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

Jeremy Cain

Centre for Ecology and Hydrology
Crowmarsh Gifford, Wallingford, Oxon, OX10 8BB, UK
Tel. : +44 (0)1491 838800
Fax. : +44 (0)1491 692424
Email: jby@ceh.ac.uk

ISBN 0 903741 00 9

Published by the Centre for Ecology & Hydrology

Crowmarsh Gifford, Wallingford, Oxon, OX10 8BB, UK

August 2001

Cover photo reproduced with the kind permission of Charles Batchelor

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Contents

Introduction to the guidelines iv-v
Glossary and acronyms vi
Acknowledgements viii

Chapter 1: Formulating integrated water resource management plans 1
Integrated water resources management 1
Formulating IWRM strategies 2

Chapter 2: How Bayesian networks can help make decisions about integrated water resource management 5
Using Decision Support Systems 5
Pros and cons of Bayesian networks 6
Are Bayesian networks for you? 7
Do you need a DSS? 7
Should you build it yourself or obtain specialist help? 8
Do you need to answer "What-if?" type questions? 8
Selecting the most appropriate DSS approach for you 8
Do you have the time and the money to use a Bayesian network? 13
Using Bayesian networks for IWRM – the underlying approach 14

Chapter 3: Bayesian networks — a technical description and key skills for their use 17
What is a Bayesian network? A description and some terminology 17
Key skills: How to turn qualitative descriptions into a quantitative model 19
Some general aims 19
A general network structure 20
Key skill 1: Choosing variables to represent ideas 21
Key skill 2: Choosing states to represent ideas 23
Key skill 3: Simplifying the network structure 25
Key skill 4: Dealing with time 26

Chapter 4: A step-by-step guide to using Bayesian networks 31
Examples for Step 4 35
Examples for Step 5 41
Examples for Step 6 45
Examples for Step 9 50
Examples for Step 10 53
Examples for Step 11 58

Chapter 5: Managing resources in the Poya Ganga — a hypothetical case study 67

Appendix 1: Filling in conditional probability tables using information types 1 and 3 95

Appendix 2: Filling in conditional probability tables using information types 2 and 4 105

Appendix 3: Further reading 121

Appendix 4: What's on the CD-ROM? 123

Who should/shouldn't use these guidelines?

These guidelines are for anyone who is faced with making difficult decisions concerning the planning and management of natural resource and other development programmes. They provide a guide to the use of Bayesian networks- an approach to decision support which is holistic, easy to use and encourages stakeholder involvement. People who have used them have found that they can help make the formulation of management strategies more rigorous and comprehensive.

While the guidelines have been developed within the water sector (and specifically to promote integrated water resource management) it should be clear that they can help in any planning process that requires a holistic view to be taken and the involvement of many disciplines. As such, it is hoped that they will be useful to those who are planning and implementing integrated development programmes, whether they are primarily water focused or not.

The guidelines should be used by those who have responsibility for management decisions or those who have been asked to make recommendations about particular decisions. They may be used by individuals but, as they have been designed to facilitate integrated water resource management, it is more likely that they will be used by small multi-disciplinary teams. The guidelines will help the team to build a Bayesian network and then use it to develop integrated management plans. Although they will do this by drawing on information from a wide range of stakeholders, in most cases Bayesian networks will not be used by stakeholders themselves.

How should these guidelines be used?

The guidelines are intended to provide a foundation from which to begin using Bayesian networks for water management problems. They are not meant to be rigid but should be adapted to your needs once you are familiar with the tool. One of the advantages of Bayesian networks is their flexibility.

Throughout it is assumed that you have access to and are familiar with the use of software for building and analysing BNs. If this is not the case, then you should obtain software from one of the suppliers listed below and follow the instructions supplied.

Two packages have been used during the development of these guidelines:

- Netica (*www.norsys.com*)
- Hugin (*www.hugin.dk*)

Fully-functional versions are available for free download, although network size is limited in both demonstration versions. Both have easy to use graphical user interfaces and detailed help files and are, as a consequence, excellent packages to begin with. They also support a full range of functionality for more advanced users.

A full list of available software, together with details of their functionality, can be found at:

http://www.cs.berkeley.edu/~murphyk/Bayes/bnsoft.html

A guide to the guidelines

The guidelines are arranged into five chapters. The first chapter provides a brief description of the concepts of integrated water resource management (IWRM) and presents some general ideas about how IWRM strategies can be formulated. The second deals with the specific ways in which decision support systems in general and Bayesian networks (BNs) in particular can and should help in the formulation of IWRM plans. Chapter 3 gives a short technical description of Bayesian networks and then presents the key skills that you will need to acquire to use them effectively. Clearly it is important that you read this chapter so you can acquire these skills, but also because the rest of the guidelines will make less sense without the information it contains!

Chapter 4 contains the core of the guidelines, divided into 12 consecutive steps. The first step is crucial in that it encourages you to identify the problem clearly. If you don't do this then you are likely to end up devising a solution to a problem that doesn't exist! Steps 2 to 8 are largely concerned with stakeholder consultation. Such consultations are essential to ensure that your final management plans can be implemented. As noted above, it will rarely be appropriate for you to use BNs directly with stakeholders and it is not strictly necessary to use them at all at this stage. However, it has been found that their use can help ensure you fully understand what stakeholders are telling you and highlight areas for further consultation. Ultimately, as with Step 1, the aim is to ensure that you have properly identified the problem.

As each situation is unique, Steps 2 to 8 provide only general guidance. It is recommended that you consult with someone who has local experience of facilitating stakeholder participation.

The remaining steps guide you in building and using a Bayesian network which represents your best understanding of the nature of the problems to be solved and contains the best information available to solve it. Many of the activities here are repeated from earlier steps associated with stakeholder consultation. Examples are provided throughout.

Finally, Chapter 5 presents a hypothetical case study, describing the development of a single BN through each of the 12 steps.

Glossary and acronyms

Adaptive management: A management approach that is flexible enough to change as a result of new information about the effects of initial interventions (*q.v.*). It is an explicit, structured and systematic process for learning from one's experience through a cycle of planning, acting, monitoring and evaluation.

Additional impacts: Factors that are changed as a result of interventions (*q.v.*) that do not affect anything else in the environmental system (*q.v.*) and are not directly related to management objectives (*q.v.*).

Bayesian network: A graphical tool for building decision support systems (*q.v.*) to help make decisions under uncertain conditions.

Bayesian network diagram: The graphical component of a Bayesian network (*q.v.*), consisting of nodes (*q.v.*) and links. It does *not* include the conditional probability tables (*q.v.*). It is more formally called a directed acyclic graph or DAG.

BN: Bayesian network.

Case data: A set of observations as to which state (*q.v.*) a variable was in given the states of its parent variables (*q.v.*).

Child variable: A variable that has links feeding into it from other variables.

Conditional probability table: A table underlying each child variable (*q.v.*) in a Bayesian network (*q.v.*) which expresses the probability that the child variable will be in a particular state (*q.v.*) given the states of its parent's variables (*q.v.*). Distributions across the states of the child variable are given for each possible combination of the states of the parent variables.

Continuous variable: A variable that can take a value between any other two values. See also "Discrete variable."

Controlling factors: Factors that cannot be changed by intervening at the scale you are considering but control the environmental system at that scale, in some way.

CPT: Conditional probability table.

Decision Support System: A collection of tools which, together, facilitate a decision making process. These tools can be conceptual and methodological as well as computer-based.

Discrete variable: A variable with a well defined, finite set of states (*q.v.*). See also "Continuous variable."

Divorcing: A technique for simplifying the structure of a Bayesian network. It involves grouping a number of parent variables so that they feed into an extra, intermediate child.

DSS: Decision support system.

Elicited probability table: The same as a conditional probability table (*q.v.*) except that they only contain some of the possible combinations of parent variable states (*q.v.*).

Environmental system: A special class of management system (*q.v.*) relating specifically to environmental management. It typically includes physical, economic, social and institutional factors.

EPT: Elicited probability table.

Implementation factors: Factors that directly affect whether an intervention (*q.v.*) can be successfully implemented.

Integrated water resources management: An approach to water management, which recognises that multiple objectives must be met through a variety of means. Achieving sustainability is fundamental to the approach.

Intermediate factors: Factors in an environmental system (*q.v.*) that link objectives (*q.v.*) and interventions (*q.v.*).

Interventions: The things you wish to implement in order to achieve your objectives (*q.v.*). They can also be thought of as management options.

IWRM: Integrated water resources management

Learning: The process by which a Bayesian network (*q.v.*) updates its conditional probability tables (*q.v.*) as a result of receiving case data (*q.v.*) about variables in the network.

Management objectives: See "Objectives"

Management system: A set of interlinked components, which are related by their common association with an entity that is to be changed through human intervention. The system is described by a set of variables, the state these variables are in and rules governing how these variables change.

Model: An aid to conceptualising and investigating the interactions between the linked components of a management system (*q.v.*). It need not be implemented on a computer, although it often will be.

Modifying parent: A parent variable (*q.v.*) whose effect on (some of) its children is dependent on the states of other parents of those children. See also "Non-modifying parent."

MP: Modifying parent.

NMP: Non-modifying parent.

Node: The element of a Bayesian network that represents a variable in the system being modelled. It is usually shown as a box, together with the name of the variable. In these guidelines the term is used interchangeably with "Variable."

Non-modifying parent: A parent variable whose effect on its children is independent of the states of any other parents of those children. See also "Modifying parent."

Objectives: Things you wish to affect through management of the water resource. These may be things you wish to improve or things you wish to prevent from worsening.

Parent variable: A variable that has links going out of it to other variables.

Stakeholder: An individual or group who can affect or are affected by a decision making process.

State: A value that may be taken by a Bayesian network variable which may be quantitative or qualitative. Variables must have at least two states. The states of a variable must represent all values that that variable can take and must not overlap with each other.

Variable: See "Node".

Acknowledgements

This document is an output from a project funded by the UK Department for International Development (DFID) for the benefit of developing countries. The views expressed are not necessarily those of the DFID.

I would like to thank the many people who have been involved in the research which led to the development of these guidelines. In particular, Charles Batchelor, John Bromley, David Jones, Chris Lovell and Patrick Moriarty all made significant contributions to the work and provided much constructive advice. In addition, Bruce Campbell, Peter Frost and Tim Lynam from the Institute of Environmental Studies at the University of Zimbabwe deserve thanks for discussing their valuable ideas on integrated modelling. Special thanks are also due to John Griffin for his hard work in preparing the final manuscript, to Brent Boerlage at Norsys for proof-reading and to John Gash for his support and encouragement throughout.

None of this would have been possible without the generous collaboration of our partner organisations. In Sri Lanka, the involvement of many staff at the International Water Management Institute was crucial to the project's success. In particular, the significant efforts of Ian Makin, B.R. Ariyaratna, G. Jayasinghe, K. Jinapala, L.R. Perera, R. Sakthivadivel and P.G. Somaratna ensured that our work was both successful and enjoyable. In addition, the support of Mr Wickramage and Mr Senevirathne at the Water Resources Secretariat, Mr Liyanage from the National Water Supply and Drainage Board and Mr P.G. Amarakoon from Monaragala District was vital to the success of the Menik Ganga study.

In Zimbabwe, the work was ably guided by Mike Clowes, Neil Lecler and Chemist Nyati at the Zimbabwe Sugar Association Experiment Station. I would like to thank them for their invaluable support and encouragement. I would also like to thank the Mkwasine Sugar Estate and the Chipiwa Farmers' Association for their generous cooperation. The work in India was made possible through the support of Mr Tucker at the Andhra Pradesh Rural Livelihoods Project, Dr Tirupataiah and Mr Narasimha Rao at the Andhra Pradesh Academy of Rural Development and the efforts of John Butterworth at the Natural Resources Institute in the UK. I would like to thank them all.

Finally I would like to thank my wife Emma for providing the illustrations in Chapter 1, putting up with my prolonged absences from home and making me a very happy man.

Chapter 1: Formulating integrated water resource management plans

Integrated water resources management

Taking an integrated approach to water resources management means recognising that a successful strategy must meet multiple objectives through a variety of means. What these objectives are and the most appropriate means of meeting them is highly dependent on the scale of the environmental system[1] to be managed (see boxes) but an effective strategy will invariably address a mixture of environmental, economic, social and political issues. Clearly, identifying all the important issues means taking an overview of the whole system to select those features that might help or hinder the attainment of management goals.

This approach underpins the whole concept of integrated water resources management (IWRM) in that it seeks to link the management of water use sectors that had previously been handled separately. It is believed that by doing this, plans can be devised to maximise the benefits obtainable from water resources without causing significant adverse effects. These benefits will only be obtained, however, if the management plan can be successfully implemented and this relies on the support of all those affected by management changes. Consequently, ideas of stakeholder participation in planning processes have become fundamental to the IWRM approach. The aim of IWRM can now be reformulated as a need to meet multiple *stakeholder* objectives through a variety of means.

---

[1] The word "system" is used extensively throughout these guidelines in three ways:

1. In the phrase "Decision Support System". This is defined as a collection of tools which, together, facilitate the decision making process. These tools can be conceptual and methodological as well as computer-based.
2. In the phrase "management system". This is defined as a set of inter-linked components that are related by their common association with an entity that is to be changed through human intervention. A system is described by a set of variables, the state that these variables are in and rules governing how these variables change.
3. In the phrase "environmental system". This is used as a special class of management system relating specifically to environmental management. It typically includes physical, economic, social and institutional factors.

Where "system" is used in isolation, the sense should be clear from the context but it will generally refer to an environmental system.

Box 1: Water resource management in a river basin

The Deduru Oya flows 140 km from central Sri Lanka to the west coast through a basin of area 2623 km^2. The basin is subject to both the south-west monsoon prevailing from April to September and the north-east monsoon prevailing from October to March. In an average year, the basin captures 4397×10^6 m^3 of water, of which 44% is lost through natural processes, 24% is removed for human consumption and 32% flows into the sea.

The population of the basin is just under 1 million people, with 10% of them living in one of the two main towns. 80% are employed in farming or fishing. 63% of families have a monthly income less than US$13.60 and only 2% of the population have access to piped water. The rest rely on shallow wells for their domestic water, often at distances of 2 to 3 km from their homes.

Land use is fairly evenly divided between plantation crops (mostly coconut), paddy rice and other irrigated field crops. 24% of the land is uncultivated and only 2% is forested. Other agricultural activities include poultry and livestock farming, and shrimp farming at the river mouth. There are also several small industries dealing with food processing, metal quarrying, saw milling, rice milling and cement production. There is widespread mining of sand from the riverbed to provide building materials. This is leading to bank collapse and increased seawater intrusion resulting in the salinisation of groundwater.

Six government institutions are responsible for different aspects of water resource use in the basin. Coordination of water related activities is carried out by the Office of the Divisional Secretary in each of the 27 divisions into which the basin is split.

Objectives of water management in this basin include: increasing agricultural productivity, improving access to clean drinking water, expanding industry and the control of saline intrusion. Strategies for increasing agricultural productivity may focus on increasing water supply through dam construction. With limited finance, however, similar results may be achieved by improving the quality of seed and fertiliser provided to farmers, and the usefulness of either solution depends on the availability of markets. Access to drinking water may be improved by drilling more wells but this may have little impact if saline intrusion is allowed to proceed unchecked. Expansion of industry will provide jobs but will lead to an increase in demand for water and a reduction in the supply to other water users. Any strategy developed must be able to be implemented and administered by the existing institutional structure.

Formulating IWRM strategies

The formulation of IWRM strategies involves deciding which management strategy to pursue from among a number of possible options. These decisions will not only be made at the beginning of a project but throughout its lifetime in response to monitoring and evaluation of project progress and changes in the environmental system. Decisions should be made by the group of people responsible for implementing them.

Some decision theorists use three phases to describe the way people make decisions. The first phase, called the “Intelligence Phase”, is concerned with identifying exactly what the problem is. The second phase, called the “Design Phase”, sets the criteria by which a decision will be made, identifies the options available and attempts to predict the outcomes of each option. The final phase, the “Choice Phase”, comprises the selection of the best option from those available. For most decisions, this procedure will be carried out informally, although a more formal approach may yield better results, particularly when the decisions to be made are complicated.

Stakeholders can contribute to each of these phases in different ways. In the Intelligence Phase, consultation with stakeholders is crucial as water management problems always relate in some way to water use by people. Therefore, as the problem relates to them, an exact problem definition can only be produced by finding out how stakeholders perceive the problem. During the Design Phase, stakeholders can help identify the options available and will be able to provide information to help predict likely outcomes (as these outcomes will often depend on stakeholder responses to

Box 2: Water resource management on an irrigation scheme

There are 199 farmers in the Chipiwa irrigation scheme in Zimbabwe, each farming 10 ha of sugar cane. The land is owned by the commercial Mkwasine estate although negotiations to transfer the title deeds to the farmers are progressing. The estate also controls the water supply to the scheme through three distribution canals and, in addition, provides agricultural extension advice to the farmers.

The scheme is arranged into several blocks, typically managed by 6 to 8 farmers. Currently, all farmers irrigate using overhead sprinklers, although some of them are interested in moving to flood and even drip. For each block, a single pump supplies a shared mainline running central to the block. This feeds a number of sprinkler lines (typically 4 to 6), which are rotated by the farmers to achieve full coverage. In any one year, yields can vary widely depending on a number of factors including farmer skill, position relative to the pump and land quality. In general, however, a farmer will obtain between 50 and 120 t ha^{-1}.

The farmers want to increase this yield and see water as the main factor constraining this – currently many farmers do not receive enough water and find it is unavailable when they most need it. They believe that the problems are a result of water distribution within the block rather than the supply in the distribution canal. Strategies to address this may involve better maintenance of pumps and pipes, improved cooperation amongst farmers and controls on water application by farmers, achieved by social or technical means. Changing irrigation techniques may also produce benefits.

different management options). Following the Choice Phase, it is important that the decisions made are endorsed by as many stakeholders as possible, otherwise the proposed plan will not receive the stakeholder support required to implement it.

As indicated above, there are certain situations when it may be beneficial to follow a formal approach to the procedure. Poor decisions can be taken because decision makers have a psychological bias towards certain information. For example, a proposal presented as having a 20% chance of failure may be rejected, while one presented as having an 80% chance of success is accepted. Equally, decision makers may pay more attention to the most recent information, events which are dramatic (though rare) and facts that support pet hypotheses rather than those which do not. Moreover, when implementing particular actions which may have more than one consequence, it can be difficult to keep track of what those consequences are.

Clearly, the more complicated the management system and the greater the volume of information to be considered, the more likely it is that poor decisions can be made. In these situations, some form of decision support may be appropriate. Since the environment is characterised by complexity, and the quantity of information needed to describe all the relevant factors is necessarily large, decision support has much to offer those involved in water management.

