# ABC-BN: A tool for building, maintaining and using Bayesian networks in an environmental management application

## Ann E. Nicholson, Owen Woodberry, Steven Mascaro, Kevin Korb

Bayesian Intelligence Pty Ltd. 2/21 The Parade, Clarinda VIC, 3169, Australia

## Abstract

The Victorian state government Department of Sustainability and Environment has a webbased application called Actions for Biodiversity Conservation (ABC) as a central resource for managing over 400 threatened species and communities. ABC maintains information for species and communities at individual locations, including lists of threats (to species and communities), actions (to mitigate threats) and population and habitat factors. Here we describe an extension to ABC, a tool for building Bayesian networks for selected populations of threatened species and occurrences of threatened communities, to model the interactions between actions, threats and population and habitat factors. The tool, called ABC-BN, also allows users to do what-if scenario reasoning, and is integrated with ongoing monitoring and reporting. Unlike most ecological BN modelling to date, which typically takes months or years to produce a single, often complex, BN, for a specific problem, the aim here was to produce a tool that would facilitate the development, maintenance and use of a large number of quite simple standardized BNs, specialized for particular instances. We describe the incremental development of ABC-BN over 3 years. A prototype version was evaluated in 2009 by building models for 100 species, and the completed ABC-BN was deployed in April, 2011.

## 1 INTRODUCTION

The Victorian state government Department of Sustainability and Environment (DSE) has a web-based application called Actions for Biodiversity Conserva-

## Adrian Moorrees, Alicia Lucas

Biodiversity and Ecosystem Services 2/8 Nicholson Street East Melbourne, VIC 3002, Australia

tion (ABC) as a central resource for managing threatened species and communities (DSE, 2009). It facilitates the management of actions documented in Action Statements prepared under the Flora and Fauna Guarantee Act 1988 and Recovery Plans prepared under the Environment Protection and Biodiversity Conservation Act 1999. ABC currently holds information on more than 400 threatened species and communities<sup>1</sup> and over 8,000 management actions at approximately 2000 locations across Victoria. ABC maintains information for species and communities at individual locations, including lists of threats (to species), actions (to mitigate threats) and population and habitat factors. By bringing together information from a range of sources, ABC is intended to make significant improvements in both knowledge about threatened species and communities and the transfer of that knowledge. The recording of actions implemented for populations and communities enables the development of prioritized lists of actions to be made on an increasingly sound basis by land and water managers. Reporting on outcomes provides a basis for applying adaptive management, whereby the effectiveness of management can be improved based on the current and previous outcomes for a population. By 2004, ABC Stage 1 was operational and able to report on whether an action to ameliorate a threat had been carried out. ABC Stage 2 (deployed in 2006), incorporated the facility to prepare actions plans within the system. However ABC did not support any modelling of interactions between actions, threats and outcome factors and had no predictive capability. More importantly, it was not able to report on the status of a species or community for which there was little or no data collected, in a way that was comparable across the state.

Bayesian networks (BNs) (Pearl, 1988; Jensen and Nielsen, 2007), are becoming increasingly popular for

<sup>&</sup>lt;sup>1</sup>The FFG Act 1988 defines a community as "a type of assemblage which is wholly or substantially made up of taxa of flora or fauna existing together in the wild".

environmental and ecological monitoring and risk assessment (see § 5.2.3 in Korb and Nicholson, 2010 for a recent survey). There have been a number of modelling guidelines published (e.g., Varis and Kuikka, 1999; Borsuk et al., 2004; Renken and Mumby, 2009), while Uusitalo (2007) reviews their features and use in modelling environmental applications. In 2008, DSE decided to extend ABC with BN technology, to provide the capacity to model the interactions between actions, threats and outcomes and overall status, and allow users to do what-if predictive and diagnostic scenario reasoning. These additional capacities would be integrated with ongoing monitoring and reporting.

