participants choose verbal phrases as a function of their frame; if they want to affirm that a particular outcome could in fact occur then they will use a term with positive directionality (e.g. 'possible', at the upper end of the verbal scale) but if the purpose is to draw attention to an events non-occurrence then a negative phrase (e.g. 'improbable', at the lower end of the scale) will be chosen (Teigen & Brun 2003). This may be because the phrases used in the text-scale worksheets and the VE software both request participants to determine "how likely" a certain response is given certain conditions. The word 'likely' creates a positive frame for the parent-child state combination requiring assessment. Positive frames may encourage positivity bias; a general readiness of participants to prefer positive over negative descriptive terms, as if positivity is the rule and negativity must be treated as an exception (Teigen & Brun 1995).

It may be possible to reduce positive framing by omitting the word 'likely' and presenting the parent-child state combination as a factual statement against which the expert applies a probability;

> "When [parent node 1] is in [state 1] and [parent node 2] is in [state 1], [child node] is [state 3]. What is the chance that this is true?"

This will be tested in future case studies. To our knowledge the presence of positivity bias in probability elicitation for BNs has not been tested directly, and may not be adequately controlled for in extant BN elicitation techniques or formats. However our observations may provide some evidence that these biases can be reduced.

For example, matrix formats (Figures 5 and 6) present experts with the entire set of parent-child state combinations all at once. So instead of considering each parent-child state combination in isolation, experts can choose to view sets of conditioning (parent) states, including the full range of possible responses (child) states across which the entirety of the 'probability budget' must be allocated. This has the advantage of making the assessment context explicit. Matrix formats may therefore provide a mechanism for participants to frame assessment requests more broadly.

5.4 Overconfidence/uncertainty avoidance

Although the apparent positivity bias was fairly easily observed, we believe a different bias was also observed during elicitations. Some experts expressed aversion to the absoluteness of the words 'certain' and 'impossible' because, in the words of one expert "nothing is certain in ecology". However, when allowed to use numeric probabilities, the same expert still responded with 1 (certain) and 0 (impossible) values. Expressions of absolute certainty were also common in verbal responses. The result suggests that these experts may also be displaying overconfidence. There were many probabilities at the very high or very low end of the spectrum (near 1 or 0), and instances of complete disagreement were observed; where one expert assessed a parent-child state combination as 'certain' and a different expert assessed the same combination as 'impossible' (divergences between experts' assessments are discussed in more detail in the following section). This observation is similar to that described by Keren & Teigen (2001) as 'the principle of definitive predictions', or 'uncertainty avoidance', where only extreme probabilities are used in responses because participants wish to appear quite clear about what will happen next.

5.5 Aggregating expert results

Subsequent to the elicitation processes described in this paper, we aggregated (averaged) subjective probabilities and evaluated the responses in two ways. Divergences between experts' probability assessments were analysed using the relative standard deviation of the average value and the Bhattacharyya distance measure (Bhattacharyya 1943).

These techniques allowed the experts, nodes and node elements for which disagreement occurred most strongly to be identified. This is necessary so that conflicting assessments can be investigated collaboratively with participants to resolve whether the causes are clerical errors or mismatching assumptions about context – in which case the distances could be expected to diminish, or if the cause of the differences is due to contrasting conceptual models among experts – in which case structural modificatifation may be required and parallel models developed.

Further, the technique showed that although each expert provided different distributions, differences across experts occurred in equal measure. A demonstrated lack of systematic bias among experts indicated that averaging was an appropriate aggregation technique. Details of this research are reported in Thomas (2008).

6 CONCLUDING REMARKS

There is currently little guidance about how to choose between subjective elicitation methods. Preferences and responses of ecological managers and scientists were informally field-trialled using a selection of probability elicitation formats. Expert's format preferences appeared to be influenced by their familiarity with the format and the complexity of the elicitation problem. Our observations indicate that none of the trialled techniques are likely to be completely impervious to bias and overconfidence. Positively framed text-based descriptions of parent-child state combinations may have contributed to the observed bias.

Acknowledgements

This work was conducted with a scholarship from the Monash University Water Studies Centre. We gratefully acknowledge assistance from the Great Barrier Reef Marine Park Authority and the Australian Centre for Tropical Freshwater Research at James Cook University. The authors would like to thank Dr Carmel Pollino for her input and our anonymous experts for their time and good faith.

References

Aragones L.V., Lawler I.R., Foley W.J. & Marsh H. 2006. Dugong grazing and turtle cropping: grazing optimization in tropical seagrass systems? Oecologia 149(4): 635-647.

Bhattacharyya A. 1943. On a measure of divergence between two statistical populations defined by their probability distributions. Bulletin of the Calcutta Mathematical Society 35: 99-110.

Brodie J., Christie C., Devlin M., Haynes D., Morris S., Ramsay M., Waterhouse J. & Yorkston H. 2001a. Catchment management and the Great Barrier Reef. Water Science and Technology 43(9): 203-211.