Chapter 2:

How Bayesian networks can help make decisions about integrated water resource management

Using Decision Support Systems

Decision support systems (DSS) can help structure decision processes and support analysis of the consequences of possible decision choices by making data easily accessible and allowing "what-if" analyses. Specific benefits quoted in DSS studies include:

- An increase in the quantity and quality of information identified as relevant to the decision
- An increase in the number of alternatives examined
- A better understanding of the management system
- New insights and improved learning
- Better decisions
- Better use of data resources
- Improved communication
- Improved documentation of the issues and justification of decisions made

However, there are a number of potential drawbacks:

- Decision makers can become over reliant on DSS to the extent that they allow it to make decisions for them
- Decision makers assume that all relevant factors are included in the DSS, when other important issues have been ignored or have arisen since the DSS was first constructed
- A poorly designed DSS can transfer power from the users to the DSS designers
- A DSS can exclude people without the technical ability to use it
- Decision making is emphasised at the expense of other management processes (e.g. people management)
- Constructing a good DSS requires significant investments of time and money

While good design and sensible use can avoid many of the problems listed above, it will always be necessary to consider whether the time and effort required to produce a DSS can be justified by the benefits expected of it. The potential benefits are listed above. However, the degree to which they are obtained depends, to a certain extent, on the complexity of the environmental system and the way in which decisions might otherwise be made about it. As each situation is unique, it is difficult to estimate the expected benefit of using a DSS before actually doing so. To help potential users to judge this, the following sections suggest a number of questions to consider.

Pros and cons of Bayesian networks

A Bayesian network (BN) is a tool that can be used to build a DSS. Like all tools, BNs can be used in a number of different ways and, clearly, the way in which they are used will have a strong effect on the quality of the decision produced. It is helpful to distinguish between (at least) two ways of using them:

1. BNs can be developed simply to provide a mathematically optimal decision on the basis of the information provided to the BN
2. BNs can be used in a way that promotes an improved understanding of the environmental system, leaving the decision makers to reach their own conclusions on the basis of that understanding.

The second approach is recommended as it supports decision makers rather than making the decision for them and allows account to be taken of other factors not included in the BN (e.g. unstated political considerations). The nature of BNs encourages users to take this second option. Specifically:

- The basis of a BN is a diagram conceptualising the environmental system to be managed. To construct this diagram, it is necessary to think carefully about how the system works as an integrated whole.

This is not easy, but improves understanding of how management options may affect the system. As a result, it is more likely that users will be making a decision based not only on the outputs of a DSS but on a full understanding of how those outputs have been produced. Moreover, it allows the user to adapt the decision recommended by the DSS in the light of factors not included in it. A further advantage of using a DSS in this way is that it helps avoid the first two drawbacks listed in the previous section.

Bayesian Networks have other features which allow them to be used in ways which can help avoid the drawbacks identified in the previous section:

- Building and using a BN does not require specialist skills. This means that decision makers can learn to use the tool to develop their own DSS. Power transfer from the users to the DSS designers is avoided, as the users are the designers!
- As BNs are diagrammatically based, it is relatively easy to understand the outputs provided by a DSS built with them. This facilitates the communication of information to people without technical abilities so they can participate more fully in the decision making process.

◆ It is relatively simple to adapt BNs to new situations. This means that, when new factors emerge which are relevant to the decision, they can be included in the DSS.

Bayesian networks were originally developed to allow the impact of uncertainty about management systems to be accounted for in the decision making process. This means that decision makers can balance the desirability of an outcome against the chance that the management option selected may fail to achieve it. This facility is particularly important for environmental management where the complexity of the natural world means that it is rarely possible to predict the exact impact of any management intervention. In an uncertain world, Bayesian networks allow users to estimate the chance that a management intervention will have a particular effect and then investigate the consequences of their uncertainty.

Unfortunately, a BN does not avoid the need for time and money to develop a DSS. In fact, using a DSS to promote understanding generally takes more time than using one simply to provide answers. Moreover, the production of uncertainty estimates requires more data to be collected. Although it is to be expected that improved understanding and the explicit recognition of unavoidable uncertainty will lead to better decisions, the time and money may not be available to do this. When this is the case, using a BN may not be the most appropriate way to proceed.

Are Bayesian networks for you?

Do you need a DSS?

As noted above, a Bayesian network is a tool that can be used to build a DSS. Therefore, the first question when deciding whether to use them is whether you really need a DSS. Consider the following questions:

◆ Are there many possible management options to choose from?

◆ Will your decision about which management option is best be based on more than one criterion?

◆ Is it difficult to predict how each of these criteria are affected by the management options you are considering?

If the answer to any one of these questions is YES, then it is likely that a DSS will be able to help you. Proceed to the next section.

Should you build it yourself or obtain specialist help?

Until fairly recently, most DSS were developed by specialist "knowledge engineers". More recently, many software packages have been produced which allow non-specialists to construct DSS for themselves. These are called DSS generators, and Bayesian networks can be considered to be among them. Therefore, the next question is whether you wish to employ specialists or construct the DSS yourself. Consider the following questions:

- Do you need the DSS to help with a well-defined problem, which you have to deal with regularly?

(as the environment is constantly changing, it is more likely that the problems associated with managing it are ill-defined and change regularly)

- Can you afford specialist help?
- Are you confident that a DSS built by someone else will fully meet your requirements?

(or consider whether a self-built DSS may be more likely to meet your requirements?)

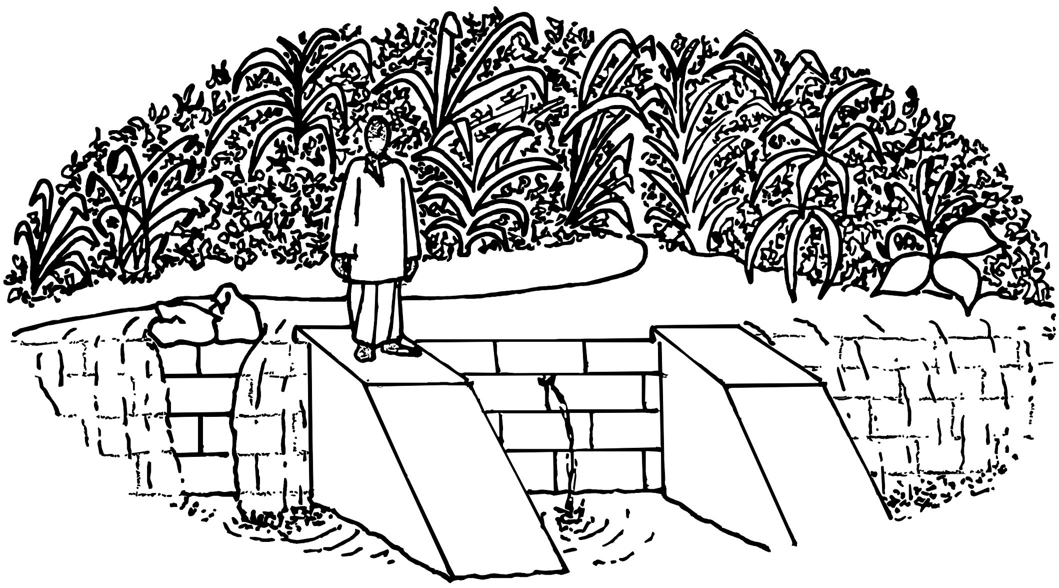
If the answer to any one of these questions is NO, then it is likely that a DSS generator will be your best option. Proceed to the next section.

Do you need to answer "What-if?" type questions?

Consider the following questions:

- Will you be able to reach a decision simply by looking at and analysing existing information?
- Do you need to predict the outcome of possible actions in order to reach a decision?

If your answer to the first question is YES and the answer to the second is NO, then you do not need a DSS with modelling capability. It will probably



be sufficient for you to collate and organise the data available to you within a database (such as Microsoft Access) and analyse it using a spreadsheet (such as Microsoft Excel). On the other hand, if your answer to the first question is NO and the second one YES, then you do need to build a DSS with modelling capability. Proceed to the next section.

Selecting the most appropriate DSS approach for you

There are many tools that can be considered to be DSS generators and will enable you construct a DSS to answer "what-if?" type questions. Obviously, you should select the one that best meets your needs. The questions below will help you identify what your needs are and you can then match these with the descriptions of different DSS generators which follow.

1. At what level do you need to represent the internal workings of the environmental system you wish to manage?

This depends on the way in which your decision will affect the factors in the environmental system which you have decided are most important (i.e. the criteria on which you will base the decision). For example, your decision may be concerned with whether to use flood or overhead irrigation for a particular scheme. The criteria on which you have decided to base your decision might be crop yield, capital costs and running costs.

If the links between the decision and the criteria are simple and obvious, then it is likely that you will not need to represent the internal workings of the environmental system in great detail. If, on the other hand, the links between the decision and the criteria are complex, then you will probably need to represent the environmental system in a more detailed way. For example, although the link between irrigation type and capital costs may be clear (a fixed price can be obtained from the supplier), the link between irrigation type and yield is less so, as yield will depend on numerous other factors (soil fertility, pest control, etc.).

The degree of detail that you will need will depend on the level of complexity that exists. However, you should also consider how much detail you, the decision maker, want to represent. It is important that you are able to understand where the outputs from the DSS come from. If you cannot do this, then you are relying on the ability of the DSS to make a decision instead of your own. So it is important that the DSS does not represent more detail than you can understand easily. There is a possible exception to this rule: if you are developing the DSS as part of a team, it may not be necessary for you as an individual to understand all parts of the DSS, as long as there is someone else on the team who does!

Because of the complexity of the environment, it is usually necessary to represent the internal workings of the system in some detail. However, it will be impossible to make an informed decision if you have too little detail to represent the most important factors or if you have too much to understand.

2. How important is it that you can communicate the reasons underlying your decision?

It may be necessary to justify your decision either to your superiors or to people who will be affected by it. If so, you should consider the best way to communicate your reasoning to them. While a written report may be suitable in some cases, a picture-based explanation may be more helpful.

3. How can people who are not directly involved in the construction of the DSS feed information into it?

As discussed above, stakeholder participation in decision-making is an underlying concept of IWRM. Therefore, it is important to account for their opinions on how a decision will affect the environment. You should consider how these opinions can be elicited from stakeholders and how they should be accounted for within the DSS.

4. Is there any uncertainty associated with your decision? If so, how important is it to explicitly represent that uncertainty in the DSS.

Again, due to environmental complexity, it is highly likely that your decision will have a great deal of uncertainty associated with it. With uncertainty comes risk. It is always important to be aware of the risk that your decision may not work — especially if the consequences of failure are serious.

Bayesian networks

Tests have shown that Bayesian networks are usually able to represent the most important factors in the system effectively. Since the networks are diagrammatically based, it is relatively easy for users to understand how those factors interact and, as a result, how the DSS produces its outputs. For the same reason, it is also fairly easy to communicate the information on which you have based your decision. As will be explained later, the diagram on which a BN is based represents simple concepts of cause and effect. Again, tests have shown that most people are able to express their ideas using such concepts. This means that information elicited from stakeholders can be used directly within the BN. Finally, BNs explicitly represent uncertainty in a way that can be clearly understood.

Although they can be used to do so, BNs are not ideally suited to situations where it is necessary to represent complexity in great detail or where concepts of cause and effect are not enough to capture ideas of how the system functions. Moreover, the representation of uncertainty requires information on what that uncertainty might be. This need increases the amount of information that has to be put into the DSS, but this is unavoidable if the associated risks are to be assessed.

Influence diagrams

Influence diagrams are very similar to Bayesian networks. The main difference is that they also support decision optimisation based on utility values assigned to the different possible outcomes of the decision. In fact an influence diagram is the more proper term for a Bayesian network when it is used according to the first approach described in the previous section. To facilitate optimisation, influence diagrams require slightly more difficult concepts to be introduced into the diagram than is the case with Bayesian networks. Consequently, it is slightly harder for people to understand an influence diagram.

As noted earlier, Bayesian networks are preferred to influence diagrams as they encourage the user to think more carefully about the decision. When the facility for optimisation is present, there is always a temptation to accept the answer provided, without considering where it came from.

Most commercial software packages that build Bayesian networks (e.g. Netica and Hugin) can also build influence diagrams. Analytica (Lumina Decision Systems Inc., ***http://www.lumina.com/software/aboutanalytica.html****) is slightly different: it uses Monte-Carlo sampling to estimate uncertainty rather than Bayesian inference. Bayesian inference is a more exact method of estimating uncertainty but this makes little difference in practice.*

Decision trees

Decision trees are another diagrammatically based approach, which capture the same information as influence diagrams but present it in a different way. An influence diagram (or Bayesian network) shows the relationships between the variables more clearly, while a decision tree shows more detail about the possible chains of events that may be initiated by a decision. However, influence diagrams are much more compact than decision trees and, consequently, easier for people to understand. Uncertainty is handled using Monte-Carlo sampling.

Commercial software packages supporting decision tree analysis include DATA (TreeAge Software Inc., ***http://www.treeage.com/****) and Decision Pro (Vanguard Software Corporation,* ***http://www.vanguardsw.com/dpbro/dpbro1.htm****).*

Mathematical modelling

On the positive side, mathematical models can easily represent any level of complexity, in terms of numbers of variables, and can handle complicated descriptions of how variables relate to each other. If you have the necessary skills then writing your own code will give you the most flexibility in creating your own DSS. Commercial software packages, such as Stella, are also available to help create mathematical models, although these are slightly more limiting than writing your own code. Stella, in particular, supports only the development of systems dynamics models. These are often useful for applications in environmental decision-making but this may not always be the case. Uncertainty analysis is possible with either approach (Stella uses Monte Carlo sampling in batch mode).

The down side of using any mathematical model is that it is hard for people not involved in its construction to understand it. Graphical interfaces can be programmed to make communication of the outputs easier although this can be time consuming. Stella overcomes this problem to a large extent by being diagrammatically based, although some of the concepts the diagram represents may be difficult for non-specialists to understand.

You can get more information about Stella from High Performance Systems Inc. at ***http://www.hps-inc.com/edu/stella/stella.htm****.*

Multi-criteria analysis (MCA)

MCA comes in all shapes and sizes. In general, however, the analysis represents only decision criteria. This means that the way in which the decision works through the system to set those criteria must be represented outside of the MCA (many commercial software packages provide facilities for doing this, although some are easier to use than others). Therefore, MCA is more suited to decisions where it is not important to understand the underlying workings of the system in detail.

More positively, MCA naturally lends itself to diagrammatic representation and its outputs can be easily understood. Moreover, the concepts it represents (that of a hierarchy of decision criteria) are fairly easy for stakeholders to understand, although it can sometimes restrict the ways in which stakeholders express themselves (stakeholders may wish to talk about things other than criteria). A further consideration is that most MCAs rely on weighted scores for input. Obtaining these scores and weights in a way that is valid for the subsequent analysis is often difficult. Uncertainty is usually addressed through a Monte-Carlo sampling of the MCA model.

An excellent guide to the use of MCA, *"Multi-Criteria Analysis: a manual"*, is available from the UK Department for Environment, Food & Rural Affairs (DEFRA Publications Sales Centre, Cambertown House, Goldthorpe Industrial Estate, Goldthorpe, Rotherham S63 9BL, UK) or electronically at ***http://www.defra.gov.uk/environment/multicriteria/index.htm***.

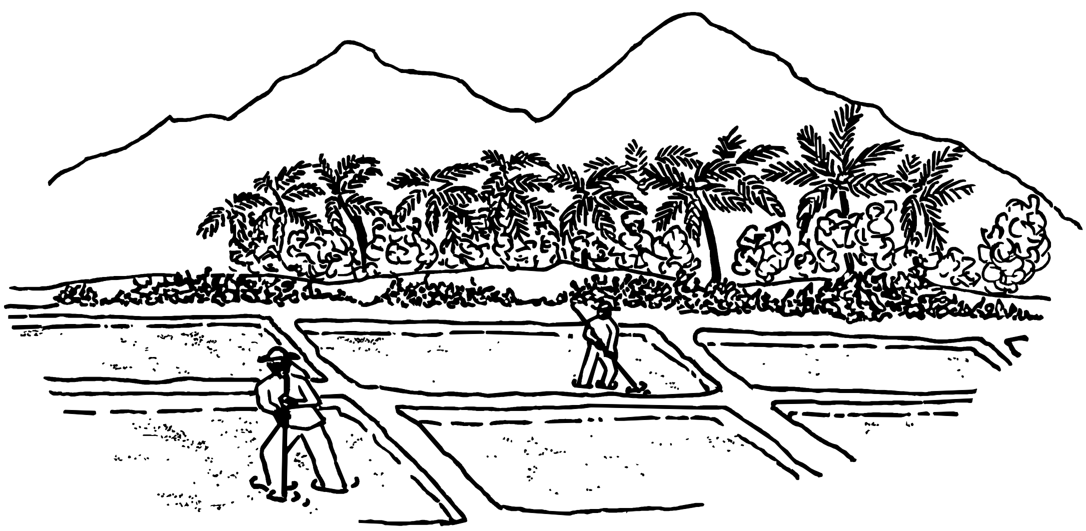
*Commercial software packages to help with MCA include Aliah Think (Aliah Inc., **www.aliah.com**) and Criterium Decision Plus (InfoHarvest Inc., **www.infoharvest.com**).*

Spreadsheets

Spreadsheets can be thought of as a simple DSS generator. They have the huge advantage that many decision makers have access to them and already know how to use them. The can be used to build mathematical models and can represent any number of variables. However, the way in which variables can be related in a spreadsheet is more limited than the other mathematical modelling approaches discussed above. Most spreadsheets also support simple optimisation.

Their main drawback is that models built within spreadsheets are not easy to understand as the dynamics they represent are hidden within the formulæ underlying each cell, rather than shown diagrammatically. Furthermore, they do not easily lend themselves to uncertainty analysis, although add-ons are available to implement Monte-Carlo sampling within them.

*Commercial software packages that provide Monte-Carlo sampling within popular spreadsheets include: @RISK (Palisade Corporation, **http://www.palisade.com/html/risk.html**) and Crystal Ball (Decisioneering Inc., **http://www.decisioneering.com/crystal_ball/index.html**).*



Do you have the time and the money to use a Bayesian network?

Constructing the diagram which is the foundation of a BN is usually a very rapid process. Even for the most complicated systems, single users who are familiar with the tool can complete this task within a day. However, building a fully functional BN that recognises stakeholder perspectives takes much longer, as it requires two major activities:

- Stakeholder consultation
- Data collection and collation

In Chapter 4 you will see that the guidelines recommend holding at least two meetings with each group of stakeholders, and then two more with all the stakeholders together. Clearly, organising these meetings takes time. In the testing that led to these guidelines, it was found that the average time taken to arrange these stakeholder meetings was about one week per group. This is not to say that you will need to spend one week with each stakeholder group, just that if you have three groups, for example, then it will take about three weeks to arrange and hold the necessary meetings. It is likely that you will want to spread these three weeks over perhaps two months. Obviously, this is highly dependent on your particular situation and the consultation approach you choose to take, but it provides a useful rule-of-thumb.