In most ecological BN modelling to date the knowledge engineering paradigm is to develop a single, often complex BN, for a specific problem, which typically takes months or years to complete (e.g., Pollino et al., 2007; Smith et al., 2007; Chee et al., 2005). Here, in contrast, the aim was to produce a tool that would facilitate the development, maintenance and use of BNs for specific populations and occurrences of species and communities with a high priority within ABC, rather than entire species and communities.

This aim led to some key design decisions. First, the tool would be for use by the DSE so-called monitors – DSE scientists and Biodiversity Officers – already maintaining information in ABC about the species for which they were responsible. The BN technology would be hidden from these users. Second, the tool would support the building of simple BNs using a template that enforced a certain causal structure between actions, threats and outcomes.

In this paper, we describe the incremental development of the tool, called ABC-BN, over a 3 year period, with the challenges of changes to the underlying BN template, and a number of substantial additions to the tool requirements. We give an overview of the ABC-BN's functionality, which is based on the iterative, incremental knowledge engineering of BNs Laskey and Mahoney (2000) A prototype version was evaluated in 2009 by building models for 100 species, and the completed ABC-BN was deployed in April 2011.

## 2 ACTIONS FOR BIODIVERSITY CONSERVATION (ABC)

The first version of ABC was developed for DSE by Spatial Vision, a software company specialising in GIS applications. Stage I became operational in 2004, with the Stage II in place from 2006. The addition of ABC-BN, to be developed by Bayesian Intelligence, became part of Stage III, along with enhancements to ABC itself, to be delivered by Spatial Vision.

ABC is a web-based interface to a database application. It is based around entities called items, which may be flora or fauna species, communities, or potentially threatening processes. Within ABC, there are different types of users, with different responsibilities and access levels. ABC is based on recording information about each item at each location (linked with a GIS) it is found; there is a **location monitor** responsible for recording the management and monitoring. The item monitor oversees the information about each item, while system administrators have additional responsibilities such as appointing monitors and adding actions and threats to the global lists. Finally, there are various stakeholders (e.g., members of other government organizations, NGOs, community groups, academics and researchers), who can be given access to the information in ABC, but may not make changes.

For each item, the ABC database contains information about the **threats** that have been identified at each location, which are ranked along two dimensions, ""likelihood" and "impact", representing priorities. It records actions (from a central list of possible actions) that have been identified as potentially useful to reduce threats to the item at that location. Finally, monitors must record monitored information in actions taken, the status of threats, and various **outcomes**, which are observations relating to the species populations or community occurrences, or the habitat which can be monitored. This information facilitates the prioritizing of species and communities based on the importance of the location and contributions of actions to mitigate threats.

#### 3 ARCHITECTURE

The key requirement for ABC-BN was that it be completely aligned with the existing ABC application, with the same web interface, the same types of users, and model the same action, threat and outcome information in the ABC database.

Figure 1 shows the system architecture for integration of ABC and ABC-BN. ABC-BN is invoked via a menu tab from ABC, with the current item as well as the user ID, passed to ABC-BN. From that point, ABC-BN interacts directly with the users via their browser, with uploads and downloads of BN models allowed via a shared DSE file system. ABC-BN is a TOMCAT application, implemented using JavaServer Pages and the NeticaJ Java API.<sup>2</sup> The BNs and ABC-BN specific information are stored in new tables added to the ABC Oracle database.