Brodie J., Furnas M., Ghonim S., Haynes D., Mitchell A., Morris S., Waterhouse J., Yorkston H., Andas D., Lowe D. & Ryan M. 2001b. Great Barrier Reef Catchment Water Quality Action Plan. Townsville, Great Barrier Reef Marine Park Authority.

Brodie J., Death G., Devlin M., Furnas M. & Wright M. 2007. Spatial and temporal patterns of near-surface chlorophyll a in the Great Barrier Reef lagoon. Marine and Freshwater Research 58(4): 342-353.

Cain J. 2001. Planning improvements in natural resources management: guidelines for using Bayesian networks to support the planning and management of development programmes in the water sector and beyond. Wallingford, Oxford, Centre for Ecology & Hydrology. 132pp

Clemen R. & Winkler R.L. 1999. Combining probability distributions from experts in risk analysis. Risk Analysis 19(2): 187-203.

Costanza R., d'Arge R., de Groot R., Farber S., Grasso M., Hannon B., Limburg K., Naeem S., O'Neill R.V., Paruelo J., Raskin R.G., Sutton P. & van den Belt M. 1997. The value of the world's ecosystem services and natural capital. Nature 387: 253-260.

Crossland C.J., Done T.J. & Brunskill G.J. 1997. Potential impacts of sugarcane production on the marine environment. In: Intensive Sugarcane Production: Meeting the Challenges Beyond 2000. (Eds.) B.A. Keating & J.R. Wilson: 423-436pp.

Duarte C.M. & Chiscano C.L. 1999. Seagrass biomass and production: a reassessment. Aquatic Botany 65(1-4): 159-174.

GESAMP 2001. Protecting the oceans from land-based activities: Land-based sources and activities affecting the quality and use of the marine, coastal and associated freshwater environment. Rep. Stud. GESAMP No. 71, IMO/FAO/UNESCO-IOC/WMO/WHO/IAEA/UN/UNEP Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection and Advisory Committee on Protection of the Sea. 162pp

Hart B.T., Burgman M., Grace M., Pollino C., Thomas C. and Webb J.A. 2006. Risk-based approaches to managing contaminants in catchments. Human and Ecological Risk Assessment 12: 66-73.

Heck Jr. K.L., Hays G. & Orth R.J. 2003. Critical evaluation of the nursery role hypothesis for seagrass meadows. Marine Ecology Progress Series 253: 123-136.

Hope L.R., Nicholson A.E & Korb, K.B. 2002. Knowledge Engineering Tools for Probability Elicitation, School of Computer Science and Software Engineering, Monash University, Technical report 2002/111.

IUCN 2000. The 2000 Red List of Threatened Species. Gland, Switzerland, IUCN.

Kahneman D., Slovic P. & Tversky A., Eds. 1982. Judgement under Uncertainty, Heuristics and Biases. Cambridge, Cambridge University Press.

Keren G. & Teigen K.H. 2001. Why is p = .90 better than p = .70? Preference for definitive predictions by lay consumers of probability judgments. Psychonomic Bulletin and Review 8: 191-202.

Laskey K.B. & Mahoney S.M. 2000. Network engineering for agile belief network models. IEEE Transactions on Knowledge and Data Engineering 12(4): 487-498.

List J.A. 2001. Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards. The American Economic Review 91(5): 1498-1507.

Marsh H., Eros C., Corkeron P. & Breen B. 1999. A conservation strategy for dugongs: implications of Australian research. Marine and Freshwater Research 50: 979 – 990.

Merkhofer M.W. 1987. Quantifying judgmental uncertainty: methodology, experiences and insights. IEEE Transactions on Systems, Man and Cybernetics 17(5): 741-752.

Mumby P.J., Edwards A.J., Arias-Gonzalez J.E., Lindeman K.C., Blackwell P.G., Gall A., Gorczynska M.I., Harborne A.R., Pescod C.L., Renken H., Wabnitz C.C.C. & Llewellyn G. 2004. Mangroves enhance the biomass of coral reef fish communities in the Caribbean. Nature 427: 533-536.

Neil M., Fenton N. & Nielson L. 2000. Building large-scale Bayesian networks. The Knowledge Engineering Review 15: 257-284.

Norsys Inc. (2007). Netica, http://norsys.com.

Pollino C.A., Woodberry O., Nicholson A.E. & Korb K.B. 2005. Parameterising Bayesian networks: a case study in ecological risk assessment. (Eds.) V. Khacitvichyanukul, U. Purintrapinban and P. Uthayopas Proc. of the 2005 Int. Conf. on Simulation and Modelling (SIMMOD05), Thailand.

Rayment G. 2005. Northeast Australian experience in minimizing environmental harm from waste recycling and potential pollutants of soil and water. Communications in Soil Science and Plant Analysis 36(1-3): 121-131.

Moxey L.M. & Sanford A.J. 2000. Communicating quantities: a review of psycholinguistic evidence of how expressions determine perspectives. Applied Cognitive Psychology 14(3): 237-255.