When you have constructed the BN diagram (based, in part, on what you have learned from the stakeholder consultation) it will specify exactly what data are required to allow it to function. Sometimes, some of these data will need to be collected or generated using models and this will clearly take time. Even if all the data are potentially available, they will rarely be in one place or in a format which you can use immediately. Therefore, collating and formatting data for entry into the BN can also take a significant amount of time. Our testing shows that this can potentially take months although this is highly dependent on data availability and the human resources you have to collect it. Data collection may also raise the need for modification of the BN diagram which may, in turn, lead to further stakeholder consultation. If care is taken in the earlier steps then such modifications should be avoided but extra time should be allowed for, in case the need arises.

Clearly, the activities outlined above also carry a financial cost in terms of staff time and the costs of arranging workshops. These should be calculated and budgeted for before BN construction begins, in order to ensure that you can complete the process. If you do not complete the process, not only do you risk making a poor decision, but you may also alienate your stakeholders, who will see no benefits arising from their efforts.

Many of the potential costs described above are related to the size and complexity of the BN created. Sometimes a small, inexpensive network can provide some useful insights, and then when funds become available it can be expanded into something more detailed. It should be noted that, since a network may develop over a number of years, these costs can also be spread over a number of years. Moreover, parts of a BN developed for one decision problem (including the information used to drive it) might well be useful in a later BN developed for another problem. In this way, development of a DSS can be seen as a valuable long-term investment in improving decision making.

So, do you have the time and money to use a Bayesian network? To answer this properly you need to compare the costs outlined above to the benefits you hope to gain. To a large extent this must be a subjective judgment dependent on your experience of previous decision making processes you have been involved in. If you can identify problems with these processes and you think BNs may be able to help you improve on them, then it is likely that you will find benefit in their use. During the testing which led to these guidelines, it proved difficult to measure these benefits in any rigorously objective way. However, the water management professionals and policy makers who were involved in the testing, all found that BNs helped make the formulation of management plans more rigorous and comprehensive.

Using Bayesian networks for IWRM – the underlying approach

If you have decided that it is appropriate to use a DSS based on Bayesian networks, to help formulate a IWRM strategy, then these guidelines will help you. As they describe a particular approach to using Bayesian networks, this section briefly explains why this approach has been developed.

At its simplest, the approach is composed of two linked activities:

1. Eliciting information from stakeholders
2. Construction and analysis of a Bayesian network containing stakeholder and any other information considered relevant by the decision maker

While the guidelines briefly outline ways to elicit information from stakeholders, this is a subject in its own right and it is recommended that you consult with an expert in this field. Further information can be found from:

http://www.oneworld.org/iied/resource/

http://www.ids.ac.uk/ids/particip/intro/introind.html

The main focus of the guidelines is on the construction and analysis of a DSS using Bayesian networks. The approach taken to this is based on the following four principles:

1. Users should be able to build their own DSS

By building it themselves, users can make sure that the decision support system meets their needs.

2. A DSS should be used to understand the nature of the decision better

A DSS should help users make a better decision not an easier one. It should be a "tool for thinking", not an automatic answer provider, and *NOT* make the decision for the user. Instead, it should encourage the user to identify all the relevant information and analyse it more deeply.

3. A DSS should be developed from stakeholder consultation

Without this, it is unlikely that you will be able to implement decisions based on it.

4. A DSS should encourage users to deal explicitly with uncertainty

It is impossible to be certain about the consequences of any environmental management decision. This must be recognised together with the effect of that uncertainty on the decision.

Chapter 3: Bayesian networks — a technical description and key skills for their use

What is a Bayesian network? A description and some terminology

Bayesian networks are composed of three elements:

1. A set of nodes representing management system variables, each with a finite set of mutually exclusive states (the terms “node” and “variable” are used to mean the same thing throughout these guidelines).

   Variables can either be discrete or continuous. A discrete variable is one with a well-defined finite set of possible values, e.g. the number of wells in a village; whether a crop is wheat, cotton or sorghum; whether a statement is true or false. In a BN, each of these values becomes a state of the node. A continuous variable is one that can take on a value between any other two values. Examples include rainfall depth, groundwater level, crop yield and price. When represented in a BN, the full range of values taken by a continuous variable must generally be broken down into sub-ranges, with each sub-range becoming a state of the node.

2. A set of links representing causal relationships between these nodes.

   Links, therefore, have direction – from cause to effect. If there is a link from node A to node B, B is described as a child of A, and A as a parent of B.

3. A set of probabilities, one for each node, specifying the belief that a node will be in a particular state given the states of those nodes that affect it directly (its parents).

   These are called conditional probability tables (CPTs) and can be used to express how the relationships between the nodes operate.

Elements 1 and 2 together form a BN diagram (or, more formally, a directed acyclic graph): the addition of Element 3 creates a fully-functioning BN. Figure 3.1 shows an example of a BN diagram (the equal probability distributions across the states of each node merely indicate that the BN is not yet fully functional). The structure of this BN diagram encodes the perception that river flow is affected by forest cover and rainfall and this, in turn, affects the amount of water stored by a dam (“Dam storage”) whose construction is being considered (“Construct dam?”). The other relationships represented by the BN diagram can be read from it in a similar way. Underlying each node in the BN (and not shown in Figure 3.1) are the CPTs. Table 3.1 shows the CPT describing the relationships between river flow (the child node) and forest cover and rainfall (the two parent nodes). It should be noted that a CPT contains entries for every possible combination of the states of the parents.

Once all the CPTs have been completed in a similar way, the BN can be compiled and used for analysis. In general terms, this is performed by altering the states of some nodes while observing the effect this has on others. As the BN is a network, the impact of changing any variable is transmitted right through the network in accordance with the relationships expressed by the CPTs. Changes in any node simply arise from the combined effect of changes in all the nodes linked to it either directly or indirectly. (In formal terms, the BN encodes a joint probability distribution over all the nodes. Every time the state of a node changes, the joint distribution is updated through the iterative application of Bayes' theorem). Changes in the BN are observed as changes in the chance that a node is in a particular state. Due to the uncertainty in the CPTs, it is rare for a node to definitely be in one state or another and it is far more common for probability distributions across all the states of a node to be observed.

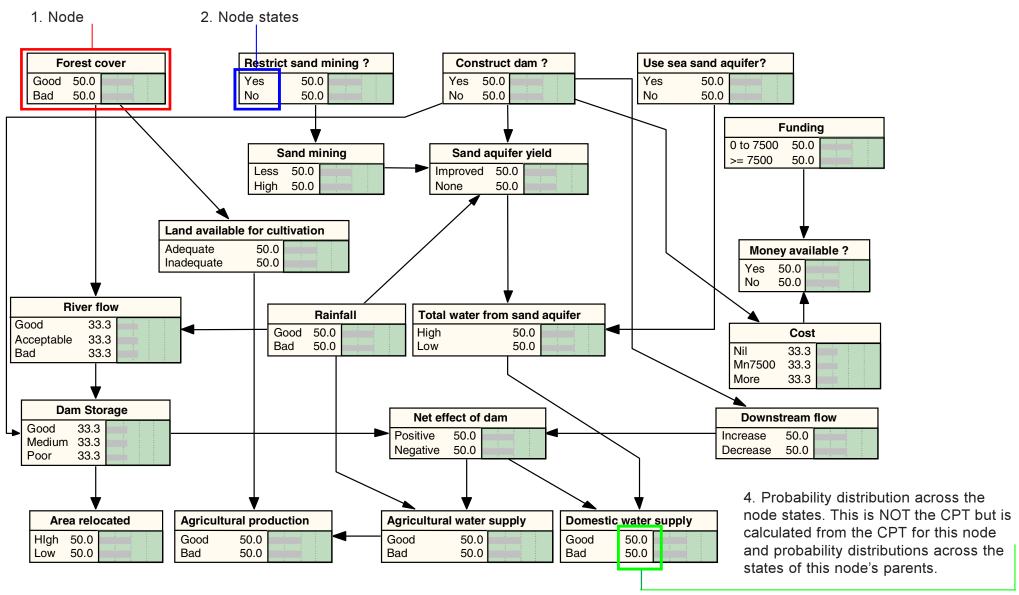


Figure 3.1: Simple Bayesian network showing the environmental system related to agricultural production in Deduru Oya, Sri Lanka

Table 3.1: Conditional probability table for the node "River flow" in Figure 3.1.

		River flow:		
Forest cover:	**Rainfall:**	Good	Acceptable	Bad
Good	Good	0.60	0.40	0.00
Good	Bad	0.00	0.10	0.90
Bad	Good	0.40	0.60	0.00
Bad	Bad	0.00	0.00	1.00

Read the table one row at a time. For example, the first row says: "If forest cover is good and rainfall is good, then there is a 60% chance that river flow will be good, a 40% chance that river flow will be acceptable, and no chance that it will be bad."

Further technical details can be found at *http://www.cs.berkeley.edu/~murphyk/Bayes/bayes.html,* together with a recommended reading list. The book "An introduction to Bayesian Networks" by Finn Jensen (UCL Press, 1996) is particularly recommended for beginners.

Key skills: How to turn qualitative descriptions into a quantitative model

Some general aims

Although building and using Bayesian networks does not need specialist skills, a degree of imagination is required to realistically represent the world (as described by stakeholders) in the form of a simple conceptual model. Experts call this "knowledge engineering" but it's not as difficult as it sounds. This section provides you with some techniques to help.

The aim in building a BN should be to ensure that all ideas are *clearly* captured by the network. The logic underlying these ideas will be represented by three things:

- ✦ The network structure (how the nodes are linked together);
- ✦ The names of the nodes;
- ✦ The names of the states of the nodes.

Ways of representing ideas with node names and state names are discussed later. However, the important point to note is that all three of the above are of equal importance in capturing the logic.

When beginning to build a BN, it is a good idea to start with the basic logic of the system and then add the details later. Remember that one of the points of using a Bayesian network is to give an overview of the whole environmental system, so begin by naming and linking nodes (and states) to represent the most important variables in the system ***as a whole***. You can add nodes later to describe how individual components of the system work.

The word "clearly" has been stressed since it is crucial that the network is also understood by those who have not been directly involved in building it. For example, while it may be clear to those involved how a consensus among farmers can affect sugar cane yields, those who are unfamiliar with the situation will need more detail to understand it fully. In Figure 3.2, the ideas that the BN is intended to represent are more easily communicated by the one on the right than by the one on the left.

While ensuring that all important ideas are represented at an appropriate level of detail, it is also useful to minimise the number of nodes and states. This is for two reasons: the smaller the network, the easier it will be for other people to understand it, and the easier it will be to fill in the CPTs. The optimal network size is achieved when ideas are represented in as concise a way as possible given the need for the network to be self-explanatory. During construction, users should ask themselves if all the variables they have included are really necessary. It may be possible to delete some entirely or combine two or more ideas into one variable. Equally it may be possible to reduce the number of states given to each variable. Ways of

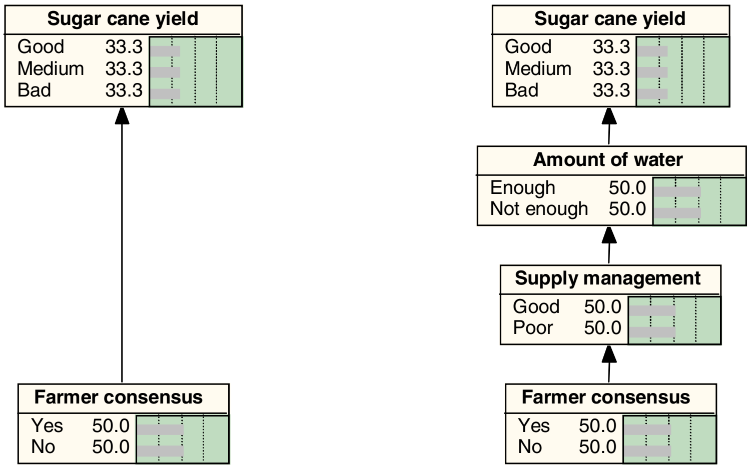


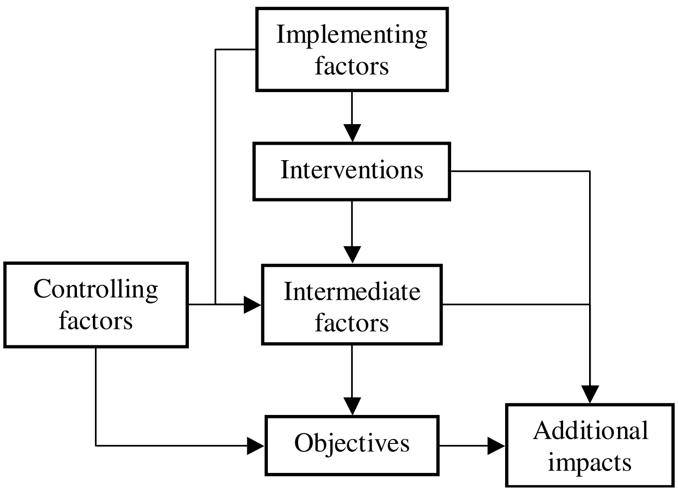
Figure 3.2: Two Bayesian networks showing relationship between farmer consensus and sugar cane yield. The BN on the right-hand side describes the relationship more clearly.

doing this will be discussed below but it should be noted that carrying out the procedures recommended, helps to produce a better final decision by forcing the user to think in depth about how the environmental system is really working.

A general network structure

You may find it easier to build a Bayesian network if you try to put the variables that you think are important into categories and then arrange them according to the structure suggested below. As you become more familiar with BNs you should adapt this structure to meet your own needs but it provides a useful starting point. Six categories of variables are suggested (shown in the table opposite).

The following structure is suggested, where the arrows show how the categories are likely to be linked. For example, by definition, implementing factors will be linked directly to those interventions which they affect. However, they may also link to intermediate factors. The arrows are intended only as a guide – other connections between nodes in different categories may be possible. If you think this is the case for your network, then do not be constrained by the diagram below.



Category	Description	Examples
Objectives	The things you wish to affect through management of the water resource. These may be things you wish to improve or things you wish to prevent from worsening. They will define the criteria on which your management choice will be based.	◆ income ◆ agricultural water resources ◆ agricultural productivity ◆ water quality ◆ access to domestic water requirements
Interventions	The things you wish to implement in order to achieve your objectives. They can also be thought of as management options. Some are implemented as "one-offs" (e.g. "construct a dam"), others over a longer period (e.g. "subsidise agricultural inputs").	◆ increase forest cover ◆ construct a dam ◆ subsidise agricultural inputs ◆ train farmers ◆ install wells ◆ regulate groundwater extraction
Intermediate factors	Factors which link objectives and interventions.	◆ river flow (linking "increase forest cover" to "agricultural water resources") ◆ yield (linking "train farmers" and "income")
Controlling factors	Factors which cannot be changed by intervening at the scale you are considering but control the environmental system at that scale, in some way.	At the scale of a river basin: ◆ population ◆ rainfall ◆ macro-economics ◆ government policy
Implementation factors	Factors which directly affect whether the intervention can be successfully implemented both immediately and in the future (depending on whether the intervention is implemented as a one-off or over a longer period).	◆ funding (for "construct a dam") ◆ land availability (for "increase forest cover") ◆ community support (for "regulation of groundwater extraction")
Additional impacts	Factors which are changed as a result of interventions that do not affect anything else in the environmental system.	Dependent on the system you are considering. For example, as well as decreasing river flow, increasing forest cover may lead to an increase in bird populations. It is unlikely that this change will affect the water resources in any way and so it may classed as an additional impact.

Key skill 1: Choosing variables to represent ideas

Variables can represent any physical, social, economic or institutional factor. They can represent tangible things like water, or intangible concepts such as a consensus among farmers. They can represent quantities of those things (water volume, for example), they can represent a property of those things (water quality) or they can represent movement of those things (water flow). They can also represent actions (irrigation, road construction, learning). Taking advantage of this flexibility is the key to capturing ideas effectively.

For example, sugar cane yield is affected by the application of nitrogen fertiliser. The degree of effect is dependent on the type of fertiliser used, the amount applied and the time at which it is applied (rainfall shortly after application can cause crop burn). Each of these factors can be represented by its own variable. In some cases, however, it may not be necessary to represent each factor as a separate node. For example, the amount of water yielded from an aquifer is dependent on the aquifer storage, transmissivity and extent. However, it will rarely be necessary to represent these properties separately, especially as it is unlikely that they can be changed through management intervention. Instead, a single node can be used to represent a combination of the three factors, as shown in Figure 3.3. This has the advantage of reducing the size of the network.

When deciding how to capture a particular idea, it is important to consider the spatial area and time period which the BN you are constructing represents. For example, if a BN represents the environmental system in a catchment over one year, then the use of this scale presents a problem in defining a variable to capture the idea of groundwater level, as this will vary greatly in space and time, and no average value is sensible. To address this, it is more helpful to think of a substitute for groundwater level, guided by the children of the groundwater node in the BN. For example, the BN on the left of Figure 3.4 is trying to express the idea that groundwater level will affect the water availability in the catchment over the whole year. A suitable substitute for groundwater level might be the number of days in a year for

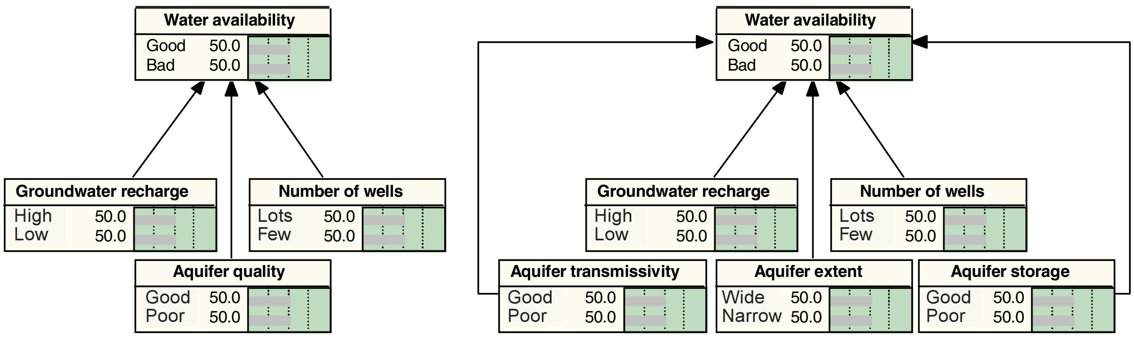


Figure 3.3: Two Bayesian networks showing factors determining water availability. In the BN on the left, the node "Aquifer quality" represents a combination of the nodes "Aquifer transmissivity", "Aquifer extent" and "Aquifer storage" in the BN on the right.