 $<sup>^2</sup>$ www.norsys.com

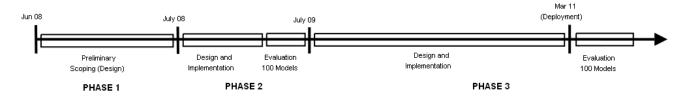


Figure 2: ABC-BN development phases

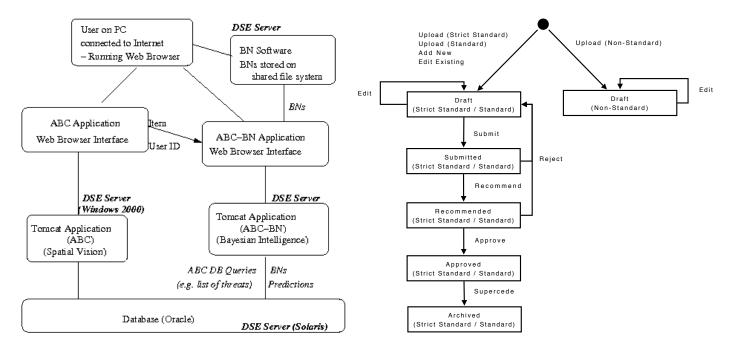


Figure 1: System architecture for the integration of ABC and ABC-BN

#### 3.1 Incremental and Iterative Development

ABC-BN was designed and built incrementally, following the prototype-based spiral software development cycle advocated by Brookes (1995) and Boehm (1988). There were three main phases (see Figure 2): a preliminary scoping phase, then two main design and development phases. These were due to (1) changes to the underlying BN template and (2) a number of substantial additions to the tool functionality.

## 4 USER ROLES AND MODEL MANAGEMENT

ABC-BN provides model management and an approval workflow, together with an audit history. The BN models are managed for each item location; there may be many draft models, but only one approved model at any one time. Models must go through a three step submission process: submitted  $\rightarrow$  recommended  $\rightarrow$  approved. When an approved model is superseded

Figure 3: Model lifecycle

(by a new approved model) it becomes **archived**. Figure 3 shows the model lifecycle.

Two user roles were added for ABC-BN: (1) **Model administrators** have privileged access to the ABC-BN module, notably approving models and the ability to edit the ABC-BN global settings, upload Netica BNs to the system, and delete any model; (2) **Outcome monitors** are nominated by the model administrators for each item location.<sup>3</sup> Outcome monitors may build and submit models for approval, while item monitors are responsible for recommending models.

Table 1 summarizes the user roles within ABC-BN., while Figure 12 shows the model management page.<sup>4</sup>

 $<sup>^3{\</sup>rm This}$  is because they may be some one other than the location monitor.

 $<sup>^4</sup>$ Figures showig screen shots of ABC-BN are placed together at the end of the paper.

Table 1: User roles

Role	Abbr.	Comments
Model Ad-	MA	Full access to the ABC-BN module.
ministrator		Able to create, submit, recommend
		and approve models.
Item Moni-	IM	Access all models. Able to submit
tor		and recommend models for locations
		under the item.
Outcome	OM	Access all models. Able to create and
Monitor		submit models for item location.
Any (Other)	M	Access all models.
Monitor		
Stakeholder	SH	Access only Approved models.

## 5 MODEL BUILIDNG

ABC-BN supports all stages of the BN elicitation process. The elicitation algorithm is shown in Table 2. Navigation around the elicitation process is flexible, the user may move to any module, and any sub-module within that module (using tab menus). Thus it provides the iterative and incremental BN construction process advocated by many (e.g., Laskey and Mahoney, 2000; Korb and Nicholson, 2010; Boneh, 2010). Figure 4 shows an example of a BN created within the system.<sup>5</sup>

#### 5.1 Structure

ABC-BN supports building models specific to each item location, based on a pre-defined structure template, incorporating actions, threats and asset factors in the ABC database. This template evolved over ABC-BN's development phases as shown in Figure 5:

- Phase 1: Template contains only threats, which are parents of a ThreatTrend node, with values {Improving, Stable, Worsening}
- Phase 2: Template contains Action nodes (all root nodes) which are parents of the Threat nodes, which are in turn parents of AssetFactor nodes, as well as two Trend nodes.
- Phase 3: Trend nodes modelling change now replaced by Status nodes representing absolute values: {Severe, Moderate, Negligible} for Threat-Status and {Good, Fair, Poor} for AssetStatus.