Starck W. 2005. ‘Threats’ to the Great Barrier Reef. IPA Backgrounder 17(1).

Teigen K.H. & Brun W. 1995. Yes, but it is uncertain: Direction and communicative intention of verbal probabilistic terms. Acta Psychologica 88(3): 233-258.

Teigen K.H. & Brun W. 2003. Verbal probabilities: a question of frame? Journal of Behavioral Decision Making 16(1): 53-72.

Thomas C.R., Hart B.T., Nicholson A.E., Grace M. & Pollino C.A. 2005. Development Of Criteria For Simplifying Ecological Risk Models (Eds.) A. Zerger & R.M. Argent Proc. of MODSIM 2005 Int. Congress on Modelling and Simulation.

Thomas C.R. 2008. Bayesian Belief Network Development for Ecological Decision Support in Data-Sparse Domains. PhD thesis, Water Studies Centre Monash University, Melbourne. 436pp.

Ticehurst J.L., Newham L.T.H., Rissik D., Letcher R.A. & Jakeman A.J. 2007. A Bayesian network approach for assessing the sustainability of coastal lakes in New South Wales, Australia. Environmental Modelling & Software 22(8): 1129-1139.

van der Gaag L, Renooj S., Witteman C.L.M., Aleman B.M.P., & Taal B.G. 1999. How to elicit many probabilities. In UAI99 – Proc. of the Fifteenth Conf. on Uncertainty in Artificial Intelligence, Laskey and Prade, Eds., pp. 647-654.

Webb, A. 2003. Unpublished data, Water Studies Centre, Monash University.

Windschitl P. & Wells G. 1995. Measuring psychological uncertainty: verbal versus numeric methods. Journal of Experimental Psychology: Applied 2: 343-364.

Woodberry O.G., Nicholson A.E., Korb K.B. & Pollino C. 2004. Parameterising Bayesian networks. In: Lecture Notes in Computer Science. (Eds.) G.I. Webb & Y. Xinghuo. Cairns, Australia, AI2004: Advances in Artificial Intelligence: 17th Australian Joint Conf. on Artificial Intelligence: p 1101-1107.

Zimmer A.C. 1984. A model for the interpretation of verbal predictions. International Journal of Man-Machine Studies 20: 121-134.

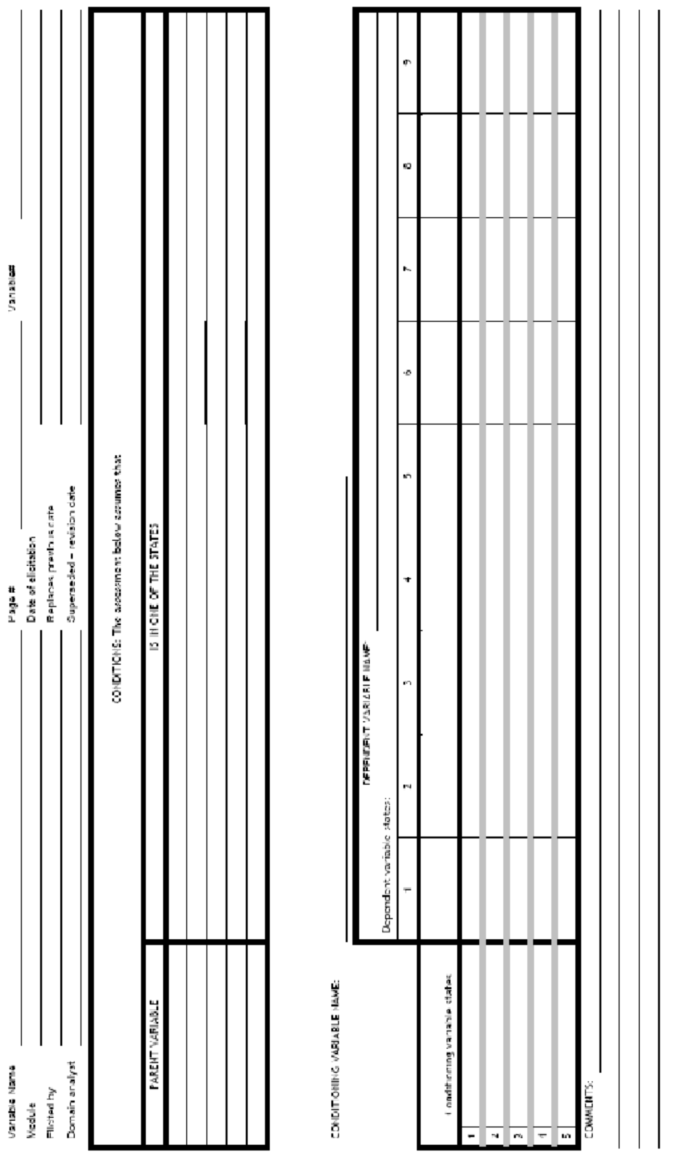

Figure 6: Example matrix worksheet, adapted from Laskey & Mahoney (2000).