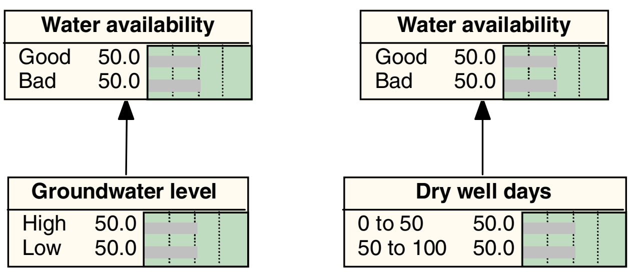


Figure 3.4: Two Bayesian networks capturing the idea that groundwater will affect water availability. It is easier to give a value that is representative of groundwater changes in space and time to the groundwater variable on the right ("Dry well days").

which wells were dry, as this idea relates directly to water availability, the child node. Therefore, while the BN on the right of Figure 3.4 expresses the same idea as the one on the left, it is easier to imagine how a value which is representative of a year, for the whole catchment, can be given to the variable representing the effects of groundwater. Occasionally it makes sense to introduce separate nodes for the values of some variables at different points in time and space. This is discussed in more detail below.

Key skill 2: Choosing states to represent ideas

As a simple guide to selecting states, for each variable, decide how to describe:

1. the state it is currently in;
2. the state towards which you think it will move under your proposed management plan;
3. any intermediate states (you may skip this if you want to – see below).

For example, “agricultural production” may currently be in a poor state and you expect your management plan might cause it to be in a good state. In changing from a poor to a good state, however, it might pass through a medium state, where agricultural production is improving but not yet enough to be called good. Therefore, you should choose the states “Good”, “Medium” and “Poor”.

Thinking about what states to give a node is an excellent check on whether the variable you have chosen properly represents the idea you are trying to capture. In the example in Figure 3.4, the variable “Dry well days” has been given states, referring to the number of days when the wells in a catchment are dry (in two sub-ranges: 0 to 50 days and 50 to 100 days). This choice of states might not be appropriate if it matters when the well dries up. By specifying the states in numerical terms, the wrong idea is captured in the network, as the implication is that it makes no difference when the well dries up. If the time the wells dried up is more important than how long they were dry for, it may be better to select states indicating that time (as in the BN on the left of Figure 3.5). If both are important, then it may be better to introduce a new node, as shown in the BN on the right of Figure 3.5.

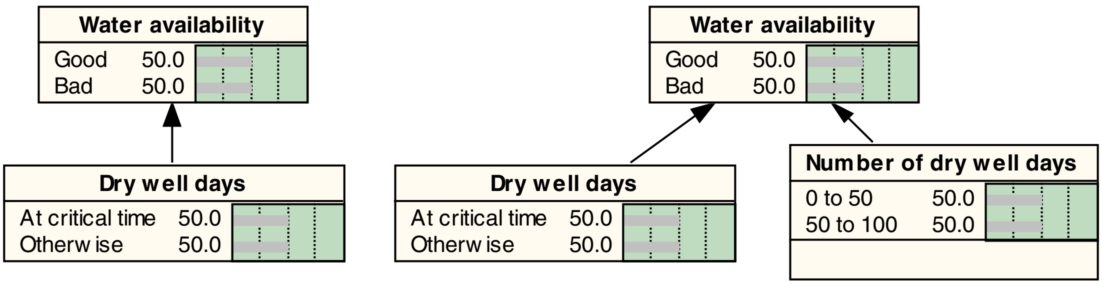


Figure 3.5: Two Bayesian networks capturing the idea that groundwater will affect water availability. The states chosen for “Dry well days” in the BN on the left suggest that the timing of the dry-up is the crucial factor. The variables and states in the BN on the right suggest that both the timing and duration of the dry-up are important.

The different state names given to the variable "Dry well days" in Figures 3.4 and 3.5 show how a variable in the environmental system may have two or more characteristics, each of equal relevance. It is important to be sure that all characteristics of a variable are identified and differentiated. However, they should only be included in the BN if they are considered to be a key factor in the functioning of the environmental system.

Another consideration when choosing states, is that the states you select fit in with the logic of the BN structure as a whole. To do this, look at the parents and children of the node. For example, in the BN on the left of Figure 3.6, the states given to "Crop water application" are logical, as all its parents affect, in some way, the quantity of water applied to the crop. In the BN on the right, however, a further parent node relating the timing of water application has been introduced. As this variable does not refer to a quantity of water, the states of "Crop water application" must be changed to reflect this new idea. In this case, the change to the state names of "Crop water application" does not require a change in the state names of "Yield", as the logic is unaffected. However, this may not always be the case.

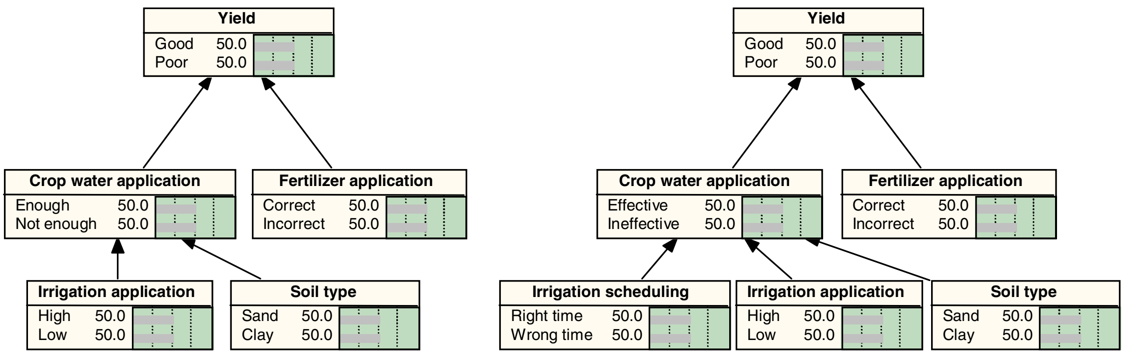


Figure 3.6: Two Bayesian networks showing factors determining yield. On the left, "Crop water application" only represents whether enough water has been applied to the crop. On the right, the same variable represents whether enough water has been applied at the right time. Consequently, the state names of the node have been changed to reflect this new idea.

As noted earlier, it is useful to minimise the size of the network by giving as few states as possible to each node. To help with this, focus on those states that are of interest to you as a decision maker. To return to the example used earlier, if you are only interested in agricultural production being good, then there is no need to include a state called "Medium". You are not interested in every state the environmental system can take. You are only interested in seeing how the system can be manipulated to reach your management objectives. However, it is important to remember that the states you choose must describe all the values that a node might take. So if you had decided to delete "Medium" as a state of "Agricultural production", then the new "Poor" state must be understood to include both the states you previously described as "Medium" and "Poor". Alternatively, continuing the example, you might rename the second state and have two states called "Good" and "Other". Focusing strongly on your needs will help keep the BN to a manageable size. In many cases, it is sufficient to give nodes only two states: one positive and one negative. Examples of this can be seen in the networks shown above.

Choosing states is fairly straightforward while you are trying to represent the basic ideas but becomes more difficult when you start to fill in the conditional probability tables (CPTs). This is because you will often need to define exact values for the states you have chosen. For example, in Figure 3.6, to fill in the CPT for the node "Yield" you will probably need to define the state "Good" as a number of tonnes per hectare (>100 t ha^{-1}, say) and do likewise for "Bad". To begin with, however, don't restrict yourself by worrying about quantifying the states you choose. It is more important to make sure the BN is logical and expresses all the necessary ideas. If you need to, it will be possible to adapt it later to help fill in the CPTs.

3 Key skill 3: Simplifying the network structure

With so many factors defining most environmental systems, BNs can become very complex. Often these complex structures can be simplified. This offers two advantages: firstly, it helps the BN be more easily understood by those not involved in its construction, and secondly, it makes it easier to fill in the CPTs.

The best way to simplify a BN is to be clear about which factors *must* be represented in the BN and which factors do not need to be. For example, if a factor is unlikely to affect the outcome of a management plan or to be changed by it, there is no need to include it in the network. Equally, if a node state is not of interest or is unlikely to be reached, you should leave it out. This sounds obvious, but it is surprisingly easy to include information in the BN that is not strictly necessary. To avoid this, review your network constantly to ensure that it does not include any unnecessary information.

Another good way to simplify a BN is to carry out "divorcing" (Figure 3.7). The network on the left includes six factors, all feeding directly into the node "Yield". However, the three nodes at the top represent factors relating to fertiliser, so they can be grouped together as parents of a single node, "Fertiliser application". They are now said to be divorced from the node

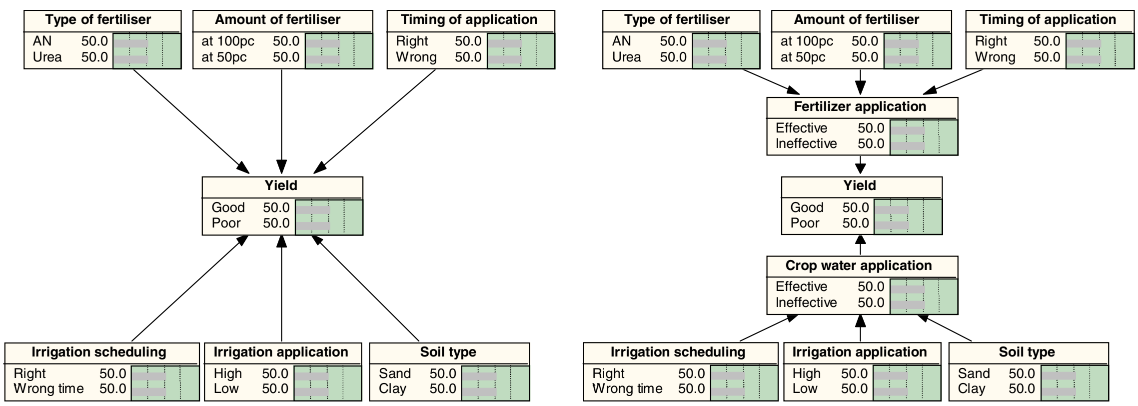


Figure 3.7: Two Bayesian networks showing the factors affecting yield. The one on the right is simplified by adding two variables, "Fertiliser application" and "Crop water application", to divorce the other six variables from the node "Yield".

"Yield" as they no longer affect it directly. The same can be done for the three nodes at the bottom, which all represent factors related to water.

Figure 3.7 shows that divorcing actually adds nodes to a network, which may not appear to be the best way of simplifying it. However, although nodes are added, the combined size of the CPTs underlying all the nodes is reduced. This is because CPT size is determined both by the number of parents of a node and the number of states each of its parents has. Moreover, divorcing makes the network easier to follow as the new variables added group the BN into logical units. This raises an important point: divorcing should not be done when it results in ignoring important interactions between parents which can influence the child. For example, in Figure 3.7, the effect that a particular type of fertiliser has on the yield may be dependent on the soil type. This effect can be captured by the BN on the left but not by the divorced BN on the right. Consequently, if there is a dependency between these two parents ("Type of fertiliser" and "Soil type"), then the divorcing procedure shown in Figure 3.7 should not be done.

Divorcing also has a further impact in that it increases the number of nodes between the interventions and the objectives. This may have the effect of "diluting" the impact of the interventions on the objectives, particularly if the CPT underlying the divorcing node (e.g. "fertiliser application" in Figure 3.7) is specified with uncertainty. By giving a probability of 100% to one child state for every combination of parent states in the CPT, the uncertainty in the relationship between nodes expressed by the CPT is effectively removed. It may be appropriate to do this for divorcing nodes, although this may not always be the case. As above, you should think carefully about your reason for carrying out the procedure and see whether logic suggests that the introduction of further uncertainty is appropriate.

4 Key skill 4: Dealing with time

A Bayesian network represents a single time period. For water management applications, this will often be a year or an agricultural season. Obviously, it is important that decisions are made with consideration given to how management choices will affect the environmental system in the future. Therefore, for Bayesian networks to be a useful tool, they need to be extended, in some way, to allow a long-term view to be taken.

It is not too difficult to do this, but it is important to consider exactly what additional information you hope to gain by doing it and design your extended BN accordingly. The BN representing a single year (for example), in the environmental system you want to manage, can be replicated for every further time step you need, with those variables from the previous time step which affect variables in the next one being linked together (Figure 3.8).

This can be done for as many time steps as you like. It should be noted that variables present in each time step may be different, but the time steps need not be of equal length. For example, in Figure 3.8, the second time step might represent how the system has changed by Year 5 (instead of Year 2, as shown). Depending on your purposes, this might be all you need. For example, to make your decision you may only want to know what the immediate effect of your intervention will be, together with an idea of

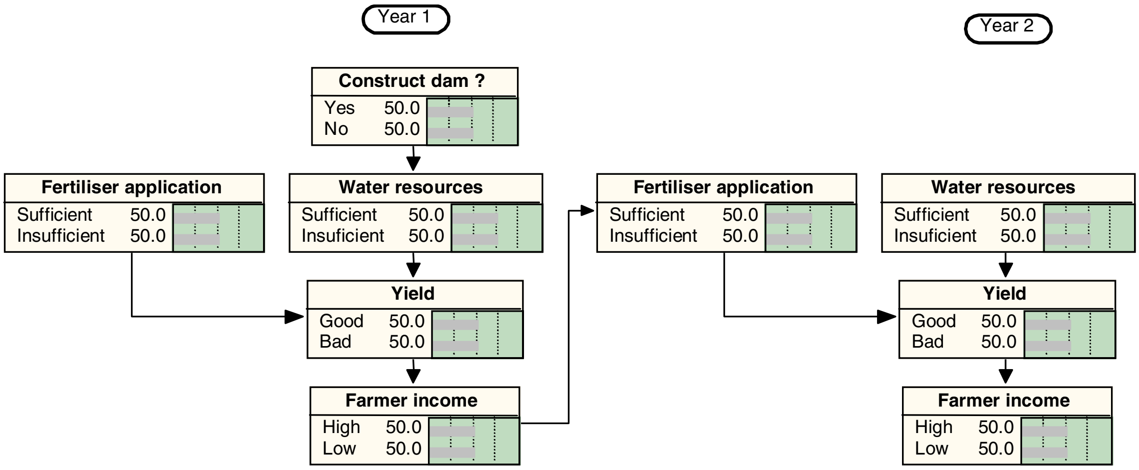


Figure 3.8: A Bayesian network representing an environmental system over two annual time steps. If farmer income increases in Year 1, then fertiliser application may increase in Year 2 as farmers are able to buy more.

whether that effect will still be felt five years after implementation. In this case, two time steps would be all you need: one representing the effects in Year 1 and the other the effects in Year 5.

When considering the planning period over which you wish to track changes in an objective, it is important to be aware that in a complex system the uncertainty associated with the value of an objective increases year on year. This can be shown mathematically, but it is also clear from common sense: looking further into the future will only increase that uncertainty. In practice, this means that it may be pointless to try and predict how an environmental system will change over more than, say, three years, as any answers will be so uncertain as to be meaningless. This uncertainty is not a consequence of using a BN, but a result of the highly complex nature of the environment, and particularly the people who live in it. People are not passive and will react to management interventions in highly unpredictable ways (even with extensive stakeholder consultation). Recognising this, many people are now advocating what has become known as “adaptive management”.

Adaptive management does not try to predict the impact of management interventions too far into the future. Instead, it tries to understand what might happen in the short term, selects an intervention on the basis of that and then monitors the impact of that intervention. If the intervention is seen to be having the effect which was expected, then all is well. However, if it is not, then the reasons for the failure can be identified and used to improve the understanding on which the first decision was made. With this improved understanding a new, and hopefully better, decision can be taken. This approach implies that, when using a BN, you need only to represent enough time steps to understand how the system, *as it is now*, is working.

This is not to say that you should only try to represent a single time step. Although this may sometimes be all you need to do, it will more often be important to understand how any “feedbacks” in the environmental system may affect your objective. Feedbacks are when an intervention produces two or more effects with opposing impacts on the objective. For example, it may be that drilling new boreholes in an area will have two effects:

1. It will increase the number of people using groundwater as they will have better access;
2. It will increase the amount of groundwater people use not only because they have better access but because the boreholes are equipped with better pumps which deliver water more quickly.

Depending on a number of other factors (the safe yield of the aquifer, for example), the amount of water extracted from the borehole may eventually become so great that the aquifer becomes depleted and the volume of water available reduces. Such effects happen across more than one time step and it is important to identify them and understand how they change in relation to each other.

A BN can help you to do this using only two steps, as shown in Figure 3.9.

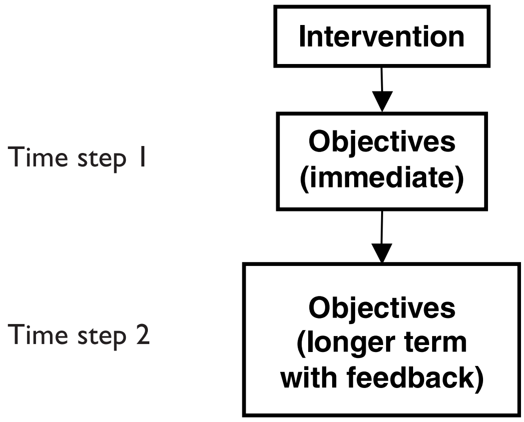


Figure 3.9: Schematic diagram of the representation of multiple time steps using a Bayesian network

Time step 1 represents the immediate effects on the objectives (i.e. the first year after the intervention). Time step 2 represents the effects on the objectives over a time period considered to be sensible, given the uncertainty, and allows any possible feedbacks to be included. For example, the total volume of groundwater being extracted after a number of years can be compared with the safe aquifer yield to see if the volume of water resources can be sustained.

A BN to do this is shown below in Figure 3.10. Here the feedback arises from the interaction between the total groundwater extracted, the safe aquifer yield and the groundwater extracted per person. If the total extraction exceeds the sustainable yield, then we would expect the amount that a person can extract to decrease as the aquifer yield is not sufficient to provide everyone with as much as they want. This interaction is represented by Time step 2 and sustainability can be judged by comparing the "Total GW extracted" variables in each Time step. If the values are significantly different then it can be concluded that feedbacks are important and the BN will also indicate the nature of that effect (i.e. positive or negative).