The elicitation process builds only BNs that follow this so-called **strict standard** template, which limits the complexity of the networks and ensures they are easily analyzed for reporting (based on the STATUS nodes).<sup>6</sup>

Table 2: Elicitation algorithm

Module	Sub-module	Algorithm
		-create standard template
		model
Structure:	Name:	-name the model
	Threat:	-select threats to be included
		in the model
	Actions	-for each threat variable
	& Asset	-select action variables affect-
		ing this threat
	Factors:	-select asset factor variables
		affected by this threat
	States:	-for all action, threat and asset
		factor variables
		-select a state space
		-if optional state description
		-give description
	Additional	-for each threat and asset fac-
		tor variable
	Links:	-select other variables, of
		the same type, affected by this variable
CPTs:		-for all threat, asset factor and
01 15.		query variables
		-for each parent state combi-
		nation
		-select verbal cues for each
		state
Scenarios:		-test model with scenarios tool
~ 501101100.		111111111111111111111111111111111111111

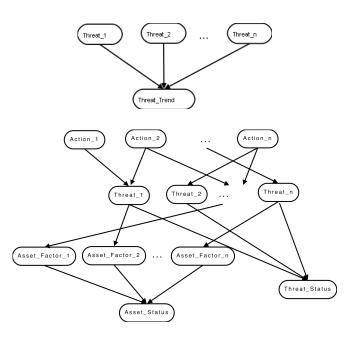


Figure 5: ABC-BN's templates from Phase 1 (above) and Phase 3 (below). The Phase 2 template was as for Phase 3, but contained Trend nodes, instead of Status nodes.

 $<sup>^5{\</sup>rm This}$  BN was produced by Phil Papas and Di Crowther of DSE, Biodiversity and Ecosystem Services, Arthur Rylah Institute.

<sup>&</sup>lt;sup>6</sup>ABC-BN also allows uploading of BNs (for example, constructed in Netica) that do not follow the strict stan-

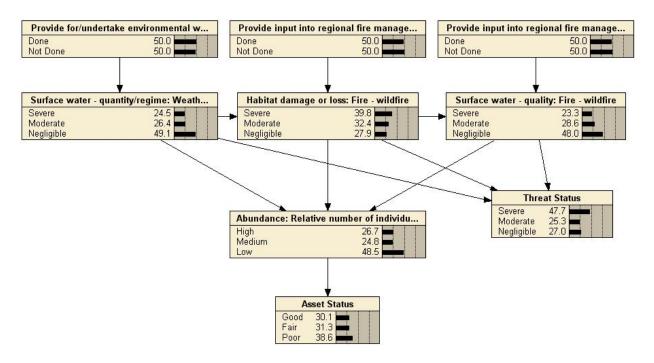


Figure 4: An example complete BN created within ABC-BN.

#### **Threat Ranking**

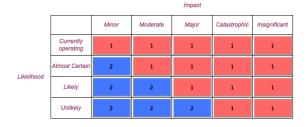


Figure 6: Threat Ranking page.

The BN nodes are selected from the values already in the database for that item and location. The discrete state space for each node is selected from a global list (intended to provide consistency across the system, and maintained by the Model Administrator), divided into "types"; for example the "Choice" type could have alternatives {Yes,No} and {Done,NotDone}, while an example "Amount" type is {High, Medium, Low}.

The user first selects threats from a list of threats for that item. Administrators can assign a threat to one of two ranks based on the impact and likelihood of a threat, (see Figure 6), which changes the way threats are presented on the Threat Selector page (see Fig-

dard: (1) standard models, which are similar to the strict standard structure, but without any constraints on the links, and may be submitted for approval; (2) non-standard models can be any Netica BN, but may not be submitted for approval, as they can't be analyzed for reporting. ure 13). Once the Threat nodes are selected, the user then selects the associated Action and Asset-Factor nodes for each (see Figure 14).

## 5.2 Parameter Elicitation

Eliciting the parameters of the BN, the conditional probabilities, is recognised as a difficult task. For each node, there is distribution for each combination of values of the parent variables; this is exponential in the number of parent variables. It can be hard for experts to express their experience/opinion in numbers, and they are often inconsistent. Hence the decision was made for ABC-BN to support the alternative qualitative assessment of probabilities using verbal anchors, developed by van der Gaag et al. (1999), and implemented in the Verbal Elicitor tool (Hope et al., 2002). The user is presented with scenario descriptions (about combination of parent values) and a selection of common chance tags, e.g., "certain", "likely" or "impossible". In this way actual probability values are not required or shown to the expert. ABC-BN contains a so-called verbal map (which is maintained by the Model Administrator) which maps verbal cues to probabilities (see Figure 15).