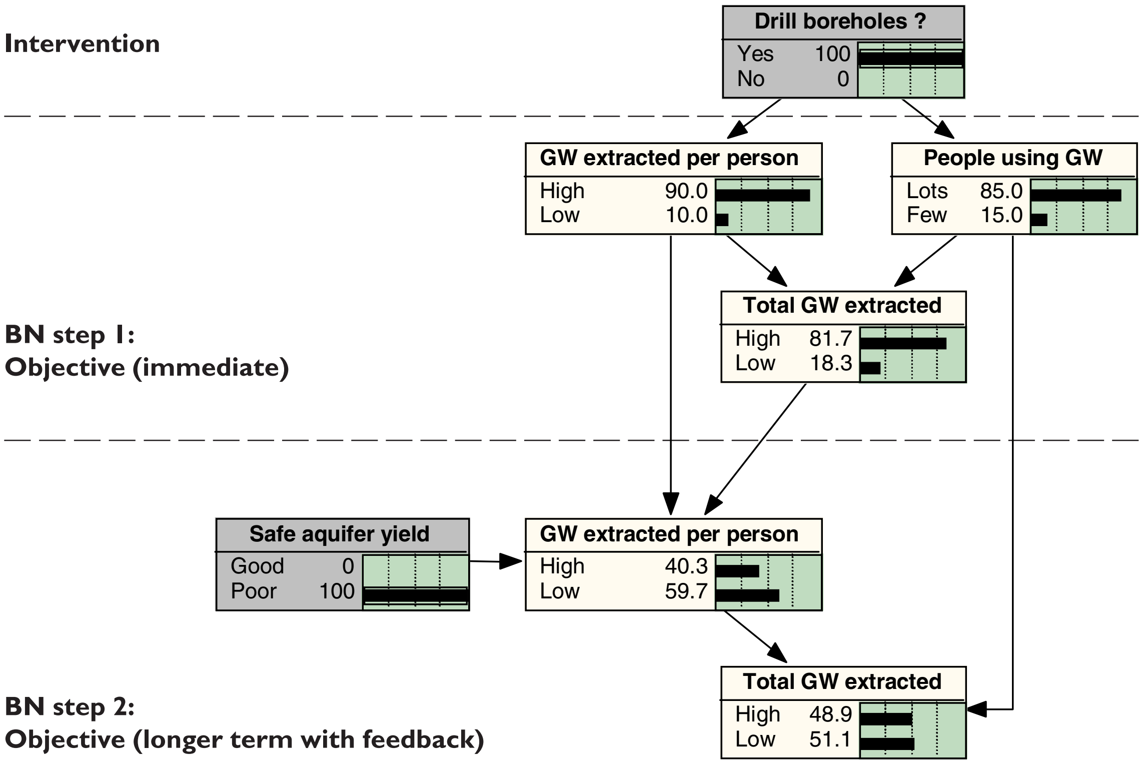


Figure 3.10: A Bayesian network representing the feedback between the volume of groundwater (GW) extracted and the number of people using groundwater over two time steps.

With this information, the intervention can be adapted to minimise negative feedbacks or encourage positive ones. For example, if the BN indicates that borehole use may exceed the safe yield of the aquifer, then the intervention can be adapted to include some way of controlling the number of people who have access to the borehole.

Chapter 4:

A step-by-step guide to using Bayesian networks

Step 1: Be clear about what you want to use the BN for
Step 2: Establish contacts with stakeholders
Step 3: Initial stakeholder group consultations
Step 4: Construct preliminary BNs *
Step 5: Further stakeholder group consultations *
Step 6: Draw conclusions from stakeholder consultation *
Step 7: Hold joint stakeholder workshop to discuss differences in viewpoints
Step 8: Complete stakeholder BNs
Step 9: Construct 'master' BN diagrams *
Step 10: Collect data and specify conditional probability tables (CPTs) *
Step 11: Use master BN to make decision *
Step 12: Hold a second joint stakeholder workshop to discuss your decisions

Step 1: Be clear about what you want to use the BN for

1. List the things you hope to improve through management of the water resource. Also list the things you do not want to worsen. These should be your management objectives and should also define the criteria on which your decision about which management option to pursue will be based.

2. List the management interventions you wish to investigate as ways of achieving your objective(s).

These guidelines will help you build a DSS to allow you to judge the impact of each management option on your decision criteria. On the basis of this, you will select the best management intervention(s). Therefore it is important that you carefully consider points 1 and 2 above. Having said this, as you proceed to build the BN, you should not restrict yourself solely to the problems and interventions you identify now. Hopefully, other ideas will come up as part of the process, particularly following stakeholder consultation. It is important that you include these ideas in your final decision.

* *Examples of steps 4-6 and 9-11 are given following the text for that step.*

3. Define the geographical area that you are considering.

4. Decide on the most appropriate period for planning. Choose the smallest period over which you think a significant change will take place, although the larger the planning period, the better. A year or an agricultural season is often the ideal choice.

5. Identify the groups of people you are hoping to affect by water management, including those who may be affected adversely.

6. Identify which groups of people will have some influence on the success or otherwise of your water management plans, without being directly affected themselves.

Step 2: Establish contacts with stakeholders

> As each situation is unique, only general guidance can be provided for those steps that involve stakeholders directly. You are recommended to consult with someone with local experience of facilitating stakeholder participation.

Your responses to points 3, 5 and 6 in Step 1 will identify the groups of stakeholders you need to involve in the development of the BN. Make sure that your responses recognise stakeholder groups who can become marginalised in decision making but who are, nevertheless, crucial to successful implementation. In particular, make sure that the poorest water users are included in the process, together with women at all levels.

Decide how many representatives of each group you wish to involve directly. In general, involving more people will be more difficult but should lead to a better result. Obviously, the number of people you involve will be dependent on the time and resources you have available but use the following as a guide:

✦ where the stakeholder group is a formal institution it may be sufficient to have a single person who will represent the official view of the institution. Care should, however, be taken that they do this and do not express their own, personal, viewpoint.

✦ where you have good reason to think that there is a high level of agreement within a stakeholder group, it may, again, be sufficient to have only one person representing it. The same caution should be applied as above.

✦ where a stakeholder group is large and / or diverse, at least three people should represent it and, preferably, more. These people should represent the range of the diversity.

In all cases, stakeholder group representatives should meet the following criteria:

✦ They must be accepted by the stakeholder group which they are representing.

✦ They must live or work within the geographical area being considered.

✦ They should possess good local knowledge.

✦ They are available for consultation and able to attend all the workshops.

When discussing their participation, explain to each stakeholder exactly how they will be involved and the commitment required of them.

Step 3: Initial stakeholder group consultations

Conduct these discussions with each group of representatives in turn (i.e. do not bring all the stakeholder groups together yet). Begin by explaining what you are trying to do, following your response to point 1 in Step 1. Ask them to comment on your objectives *as they relate to the group the stakeholder represents.* Examples of questions to ask include:

✦ do they think that achieving these objectives is important?

✦ what other objectives do they think are important?

✦ do they think that these are more or less important than your objectives?

Depending on the responses to these questions, you may wish to widen your original objectives. When deciding whether to do this, remember that it will be very difficult to implement a management plan without stakeholder support.

Next, ask them to describe the ways in which they think the agreed management objectives can be achieved. Allow the discussion to proceed freely. Ensure, however, that you elicit enough information to be able to complete Step 4. To help you do this, think in terms of “cause and effect”. If a stakeholder mentions something and the answer is not obvious, ask “What causes that?” and “What effect does that have?”

Once the stakeholders have had the opportunity to discuss their ideas, suggest the management interventions you identified in response to point 2, in Step 1 (if they have not already been mentioned). Allow the stakeholders to comment on whether they think your favoured interventions will work and encourage them to explain why. Unless they ask, wait for an initial response before explaining the reasoning behind your favoured interventions.

It is important not to influence their comments at this stage, as you need to understand how people might react in response to the implementation of management plans. Of course, part of an intervention might involve explaining the reasoning behind it to all those affected, so it is also useful to see whether such an explanation would be considered convincing.

Take careful notes of all questions and answers and, if possible, record the discussions on tape.

Step 4: Construct preliminary BNs

The objective of this step is to capture the information that you have elicited from the stakeholders in the form of a BN diagram. This is not strictly necessary, but it is useful to do this for two reasons:

1. The logic imposed by the BN will highlight ways in which you may not have fully understood what the stakeholders were telling you.

2. Capturing stakeholder information in a BN allows you to communicate it to others more easily.

This should be done for each stakeholder group based on the discussions in Step 3. This is best done as soon as possible after the discussion. You will use this BN in Step 5 to check that you have understood what the stake-holders have told you. If this is not the case, then you will be able to adapt it.

Follow the general structure presented in Chapter 3. Choose variables and states together to represent the things you have discussed with the stakeholder groups and then try to place them into one of the six categories. When you have done this, arrange them according to the general network structure and link them (following the logic described by stakeholders: see Example 4.1).

Where the factors raised by the stakeholders in Step 3 can be grouped together then do so, making them parents of a single child, which represents the thing they all have in common (divorcing). In terms of the BN structure, this will help reduce the number of parents any one child has and will make subsequent steps easier. It also helps clarify the perceptions of the stakeholders and allows their ideas to be communicated more effectively (see Example 4.2).

When you think you have finished, check that the BN diagram is logical and complete. It is likely that as you have developed the network, you will have changed your mind about how best to represent the environmental system. Double check that you are happy with the current network by performing the following checks:

1. Identify the nodes in your diagram which have no parents

✦ these should either represent interventions or controlling factors. If neither of the above is true, then these nodes should probably have parents. Check whether the stakeholders have mentioned factors that do, in fact, affect these nodes. If they did, you should make them parents of the node being considered by adding a new node to the BN, if necessary. If they did not, you should make a note to discuss this with them in Step 5 (see Example 4.3).

2. Identify the nodes in your diagram which have no children

✦ if these nodes describe all the stakeholder management objectives, then the BN diagram is probably logical and complete (as regards the objectives, at least)

✦ if they do not then you will need to think about how nodes which do describe the management objectives can be included

✦ if you have more childless nodes than management objectives then you should see if you can link the childless nodes to management objectives (but only if you believe this to be logical)

✦ if you have more childless nodes than management objectives but cannot link them logically to a management objective you should consider whether they are important additional impacts of the management interventions

✦ if you have more childless nodes than management objectives and neither of the above two cases is true, then some of the childless nodes are probably unnecessary. In this case, they should be deleted, but check with the stakeholder groups in Step 5 before doing so (see Example 4.4).

3. Consider whether the nodes and the state names you have chosen represent things that can be sensibly described for the geographical area and time scale you chose in Step 1.

✦ if they do not, think about how you can redefine them so that they do. This will often be a case of simply changing their name or the names of their states (see Example 4.5).

Once finished, consider whether the preliminary BN diagram will be easy to use in Step 5. If it is not, then redevelop it so that it is. To help with this, think of the questions you will need to ask in Step 5. Will the stakeholders understand them and be able to answer them easily? Try different ways of naming a node to represent the idea you are trying to capture. Also try different names or values for each node state to see if it makes the questions easier to answer (see Example 4.6).

Examples for step 4

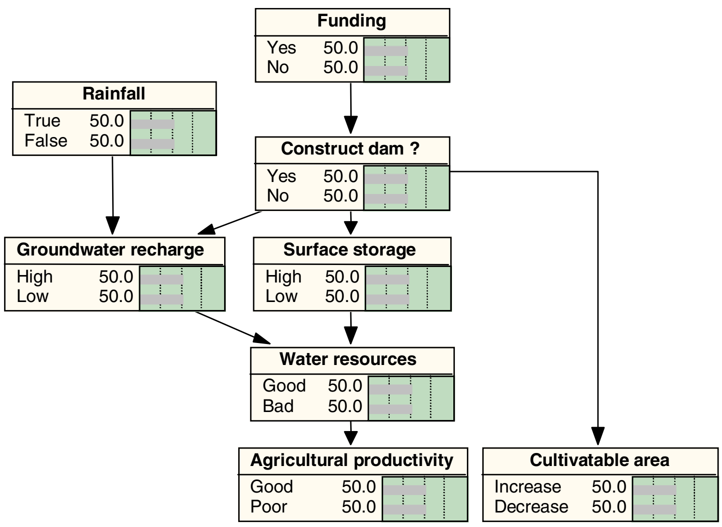
Example 4.1

Stakeholders identified increasing agricultural productivity as their main objective and the construction of a dam as their preferred intervention. When asked how constructing a dam (intervention) would affect agricultural productivity (objective), stakeholders explained that this would happen due to an increase in available water resources through improved surface water storage and increased groundwater recharge, although this would be dependent on rainfall. They also pointed out that dam construction would probably change the cultivatable area (both by removing land from production and possibly increasing irrigation command areas). Clearly dam construction would need funding to be implemented.

On the basis of this description, the variables were categorised as follows:

Variable	Category
Funding	Implementation factor
Construct dam?	Intervention
Surface storage	Intermediate factor
Groundwater recharge	Intermediate factor
Agricultural production	Objective
Cultivatable area	Additional impact
Rainfall	Controlling factor

And, following the general network structure, a simple BN could be constructed as shown below:



Example 4.2

When discussing the factors which affect yield, stakeholders noted the following, among others:

- *Fungicide application*
- *Pesticide application*
- *Heat treatment*
- *Trash burning*

While all of these could be made immediate parents of a node called "Yield", this would mean that "Yield" would have a large number of parents when the other factors relating to it were also included. To make the BN diagram clearer, the factors can be classed together as things which relate to pest and disease control. Consequently, a new node was created called "Pests and disease" which was made a child of four nodes representing the factors above and then linked directly to "Yield". This is shown in Figure 4.1.

Often there may be a number of different ways of classifying similar factors together. You should choose the way which is easiest for you and the stakeholders to understand.

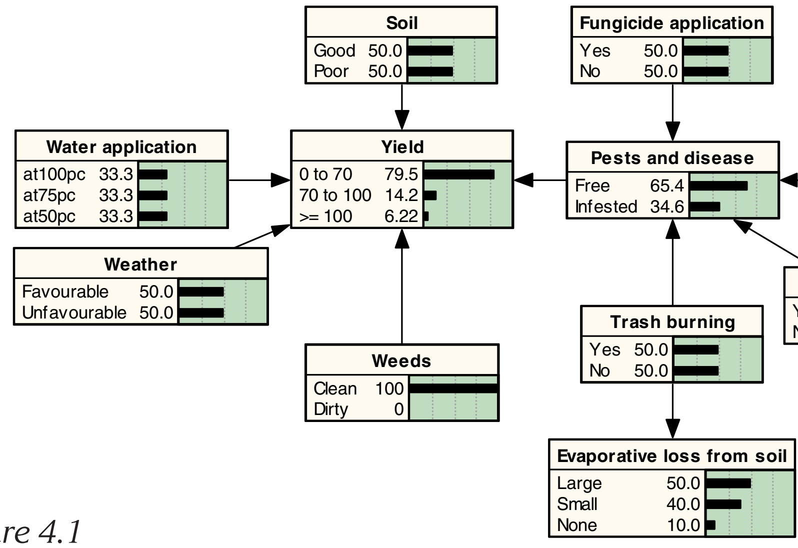


Figure 4.1

Example 4.3

Consider the Bayesian network shown in Figure 4.1. The nodes with no parents are:

- *Fungicide application*
- *Soil*
- *Water application*
- *Weather*
- *Weeds*
- *Trash burning*
- *Heat treatment*
- *Pesticide application*

Out of these, the following are the management interventions being considered:

- *Fungicide application*
- *Water application*
- *Trash burning*
- *Heat treatment*
- *Pesticide application*

Of the remainder, the nodes "Soil" and "Weather" are controlling factors.

This leaves the node "Weeds" which may appear to be an intervention. However, by examining the states, you can see that the node is meant to represent whether a field is clean of weeds or not. Obviously, this is not an intervention but a description of the condition a field is in, so you should consider which factors affect whether a field is clean or dirty.

In this case, further stakeholder consultation indicated that weed control was achieved by herbicide use, hand weeding or mechanical cultivation, so these factors were included in the BN as parents of the node weeds. As each of these new parents were potential interventions, no further changes were required. The new BN is shown in Figure 4.2.

Note that the stakeholders had originally described pest and disease control in terms of its interventions, but had described weed control only in terms of whether weeds were present or not. It is common for stakeholders to concentrate in more detail on a particular area in this way. Therefore, it is important for you to make sure that every area has received the attention it deserves.

Example 4.4

In Figure 4.1, the nodes with no children are:

- *Yield*
- *Evaporative loss from soil*

The node "Yield" represents the management objective of achieving an increase in yield. However, the node "Evaporative loss from soil" is not a management objective so there is one more childless node than management objective. It would be logical to link this to the management objective to express how reducing soil evaporation will increase soil water content and, consequently, yield for a given irrigation application. Further discussion with the farmers, however, suggested that they did not believe this to be an important factor in determining yields and so the node "Evaporative loss from soil" was deleted from the BN. The resulting BN is shown in Figure 4.2.

Example 4.5

Consider a BN which has a node in it called "Groundwater level". In many environments, this will obviously be a useful factor when considering water availability but it is a very hard thing to describe over a wide area and a long time period as it will vary so much. To avoid this problem, think about how the groundwater level actually affects your management objective. For example, if your management objective is to ensure that there is an adequate supply for domestic use then you are not interested in groundwater levels as such but in the number of days water is available in a particular well. Thinking about it like this, clearly defines both a place (i.e. a particular well owned by the stakeholder you are consulting) and a time (i.e. number of days the well is wet in the year). Don't worry that you will get different answers from different stakeholders (who may have different wells) as these differences will be accounted for later. Just make sure that the node represents an idea that is easy to quantify.

So, in this example, you could give the node "Groundwater level" states of "Good" and "Poor", where "Good" is defined as "there is water in the well for more that 300 days each year" and "Poor" as "there is water in the well for less than 300 days per year". Alternatively, you could rename the node "Days for which well is wet" and give it states of ">300" and "<300".

Example 4.6

In the stakeholder discussions prior to the development of the BN shown in Figure 4.1 and Figure 4.2, the stakeholders stated that fertiliser application also affected yield. This issue was complicated by the fact that several different sorts of fertiliser were available:

- *Manure*
- *Filter cake*
- *Single supers*
- *Nitrogen*
- *Compound D*
- *Potash*

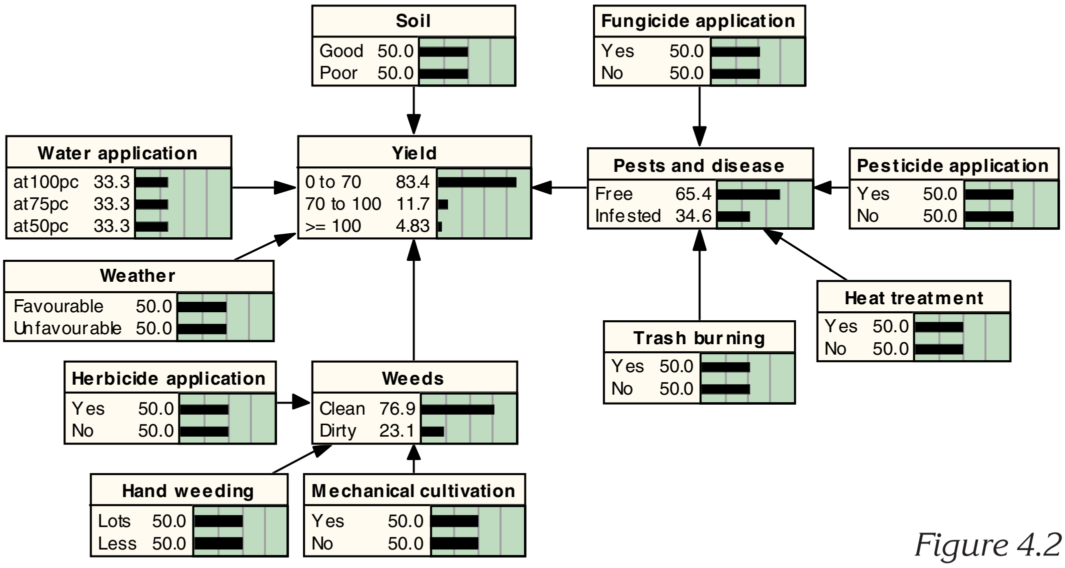


Figure 4.2

This was further complicated by the fact that there were two different sorts of nitrogen and the stakeholders considered that the amount and timing of their application was particularly important. As in Example 4.2, it was decided to group these factors. In the first instance, all the sorts of fertilisers

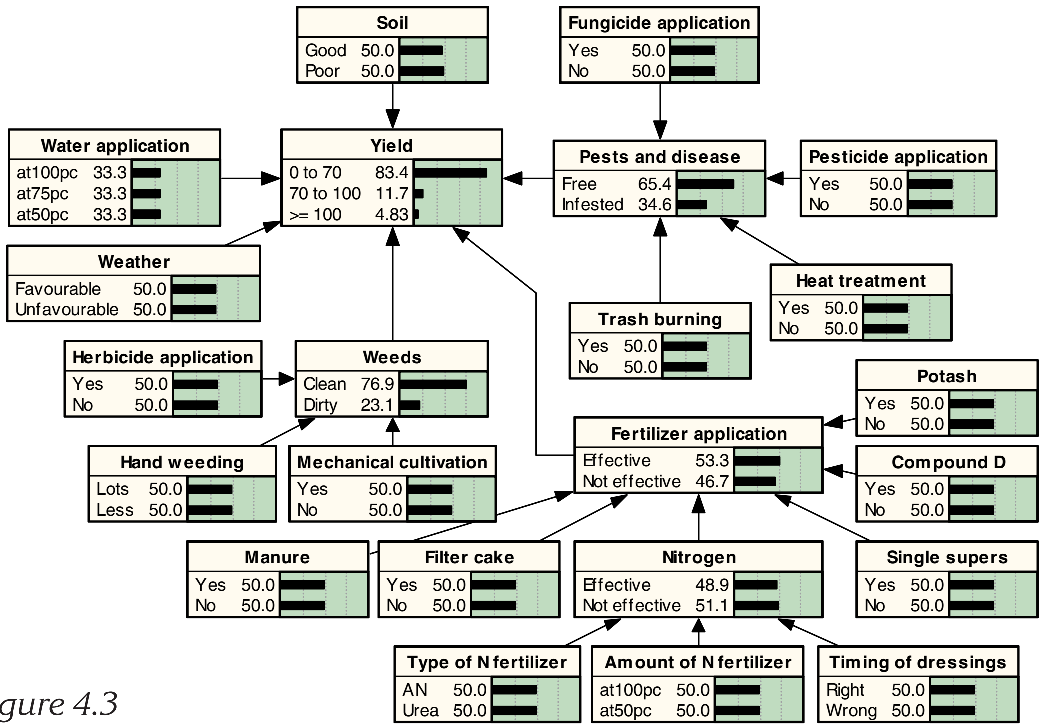


Figure 4.3

were grouped under a single child node called "Fertiliser application" and the nodes representing them were given states of "Yes" and "No" to indicate whether they had been applied or not. It seemed clearer to group the factors relating to nitrogen fertiliser (type, amount and timing) under the node representing its application (called "Nitrogen"), however the state names "Yes" and "No" no longer made sense as the node "Nitrogen" now represented a combination of the type, amount and timing of the fertiliser and not just the amount alone. Consequently, states of "Effective" and "Not effective" were chosen as this made more sense when talking about the type and timing of the fertiliser application as well as the amount. These states also made sense for the node "Fertiliser application" and so were used there as well. The resultant BN is shown in Figure 4.3.

Step 5:Further stakeholder group consultations

Arrange a meeting with each stakeholder group separately. The meeting has two objectives:

1. to check the validity of the relationships you deduced from Step 3 and included in the preliminary BN you constructed in Step 4;
2. to define the states for each node.

If you are confident that the stakeholder group will understand the logical structure provided by the BN diagram, then Objective 1 can be achieved by discussing the diagram directly with the stakeholders. Where this is not the case, the BN structure should be described by using a series of lists stating the dependencies of each node in turn (see Example 5.1).

In either case, discussions should take place with the group as a whole. It should be stressed that you are only trying to present what they told you in Step 3 and that it is important they tell you if they think you have misunderstood. You should encourage and allow plenty of opportunity for the group to do this. In this respect, it is crucial that the meeting schedule allows enough time to complete the process without needing to rush. Any changes suggested by the stakeholder group should be accepted and the preliminary BN diagram changed accordingly.

Before the general discussion begins, make sure you talk with stakeholders about any issues that arose in Step 4. Specifically:

- ✦ any parentless nodes you identified which were neither management interventions nor controlling factors;
- ✦ any childless nodes you identified which were neither management objectives nor additional impacts.

As a result of this preliminary discussion, you may wish to delete some nodes or add new ones, depending on how the stakeholders answer your questions from Step 4. If this is the case, you should adjust your lists and/or BN to reflect these changes, before moving on to the general discussion.

During the general discussion, check that the group is happy with the states you have given each node. This should be done in relation to the parents and children of the node you are discussing (as described in Key Skill 2 in Chapter 3). If they are not satisfied suggest alternatives. It is important to limit the number of states to as few as possible. Two are ideal (and can usually be achieved with imagination) however three are acceptable if limited to a small number of nodes. More than three states should only be allowed in exceptional circumstances. When the stakeholders are happy with the names, get them to define what the name signifies in terms that can be understood by people outside of the group. This will probably require some form of quantification (see Example 5.2).

Once the stakeholders are satisfied with the BN diagram, review it with them by using the questions below. This is necessary to make sure that the links in the preliminary BN you built in Step 4 properly represent how the stakeholders see the variables working together.

Starting at the management objectives, think about the relationships that a child has with its parents by asking the following two questions for each of its parents. Don't ask these questions exactly as written. Instead adapt them to use the names of the nodes and states being considered. This should make the question much clearer. Do this for each child in the network.

1. Think about the effect that changing the state of this parent node would have on its child. Could that effect be altered by any other factor (see Example 5.3)?

✦ If the answer to question 1 is yes, ask which factor could affect it. If that factor is already represented as a node in the BN and is a parent of the child being considered then there is no need to change the BN. You should note which parent is having the effect as this determines how the CPTs should be filled in (see Appendix 2). Such a node will be referred to as a “modifying parent”.

✦ If the factor is already a node in the BN but not a parent of the child, then it should be linked to the child (note it as a modifying parent, as above).

✦ If the factor is not already in the BN, then it should be added and linked to the child being considered. Remember to discuss what states it should have with the stakeholders (note it as a modifying parent, as above).

2. Would fixing one of the parent nodes in any one of its states remove the effect that changing the state of another parent has on the child (see Example 5.4)?

✦ If the answer to question 2 is yes, then this indicates that the structure of the BN is not entirely logical. To correct it, identify the parent whose effect is removed (let us call this the “nullified parent”) and break the link between it and the child. Identify the parent who caused the nullified parent to become so and link this together with the nullified parent (now separated from its original child) to a new node. Make this new node a parent of the original child and ask the stakeholders to give it a suitable name and suitable states.

Examples for Step 5

Example 5.1

Take the BN in Figure 4.3 as an example. Start with the node representing the management objective: "Yield" in this case. At the top of a piece of paper write "Yield" as a heading and then list underneath it all the nodes on which it is directly dependent (i.e. its parents). Then do this for each of the parent nodes in turn and then for their parents, moving back through the BN from the management objective node to the intervention nodes. This will produce a series of lists, as follows:

List 1 ***Yield*** Soil, Pests and disease, Fertiliser application, Weeds, Weather, Water application

List 2 ***Pests and disease*** Fungicide application, Pesticide application, Heat treatment, Trash burning

List 3 ***Fertiliser application*** Potash, Compound D, Single supers, Nitrogen, Filter cake, Manure

List 4 ***Weeds*** Mechanical cultivation, Hand weeding, Herbicide application

Present each of these lists in turn to the stakeholders, starting with list 1. Ask them to confirm that the factors listed under each heading directly affect the heading. Allow them to add other factors and remove any you have included.

Example 5.2

In Example 4.6, states of "Effective" and "Not effective" were given to the node "Fertiliser application". During the subsequent workshop, the stakeholder group agreed that these were sensible names and defined them as follows:

Effective: Fertiliser is applied such that the yield is above 70 tonnes/hectare, all other things being equal.

Not effective: Fertiliser is applied such that the yield is less than 70 tonnes/hectare, all other things being equal.

In Example 4 the states of "Weeds" were defined as "Clean" and "Dirty". These were defined as follows:

Clean: Small numbers of weeds seen in field

Dirty: Large numbers of weeds seen in field

While this second definition is much more qualitative than the first, the stakeholders felt that the concept of a clean field was a very easy one to understand.

Example 5.3

Consider the BN in Figure 4.3. When asked whether any of the parents of the node "Pests and disease" would be altered by any other factor, the stakeholders stated that the use of pesticides and trash burning both became more important when the soil was poorly drained. The factor "soil" was already present in the BN but was not a parent of "Pests and disease" and so it was necessary to link the two accordingly (Figure 4.4). The node "Soil" was noted as a modifying parent.

When asked whether any parents of the node "Nitrogen" could be altered by another factor, the stakeholders stated that getting the timing of fertiliser

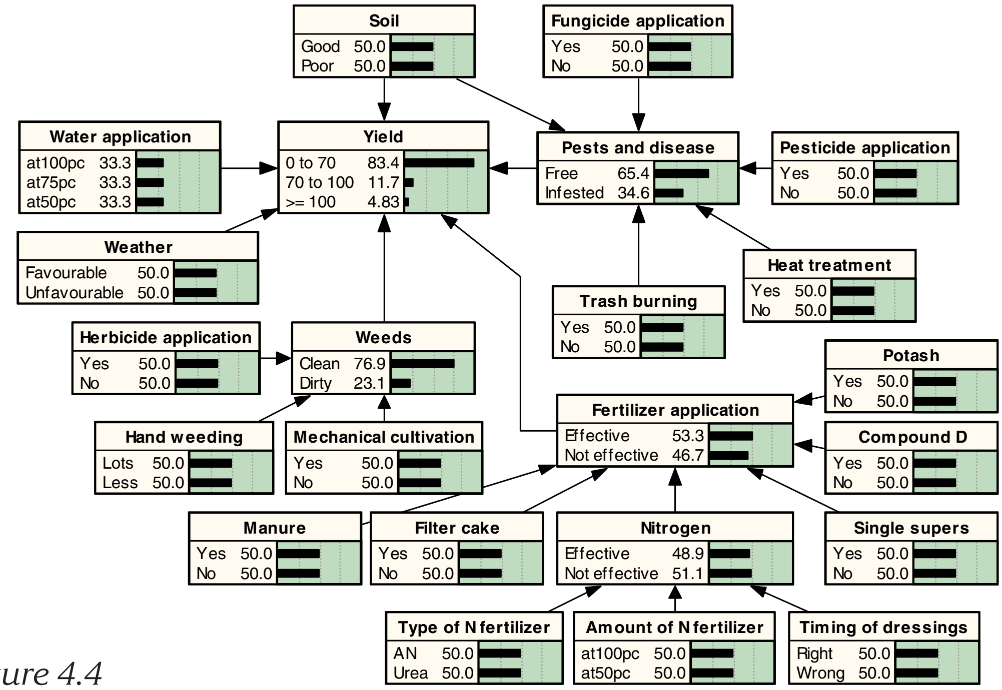


Figure 4.4

dressings wrong is much more important when using urea than when using AN. The node "type of fertiliser" was already present in the BN and was already a parent of Nitrogen, so no action was taken, although it was noted as a modifying parent.

Example 5.4

Consider the BN in Figure 4.5. When asked if any of the parents of the node "Water delivery down mainline" would remove the effect of changing the state of another parent, the stakeholders stated that position would have no effect if flood irrigation were used. Consequently, the structure of the BN was changed as shown in Figure 4.6.

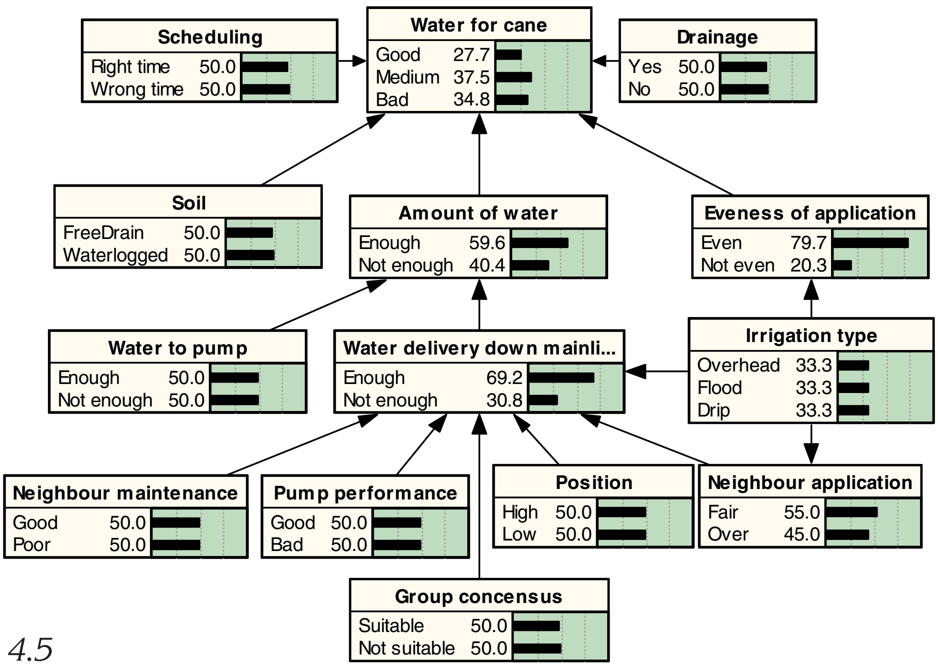


Figure 4.5

The stakeholders also stated that "Irrigation type" was a modifying parent for all the other parents of "Water delivery down mainline". Therefore, the link between it and "Water delivery down mainline" was not deleted.

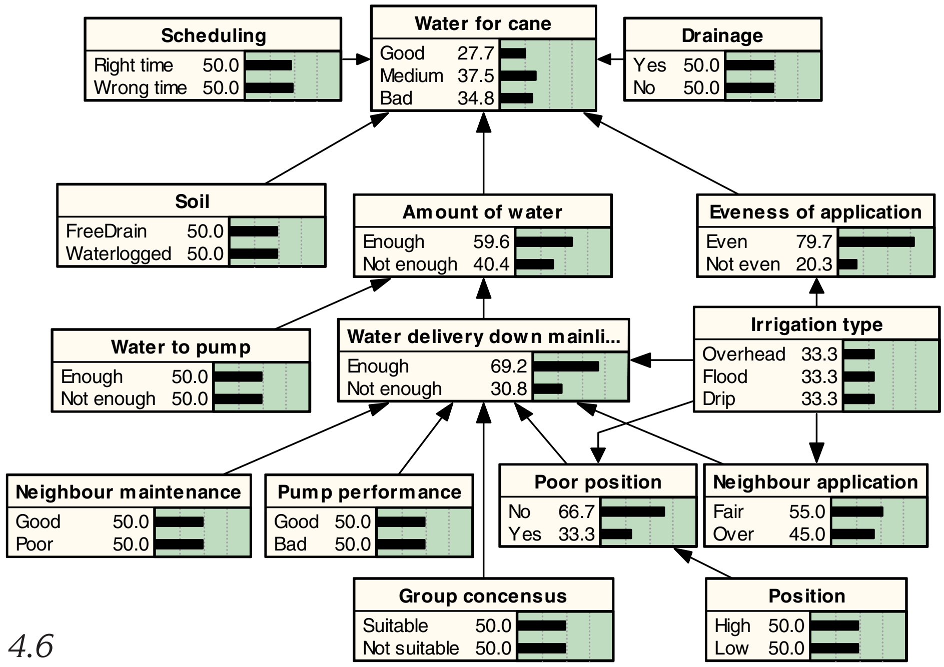


Figure 4.6

Step 6: Draw conclusions from stakeholder consultation

Following completion of Step 5, you should have a BN diagram accurately representing the perceptions of each stakeholder group. It is possible to use these diagrams to identify:

1. potential issues of consensus between groups
2. potential issues of conflict between groups
3. any remaining questions you have about how the stakeholders see the proposed management strategies affecting themselves and the environment they live in.

When you have identified them, these issues will be discussed at a stakeholder workshop (see Step 7) to find out the reasons for them.

As in Step 4, categorise the variables according to the scheme presented in the introductory notes (you may have changed your mind about where some of the variables belong since Step 4). Then, for each stakeholder BN, write down the objectives as headings and list the interventions that affect them underneath. If an intervention affects an objective, you should be able to follow a continuous path of nodes and links from intervention to objective. Highlight those interventions that affect an objective in more than one way (i.e. there is more than one path between intervention and objective) and also those which affect more than one objective (i.e. they appear in more than one list) (see Example 6.1).

Use these lists, the BN diagrams and the questions in Table 4.2 to compare stakeholder viewpoints.