For Phase 2, ABC-BN provided only this qualitative elicitation, and the users could not see the probabilities stored into the CPT.<sup>7</sup> Feedback from the users during the Phase 2 evaluation indicated that this was

 $<sup>^7{\</sup>rm That}$  is, unless they downloaded the BN for viewing in Netica!

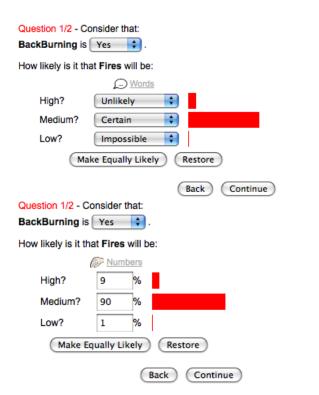


Figure 7: Example of single question parameter elicitation: (above) Verbal cue; (below) Number

quite frustrating. Also, by their very nature, the probabilities elicited using verbal tags will be inaccurate, as the cues are limited to only some probabilities, and they are nearly always changed during normalization. Hence in Phase 3 the parameter elicitation was changed to a hybrid system, combining both qualitative and quantitative elicitation, catering for an iterative stepwise refinement of the probabilities.

Basic elicitation is done via single questions, where the user can switch between "Words" and "Numbers"; an example of these alternatives, for a FIRE threat node, with a single BACKBURNING action parent node, is shown in Figure 7. In single question mode, the "Continue" option takes the user to the next combination of parent values - that is, the next row in the CPT. Note that the user is told the question is "X/Y", meaning Xth question out of Y total questions.

Feedback on the Phase 2 prototype also indicated that the users found it hard to "calibrate" their answers when answering these single questions, so Phase 3 included an alternative table interface, allowing visualisation and flexible navigation of the CPT (in either word or number form). An example is shown in Figure 8. Navigation of the table can be done by either clicking on a summary table, or by selecting a parent state combination via drop down boxes, including an

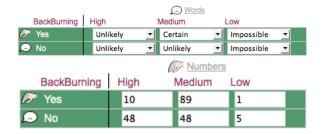


Figure 8: Example of table parameter elicitation: (above) Verbal cue; (below) Number

"any" state. Users can switch between a single question (i.e., a parent state combination) and the full table view. ABC-BN also provides visualization of the parameter elicitation progress, incorporating a summary table (not shown) with colour coding for questions that are either complete, incomplete or in need of revision (i.e., when there have been structural changes).

## 6 Reasoning in ABC-BN

ABC-BN's reasoning component (developed in Phase 2), called "Scenarios", allows the user to enter any combination of exact or likelihood evidence and provides a visualization of the posterior distributions after belief updating. The screen is layed out in four columns – Actions, Threats, Asset Factors, and Status – corresponding to the layers in the BN template, while the arcs in the underlying network are hidden is shown in Figure 9.

To enter evidence, the user clicks on a node's "Add Evidence..." button, which changes the node view to show a slider for evidence. This rectangular slider is split into as many parts as there are states, where the length of each part represents the probability of a state. The user can either drag the dividers, or enter number values, to enter the desired likelihood evidence. Alternatively, the user can simply select a state if the evidence is exact.

In Phase 3, the reasoner was extended with a similarity analysis report, which summarizes the current scenario, that is, the model and the entered set of evidence; an example is shown in Figure 10. As well as reporting the calculated probabilities for all the nodes, column 1 shows the difference between the posteriors and the initial distribution (before evidence was entered), while column 2 shows differences between the final distribution and distribution given *only* the Action node evidence.

<sup>&</sup>lt;sup>8</sup>The "any" state is useful if one parent (or a set of parents) dominate in a particular combination, rendering other parents inconsequential.