Table 4.2: Questions to guide analysis of stakeholder BNs

Question	How to answer
Questions related to objectives:	
1. Do objectives appear to be different because different groups have expressed the same objective in different ways?	Examine the nodes surrounding the objective in the BN diagram, to see the context in which the stakeholders have placed the objective (see Example 6.2).
2. Do objectives appear to be different because one group is taking a broader perspective than another?	Examine the BN to see if one group's objective is a parent (or grandparent) of another group's. Also look at the context to see if any group is considering two objectives from another group as a single objective (see Example 6.3).
3. Do apparent differences genuinely reflect different stakeholder objectives?	Probably true if neither of the above are true.
4. Are objectives that appear to be the same, genuinely the same?	Examine the states of each node as well as their context (see Example 6.4).
5. Where objectives are the same, what level of achievement do the stakeholders expect?	Check the state definitions for the objective node.
Questions related to interventions:	
6. Where 2 stakeholder groups have the same objectives, do they agree on the interventions that affect those objectives?	Compare your lists of objectives for each group.
7. Where interventions appear to be the same for two stakeholder groups, do they affect the objective in the same way?	Check the chain of nodes between the intervention and the objective to see if they are substantially the same (see Example 6.5).
8. Where interventions appear to be different, is this because they are genuinely different or because groups have expressed themselves in different ways or taken broader perspectives?	Examine node contexts as described for objectives, above.
9. Do stakeholder groups agree on which interventions can affect a single objective in more than one way?	Compare those interventions that you have highlighted in the lists for each group.
10. Do stakeholder groups on which interventions affect more than one objective?	Compare those interventions that you have highlighted in the lists for each group.
Questions related to implementation factors:	
11. Where stakeholder groups have common interventions, do they agree on the factors required to implement them?	Compare the lists of implementation factors that you have drawn up for each group.
Questions related to additional impacts:	
12. Where stakeholder groups have common interventions, do they agree on the additional impacts that may arise from implementing them?	Compare the lists of additional impacts that you have drawn up for each group.
Questions related to controlling factors:	
13. Do the stakeholder groups agree on those factors that cannot be changed by intervening at this scale but are still important?	Compare the lists of controlling factors that you have drawn up for each group.

Examples for step 6

Example 6.1

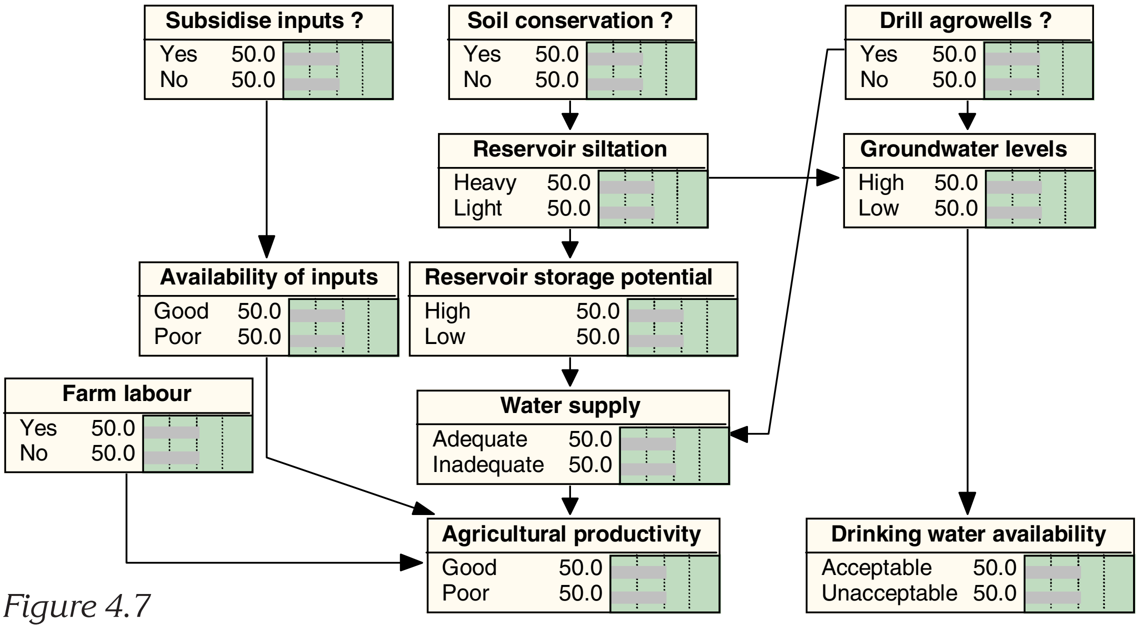


Figure 4.7

Using Figure 4.7 as an example, the nodes could be classified as follows:

Variable	**Category**
Subsidise inputs?	Intervention
Soil conservation?	Intervention
Drill agrowells?	Intervention
Reservoir siltation	Intermediate factor
Groundwater levels	Intermediate factor
Availability of inputs	Intermediate factor
Reservoir storage potential	Intermediate factor
Farm labour	Controlling factor
Water supply	Intermediate factor
Agricultural productivity	Objective
Drinking water availability	Objective

and the following lists would be compiled:

Agricultural production	**Drinking water availability**
Subsidise inputs	Soil Conservation
Soil conservation	Drill agrowells
Drill agrowells	

"Soil conservation" and "Drill agrowells" are highlighted as they affect more than one objective.

Example 6.2

Figure 4.8 shows two BNs with seemingly different objectives: the one at the top has an objective of "Yield", while the one at the bottom has an objective of "Agricultural productivity". Looking at the nodes surrounding the objectives, however, suggests that they are, probably, the same. It is unlikely that *the stakeholders were trying to represent different ideas. Rather, they simply chose different names to represent the same idea.*

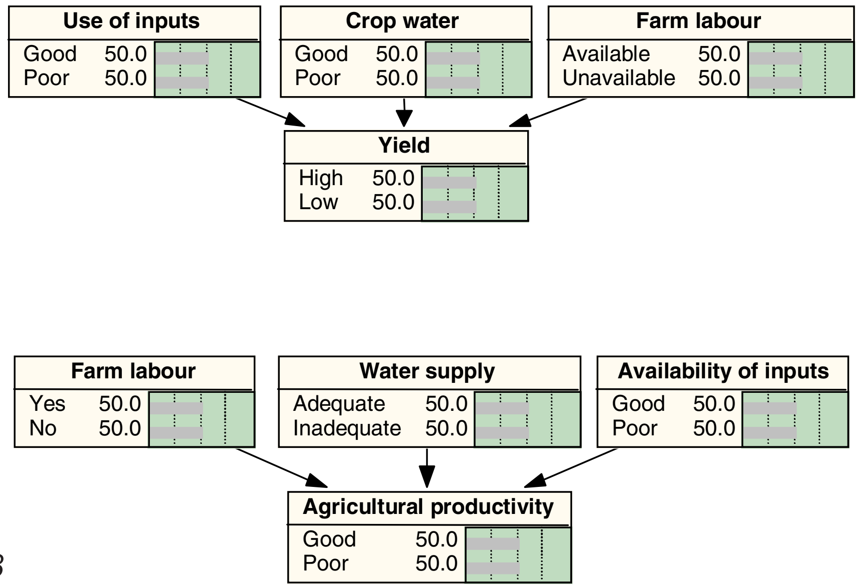


Figure 4.8

Example 6.3

Figure 4.9 shows two BNs with seemingly different objectives: the one at the top has an objective of "Agricultural productivity", while the one at the bottom has an objective of "Farmer income". However, it can be seen that, in the bottom BN, "Agricultural productivity" is a parent of "Farmer income". Therefore, while these BNs have objectives which are definitely different, it is likely that the different stakeholder groups were trying to express the same idea — the group who produced the BN at the bottom simply took a wider perspective on the problem. Note, however, that the introduction of a "market price" variable may have important consequences for the success or otherwise of any interventions.

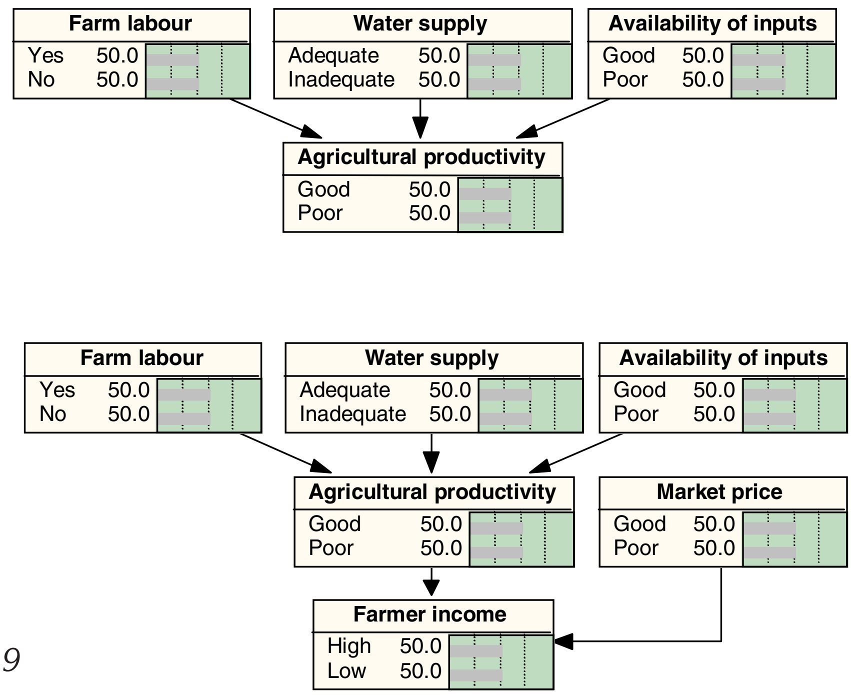


Figure 4.9

Example 6.4

Figure 4.10 shows two BNs that seem to have the same objective: they have the same names and the same states. However, looking at the context of the node suggests that the top network represents a surface water system while the bottom one represents a groundwater system. In this case, it should be concluded that the stakeholders are considering different objectives.

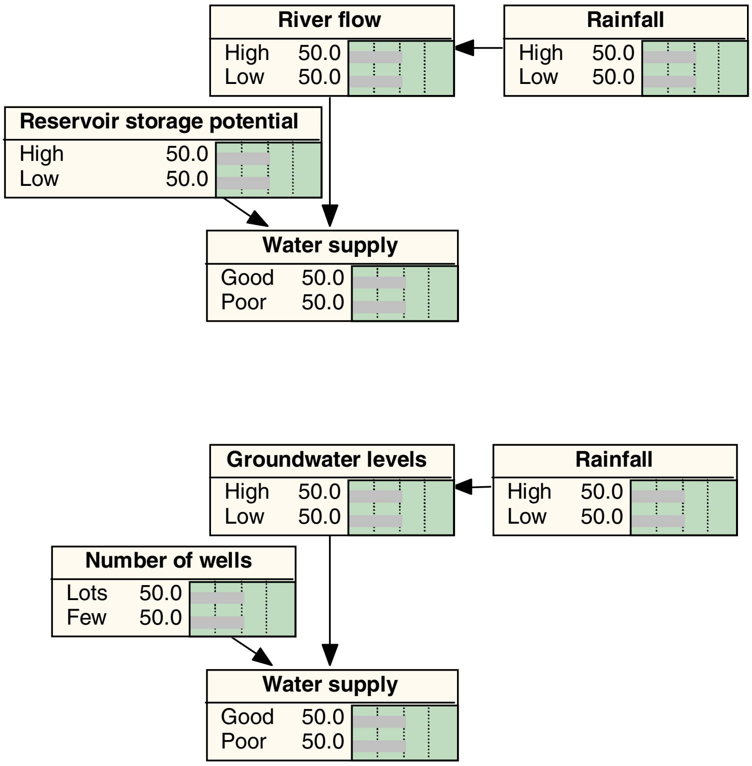


Figure 4.10

Example 6.5

Figure 4.11 shows two BNs that have the same interventions and objectives (the intervention being the handover of water management to farmer control). However, studying the chain of intermediate nodes between intervention and objective (only one, in this case) suggests that each stakeholder group has different ideas as to how the intervention will affect the objective. The stakeholders who produced the BN on the left believe that farmer control will affect the timeliness of water delivery while the group who produced the BN on the right believe it will affect the fairness of water distribution. Clearly, both these ideas may be true and it will be important to check this with all stakeholders.

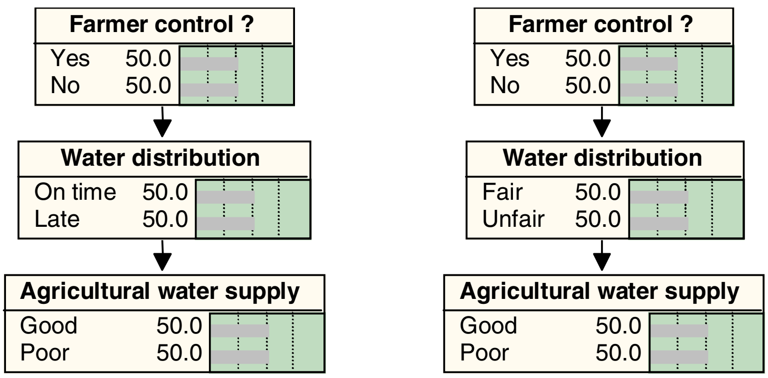


Figure 4.11

Step 7: Hold joint stakeholder workshop to discuss differences in viewpoints

If the stakeholder BNs have been constructed carefully, they should capture the perceptions of each stakeholder and help identify the areas of agreement and disagreement between groups in a formal manner. Therefore, the objective of this workshop is:

- to allow different stakeholder groups to discuss the reasons for any differences in the interventions they favour

All stakeholder groups should be present at this workshop so that they can discuss the issues together.

As with the previous workshop, begin by clarifying any issues that have arisen since the previous workshop through the analysis done in Step 6. Having done this, present the main areas of agreement and the main differences you identified between groups in Step 6. These should relate to:

- objectives – what they are and the level of achievement expected
- interventions – what they are, what objectives they affect, how they affect those objectives, which interventions are most important and why
- implementation factors – what they are and which interventions they affect
- additional impacts – what they are and what causes them
- controlling factors – what they are and what they control

The emphasis at this stage is on explanation rather than argument. You want to understand the reasons for conflict rather than resolve it, although this may also happen. You should facilitate this by discussing the reasons underlying each stakeholder group's perspective, which are revealed by the BNs (this should have been done, to a large extent, while the workshop's first objective was being achieved). Make the stakeholders aware that they can agree to disagree, as long as they can explain to you their reasons for doing so.

Begin by confirming with the stakeholders that the areas of agreement you have identified are correct. If any further disagreements arise at this stage discuss and note the reasons why. Next, discuss the remaining differences. Check with the stakeholders that the differences are real and not just a consequence of expressing the same idea in different ways. If the differences are real, ask the stakeholders whether they think these differences would seriously affect the implementation of the interventions that relate to them. Discuss possible compromises with them.

At the end of the workshop, explain what the next steps in the decision making process will be and how the information they have provided will be used to make a final decision as to which management plan to adopt.

Step 8: Complete stakeholder BNs

It is likely that the workshop in Step 7 will have identified ways in which the stakeholder BNs do not adequately represent the perceptions of each stakeholder group. If this is the case, you should alter the stakeholder BNs to reflect the conclusions reached during the second workshop. Do this by adding and deleting nodes and links as appropriate. Repeat the checks you performed in Step 4 to ensure the logic of the network is consistent.

When this step is completed, the stakeholder BNs will then provide a record of the viewpoint of each stakeholder group that can be investigated and referred to during the following steps.

Step 9: Construct 'master' BN diagrams

A 'master' BN diagram is one which you, the decision maker, will use to choose the management interventions you consider to be the most likely to achieve your objectives. When it is completed (following Step 10), it will be a fully-functional BN that you can use to develop your understanding of the environmental system.

Initially, you should develop a single master BN diagram based on a combination of the stakeholder BN diagrams and your own understanding of how the environmental system works. Remember that you should begin by ensuring that you capture the basic logic of the whole environmental system you are considering (see Key skills in Chapter 3). You should include all the objectives and interventions suggested by the stakeholders (plus your own) unless they can be clearly ruled out as unworkable, given the current situation. You may wish to alter the number of intermediate nodes suggested by the stakeholder groups (although this issue will be examined in more detail in Step 11). Ultimately, the BN should represent your understanding of what the issues are and how they can be solved, although it should be influenced and informed by the stakeholder BN diagrams that you have already built. If it is not, then you will risk failing to implement your management plan because of a lack of stakeholder cooperation.

You should make sure that you represent wider issues that may not have been raised by the stakeholders. Three important examples are discussed below, together with ways of including their analysis in the master BN.

Equity

Each stakeholder group should have only considered the impacts of management interventions on itself. As you are responsible to all the stakeholder groups, however, it is important that you consider the impacts as they affect all stakeholders. Achieving as equitable a distribution of impacts as possible may be an appropriate objective in itself.

To do this, it might be necessary to replicate your master BN diagram for each stakeholder group. This may seem pointless, as the diagrams will all be the same, but it is likely that the conditional probabilities that will be entered in Step 10 will need to be different, unless each group focused on different management objectives. When this has been done, you will have a fully-functioning BN for each stakeholder group. They will look the same but will provide different outputs indicating the different chances that you can achieve your objectives for each stakeholder group (see Example 9.1).

It is likely that the majority of these master BNs will not be greatly different. Where they represent factors that affect all stakeholder groups (such as rainfall, for example) there will be no need to change the CPTs. You will only need to change them when factors affect different stakeholder groups in different ways.

When you come to using the BNs to make a decision (in Step 11) you will need to use each in turn and note the differences in their responses before making a final decision.

Sustainability

From a practical viewpoint, the idea of sustainability means that the benefits you hope to gain from your management plan can be maintained in the long term without causing severe adverse impacts on the environment as a whole. Therefore, to assess the sustainability of your management plan, you must consider the impact it will have over a number of years on your objectives and any additional impacts. The way to do this has been described in Key Skill 4 in Chapter 3.

Depending on the nature of your intervention, it may also be important to track the changes in the implementation nodes. If your intervention is a "one-off" (see *Key Skills: A general network strategy* in Chapter 3) then changes in the implementation nodes may not directly affect sustainability. However, if your intervention is being implemented over a longer time period, then it is important to check that the implementation nodes remain in a state that continues to support the implementation of the intervention.

Unfortunately, the number of time steps you can include is limited by the complexity of the BN. When BNs are replicated in this way, they can quickly become more complex than can be handled by even the most powerful of computers. In practice, however, you will have no problem in representing at least the two time steps recommended in the introductory notes.

Wider human and environmental consequences

Carefully consider whether all the additional impacts of your proposed management plan have been included. In particular, consider the impact of your plan on all aspects of peoples' livelihoods and on the natural environment. If there are impacts which you have not yet considered, add nodes into the BN to represent them.

When you think you have finished the BN diagram, think again about your responses to the points in Step 1, including any changes you have made following the stakeholder consultations. Check if the BN you are constructing meets your requirements in terms of the decision you have to make.

Examples for step 9

Example 9.1

Figure 4.12 shows two BNs representing the impact of two interventions on each of two different stakeholder groups. The BN on the left represents a group of farmers largely dependent on a reservoir for irrigation water and situated on land overlying a low-yielding aquifer. The BN on the right represents a group of farmers whose land is situated above a high-yielding aquifer. While the BN diagrams are the same, the CPT associated with the node "Agricultural water supply" in the left hand BN reflects the fact that installing wells in a low yielding aquifer is likely to have a much lower impact than doing the same in a high yielding aquifer. Consequently, the probabilities that each group will have sufficient water are different (30% for the farmers on them left as compared to 70% for the farmers on the right).