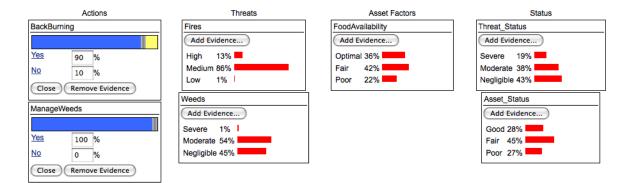


Figure 9: Scenario tester page

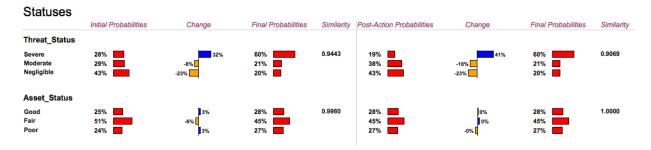


Figure 10: Similarity analysis report

#### **Count of Threat and Asset Status Outcomes**



Figure 11: Outcome summary page

#### 7 OUTCOME SUMMARY

This component was added in Phase 3, integrating the BN models with the monitoring and reporting functions of ABC. For each model, the user enters evidence for the BN, for any of the ACTION, THREAT or ASSET-FACTOR nodes, along with meta-data. This evidence is stored in the database with the BN. The model is then re-run with this evidence, and an Outcome Report is produced. This report contains a summary of the pos-

terior probabilities of the two STATUS nodes (given the evidence), plus metadata on all included factors (actions, threats and asset factors), such as the source of the information (observed data, expert or/and literature).

As well as reports on individual items, the system also produces a summary table with categorization, across a user-specified set of items; an example is shown in Figure 11. This table is an aggregation of the classification using the highest posterior probabilities for STATUS nodes computed for each BN, given the outcomes evidence added by the monitor. Results are defined as Inconclusive if the difference between the two most probable states is less than the "inconclusive threshold", a global variable (modifiable by administrators).

#### 8 EVALUATION

The prototype developed in Phase 2 was trialled in May-June, 2009 in 5 elicitation workshops around the state. It was used to generate BNs of 100 threatened species and communities. In these trials, the facilitators were DSE personnel<sup>9</sup> overseeing the development of ABC-BN, while the workshop participants were other ABC monitors. These gave very positive

<sup>&</sup>lt;sup>9</sup>The facilitators were authors Moorrees and Lucas – ABC administrators, as well as item and location monitors.

feedback, as well as suggestions for improvement that were adopted for Phase 3.

Phase 3 was deployed early April, 2011, with the first round of outcomes for 100 models due at the end of June. We intend to undertake a more formal evaluation of both the tool, and of the individual models developed by it. The tool can be evaluated by measuring the rate of elicitation (e.g. number of models, number of parameters), with qualitative feedback from the experts regarding their level of comfort using the tool. The quality of the models is harder to assess, but we intend to have a selection (10%) reviewed by experts not involved in the elicitation or approval process.

## 9 CONCLUSION

We have described ABC-BN, a tool for building, maintaining and using Bayesian networks in an existing environmental management application. In contrast to most knowledge engineering of ecological applications to date, ABC-BN supports the construction of a large number of simple, standardized BNs over a relatively short time period. It supports iterative and incremental model construction, including hybrid qualitative and quantitative parameter elicitation. It allows users to do what-if predictive and diagnostic scenario reasoning, and the BN models are integrated with monitoring data and reporting, providing outcome status reports for both individual items as well as status summaries for sets of items. Overall, the tool allows managers to estimate the costs associated with actions to mitigate threats and objectively assess the relative importance of actions for different species populations and community occurrences.

### References

- Boehm, B. W. (1988). A spiral model of software development and enhancement. *IEEE Computer*, 61–72.
- Boneh, T. (2010). Ontology and Bayesian decision networks for supporting the meteorological forecasting process. Ph. D. thesis, Clayton School of Information Technology, Monash University.
- Borsuk, M., C. Stow, and K. Reckhow (2004). A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. *Ecological Modelling* 173 (2-3), 219–239.
- Brooks, F. (1995). The Mythical Man-Month: Essays on Software Engineering (Second ed.). Reading, MA: Addison-Wesley.
- Chee, Y., M. Burgman, and J. Carey (2005). Use of a Bayesian network decision tool to manage environmental flows in the Wimmera river, Victoria. Report No. 4, LWA/MDBC Project UMO43: Delivering Sustainability Through Risk Management, University of Melbourne, Australia.

- DSE (2009). Actions for biodiversity conservation (abc): Managing our threatened species and communities. http://www.dse.vic.gov.au/DSE/nrenpa.nsf/LinkView/487AAFDF542DDAC6CA25703F001C16FEB44D7EEB86E4BDFDCA257115001408B8.
- Hope, L., A. Nicholson, and K. Korb (2002). Knowledge engineering tools for probability elicitation. Technical report 2002/111, School of Computer Science and Software Engineering, Monash University.
- Jensen, F. V. and T. D. Nielsen (2007). Bayesian networks and decision graphs (2nd ed.). New York: Springer Verlag.
- Korb, K. B. and A. E. Nicholson (2010). *Bayesian artificial intelligence* (2nd ed.). Chapman & Hall/CRC.
- Laskey, K. and S. Mahoney (2000). Network engineering for agile belief network models. *IEEE: Transactions on Knowledge and Data Engineering* 12(4), 487–498.
- Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems. San Mateo, CA: Morgan Kaufmann.
- Pollino, C., O. Woodberry, A. Nicholson, K. Korb, and B. T. Hart (2007). Parameterisation of a Bayesian network for use in an ecological risk management case study. *Environmental Modelling and Software* 22(8), 1140–1152.
- Renken, H. and P. J. Mumby (2009). Modelling the dynamics of coral reef macroalgae using a Bayesian belief network approach. *Ecological Modelling* 220 (9-10), 1305-1314.
- Smith, C., A. Howes, B. Price, and C. McAlpine (2007).
  Using a Bayesian belief network to predict suitable habitat of an endangered mammalThe Julia Creek dunnart (Sminthopsis douglasi). Biological Conservation 139, 333–347.
- Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Mod*elling 203(3-4), 312–318.
- van der Gaag, L. C., S. Renooij, C. L. M. Witteman, B. M. P. Aleman, and B. G. Taal (1999). How to elicit many probabilities. San Francisco, CA, pp. 647–654.
- Varis, O. and S. Kuikka (1999). Learning Bayesian decision analysis by doing: lessons from environmental and natural resources management. *Ecological Modelling* 119, 177–195(19).



Figure 12: Model management page

Which threats do you want to include? Short names may only be 30 characters or less - all of which are letters, digits or underscores, and must start with a letter.

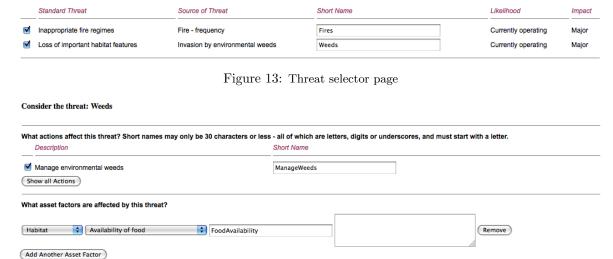


Figure 14: Action and Asset Factor selector page

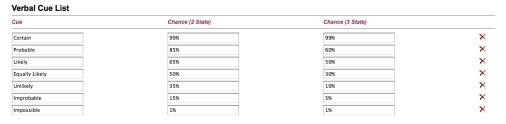


Figure 15: Verbal cues page

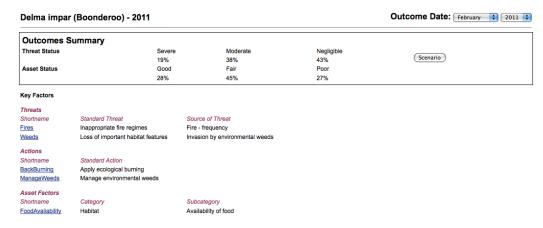


Figure 16: Example outcome page