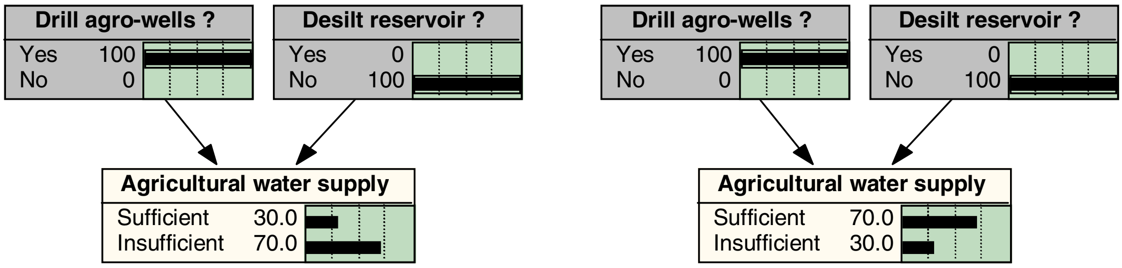


Figure 4.12

Step 10: Collect data and specify conditional probability tables (CPTs)

Having constructed the master BN diagram, it must now be turned into a fully-functioning BN that can be used to help make decisions. This is done by filling in the CPTs using the best and most appropriate data available.

A CPT underlies every node in your BN. To fill them in, you will need data linking each node to its parents. In other words, the data in the CPT must describe how a node changes in response to changes in its parents (see Example 10.1). Each row in a CPT implies a question. For example, referring to Table 3.1, the question for the first row would be: "If forest cover is good and rainfall is good, what is the chance that river flow will be good, acceptable or bad?" If you find it difficult to frame this question, then it is likely that your master BN diagram is illogical and you may need to alter either the structure or the states of the nodes represented in the CPT. If you find it easy to frame the question, then it is likely the logic of your master BN is correct.

Each of these questions (for each row of the CPT) suggests the data you need to collect to fill in the CPT. You should consider carefully the best place to get these data from. In some cases, the best information may come from stakeholders. This will usually be true when a node represents how people react to changes in the environment. For example, a node may represent how much more fertiliser farmers would use if their income was increased. The best way to find this out may be to ask farmers. In other cases, the best answers may come from more objective data that are available or can be collected.

In general, four types of information can be identified. These are:

Information Type 1 Raw data collected by direct measurement (e.g. groundwater depth measured by piezometer, population measured by census, income measured by accounting).

Information Type 2 Raw data collected through stakeholder elicitation (e.g. stakeholder perceptions of groundwater depth, population and income).

Information Type 3 Output from process-based models calibrated using raw data collected by direct measurement.

Information Type 4 Academic "expert" opinion based on theoretical calculation or best judgement.

Use the guidelines below to help you decide which is most appropriate. If you are using raw data, remember that the more information you have, the lower the uncertainty will be. Also remember that if you have a master BN for each stakeholder group (to allow you to consider equity issues — see Step 9), then information of Type 2 must come from the stakeholders for whom the BN had been constructed.

- If you have enough raw data to avoid high uncertainties, always use Type 1 in preference to Type 3, and Type 2 in preference to Type 4.
- If you have enough raw data of Type 1 and of Type 2 to answer a question, you need to consider carefully which it is best to use. In general, information Type 1 is more reliable than Type 2 but there may be cases when Type 2 is more appropriate. For example, there may be areas in which your decision should be based more on how stakeholders perceive changes taking place than on how they might actually take place.
- Where you have enough information of Types 2 and 3 to answer a question, you should consider which is better as above. In the Type 3 case, you should also think about how accurate you believe the model output to be.
- If there is insufficient information of Type 1 and you consider use of Type 2 to be inappropriate, you should use information Type 3, if suitable models are available.
- If no other information type is available, use information Type 4.

If you have the time and the resources you may wish to ensure you have enough information of Types 1 and 2 by carrying out suitably designed data collection programmes. However, in many cases, collecting enough information of Type 1 will take longer than you are likely to have available.

There are different approaches you can use for filling in a CPT depending on the type of information you decide to use. These approaches require some basic mathematical manipulation and, while software has been provided to help with this, it is recommended that you become familiar with the calculations it performs. These calculations are described in the appendices to these guidelines:

- For information Types 1 and 3, refer to Appendix 1
- For information Types 2 and 4, refer to Appendix 2

N.B. Make sure you are certain how much data you need and what they need to describe before proceeding with data collection. Understanding the information in the appendices will help you to do this. Being clear about this will save you a lot of time and money.

When the BN represents multiple time steps, remember that parents in time steps after the first one, may have additional children to represent changes which occur between time steps. In this case the CPTs will have changed and you will need to find additional information to complete them (see Example 10.2).

There is no need to calculate probabilities for those nodes that have no parents. If the BN has been structured correctly, then these nodes will represent either factors that you hope to control directly through management or factors over which you have no control at all. When you come to use the BN to make a decision, you will change the states of these nodes to see how this affects your management objectives. However, for the BN to work, you do need to fill them in. Do this, in the first instance, by giving each possible state an equal probability. You will later be able to change these to see what effect this will have on your management objective.

Examples for step 10

Example 10.1

A network from Example 6.4 is repeated in Figure 4.13. To fill in the CPT for the node "Water supply", you will need information describing how the water supply varies in response to groundwater levels and the number of wells. The exact information you will need is defined by the state combinations of the parent nodes (references to information types are explained in the main text).

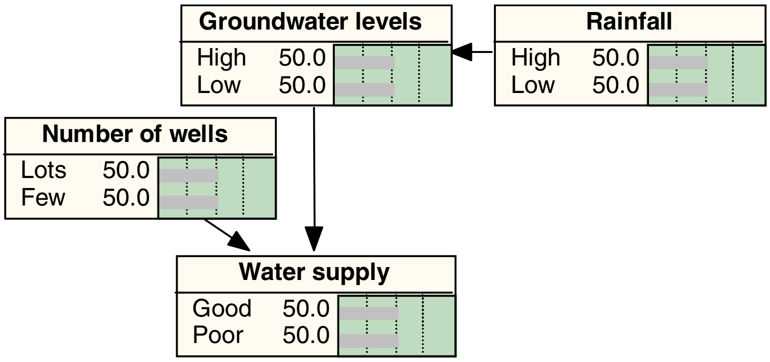


Figure 4.13

Remember, that by this stage, you should have given the state names quantitative definitions. For example, you may have decided to define the states of "Water supply" as follows:

Good: 90% or more of farmers are able to meet crop water requirements

Poor: Fewer than 90% of farmers are able to meet crop water requirements

In the example below, it is likely that measured data will only be available for when the variable "Number of wells" is in the state "Few" (assuming this is the current state). Information describing the changes likely to take place when more wells are installed will probably be of types 3 or 4.

Groundwater levels	**Number of wells**	**Water supply**
High	Lots	Information indicating whether water supplies are good or poor (the more, the better). This is most likely to be available as an output from a model (information type 3) or an assessment by an expert (information type 4).
High	Few	Information indicating whether water supplies are good or poor (the more, the better). Ideally, this will be available in the form of time series data measured over a range of groundwater levels (information type 1). If this is not available, it may be appropriate to use the estimates of local people (information type 2).

Low	Lots	Information indicating whether water supplies are good or poor (the more, the better).
Low	Few	Information indicating whether water supplies are good or poor (the more, the better).

Example 10.2

Figure 4.14 shows a BN with two annual time steps. In the second time step, the node "Fertiliser application" has gained a parent that it did not have in the first one, namely "Farmer income". Consequently, information is required to describe how fertiliser application might change in response to changes in farmer income.

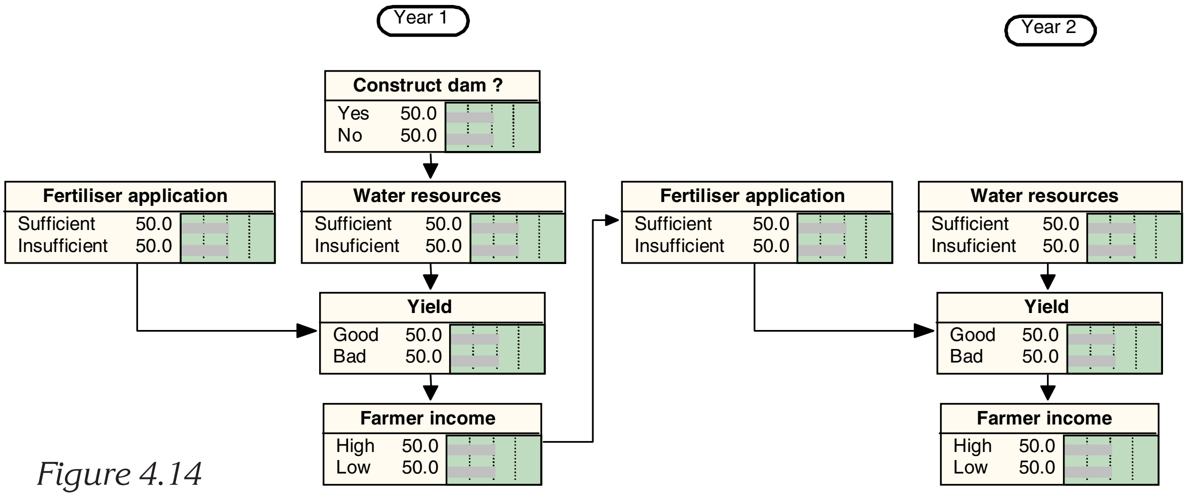


Figure 4.14

Step 11: Use master BN to make decision

To use the BN, you will change each of the intervention nodes and examine the effect it has on each of the objective nodes. This effect will be shown in terms of the change in the probability that the desired state will be achieved. (see Example 11.1).

It is crucial, however, that you also use the BN to understand why these changes are taking place.

If you have taken care to ensure that the conditional probability tables contain the best information available, then the changes in the probabilities that you see in the objective nodes, will provide a good quantitative guide. However, these values should be treated with caution until you are certain you understand how they arise.

The real value of the BN lies in the way it helps you understand your management problem in a more integrated way. It should be used as a "tool for thinking" not an automatic answer provider.

To achieve this understanding, follow the approach outlined below. It may seem complicated but, once you are used to it, it is fairly straightforward. Some software packages allow this process to be automated to an extent (e.g. the "Sensitivity to Findings" function in Netica) but, before using these, be sure you understand the information they provide. Please refer to the software manufacturers instructions for further details.

1. If you have more than one time step, start by concentrating only on the first one. Set all the intervention nodes to the state they are currently in (i.e. before any proposed intervention has taken place). Also set all controlling factors to the state you think they are currently in. If you are not certain which state a particular node should be in (either an intervention or a controlling factor), then give different probabilities to each of its states (see Example 11.2).
2. If you have multiple management objectives, start with the one that you think is most important. Change the states of each of the management interventions in turn and note the change in the probability that the management objective is in the state you would like it to be in (i.e. the desired state). When you have noted the impact that changing the state of the intervention has on the desired state of the management objective, set the intervention node back to its current state. Write all the interventions down in a list, together with the change brought about in the management objective node by each intervention node (i.e. the increase or decrease in the probability that the management objective is in its desired state — see Example 11.3).
3. Study your list and, starting with the intervention that has the most desirable impact, ask yourself why it has a greater effect on the management objective than the other interventions. There will be two principal reasons for this – either the structure of the BN favours it or the conditional probability tables favour it. This is discussed below:

Considering the structure

Generally, if distance is measured in terms of the number of nodes lying between the intervention and the objective, then those interventions that are closer to the objectives will have a greater impact. This is because each intermediate node has a conditional probability table, which introduces more uncertainty into the effect that the intervention ultimately has on the objective. It is also likely that some of the intermediate nodes will have additional parents representing factors which must also be considered when investigating how the intervention affects the objective. This will “dilute” the effect that a particular intervention will have on an objective.

Both of these features of BN structure will tend to reduce the impact that the intervention has on the objective and you should ask yourself whether they represent the real situation properly. For example:

- Does the intervention genuinely have as direct an effect on the objective as you have represented it in the BN (i.e. do you think that the number of intermediate nodes between the intervention and objective is representative of the real situation)?
- Have you oversimplified the BN and left out additional factors that should be considered (i.e. does the effect of the intervention on the objective need diluting)?
- Is the structure surrounding the other interventions (which the BN suggests have less impact on the objective) too complicated in comparison to this intervention (i.e. do the less effective interventions have a greater number of intermediate nodes between them and the objective? See Example 11.4).

Depending on your answers to these questions you may wish to change the

BN structure to represent the real situation more accurately. Where you think additional factors should be considered you should add new nodes. Where you think the structure is over-complicated, you should remove the nodes that are least relevant. When doing this, you should ensure that no important details are lost.

Considering the CPTs

The values contained in the CPTs will determine the strength of the impact that an intervention has on an objective. If you have taken care in calculating these values, then you should be reasonably satisfied with them. However, the combined effect of the CPTs underlying the intermediate nodes can often be surprising.

Examine this by changing the state of the intervention node and noting its effect on its immediate child (i.e. the one linked directly to the intervention node). When you are satisfied that you understand why the change you have observed is taking place, change the state of the child directly and note the effect on *its* immediate child (i.e. the grandchild of the intervention). Proceed in this way until you reach the management objective. Take particular care to study those nodes which have additional parents to the one which you have already considered. You may need to change the states of these nodes too before you can really understand why the changes you observe are taking place (see Example 11.5).

If you are unhappy about any of the changes you observe, then study the values in the CPT directly to see what is causing the problem. If there is a particular combination of parent states which produces a change you think is wrong, then find the values relating to that state combination in the CPT of the child (see Example 11.6). Find where the data came from that were used to specify these values and check that they are correct. If they are correct, then you should accept that the BN is correctly representing this change.

During this process, some of the child variables you are examining may change so that the probabilities for each of their states are roughly equal. When this happens, note the combination of parent states which have produced this change and consider carefully what this result may mean. It may mean that for that particular combination of parent states, you believe the child has an equal chance of being in any of its states. On the other hand, it may suggest that you had insufficient information to fill in the row of the child's CPT for that particular combination of parent states (see the discussion on "Uncertainty about uncertainty" in Appendix 1). You should make sure you know which of these two possibilities applies to this particular combination of parent states before proceeding.

Reaching a conclusion

If, after performing these tests, you are happy with both the structure of the BN and the CPTs associated with the intervention then you should conclude that the BN is an accurate reflection of reality. If you are unhappy, you should alter the BN (as directed above) until you are satisfied.

4. Repeat the procedure, above, for all the interventions that affect the management objective node. It is as important to consider the options that appear to have no effect as those that have the greatest effect. This may uncover errors in the BN but, more importantly, it will also reveal

weaknesses in the management plans that you are considering (see Example 11.7). Remember to update your list of interventions, and their effect on the desired state, whenever you make any changes to the BN.

5. Consider the effect of combinations of interventions on the management objective node. Obviously, implementing two interventions will have a greater effect than implementing either individually but, specifically, you want to look for groups of interventions whose combined impact is greater than the sum of their individual effects. This will occur where implementing one intervention will remove a constraint on another intervention working.

You should already have a good idea of where this is likely to happen from your investigation of the individual interventions, as described above. You may have found intermediate nodes with more than one parent, where changing the state of a second parent (which was not directly related to the intervention you were considering) improved the effect that the intervention had on the management objective. Where this was the case, check to see if any other interventions are related to this second parent and might augment the effect of the first intervention (see Example 11.8). If you remember the nodes that you identified as modifying parents in Step 5, then this will also indicate promising combinations of interventions.

6. Add these "intervention combinations" to your list only if their joint impact is greater than the sum of their individual effects. Once again, note the change they produce in the desired state of the management objective.
7. Examine the changes produced by each intervention (or intervention combination) over all the time steps represented in your BN. Delete from your list those interventions which do not have a lasting positive impact or which have adverse additional impacts which you consider to be unacceptable.
8. Repeat the entire procedure from Point 2 for each management objective.

By this stage, you should have a list of interventions (and intervention combinations) for each management objective, together with the improvement that each produces in that management objective. Compare these lists and highlight those interventions that affect more than one management objective, together with those that affect a single intervention to a large degree. Doing this will indicate those interventions which the BN suggests will do the most to achieve your management objectives.

Be particularly careful with interventions where more than one state of the intervention can be implemented at the same time. For example, you may have an intervention to plant either wheat or cotton. If your BN suggests that both of these options will have a similar impact on your management objectives, then there may be a good case for implementing them both (i.e. planting *both* wheat and cotton). This is particularly true with something like a cropping strategy as planting only a single crop can be highly risky and also tends to encourage disease. Using the BN in what is called "diagnostic mode" may help you to consider these situations (see Example 11.9).

As has been emphasised before, it is important that you don't just rely on the outputs from the BN. Instead, you should consider all these interventions in the light of the improved understanding about the environmental system that you should have developed through using the BN. In particular, you should be aware of how your confidence in the information you have entered into the CPTs affects the outputs from the BN (see Appendix 1 for more information). In addition, you should also consider any factors which affect the decision and have not been included in the BN (cost, perhaps or political considerations). On this basis, you should select the intervention (or interventions) you believe to be the most appropriate.

Examples for step 11

Example 11.1

Figure 4.15 shows a BN with a single objective and two interventions (this is the right hand network shown in Figure 4.12). Both interventions have been set to the states representing the current situation in the environmental system.

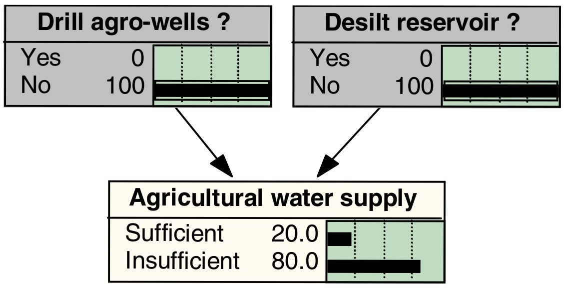
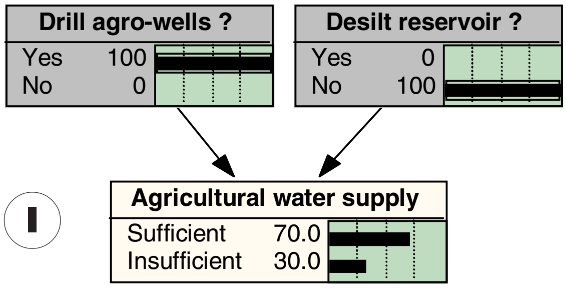


Figure 4.15

Figure 4.16 then shows how the probability that the objective (agricultural water supply) is in the state "Sufficient" changes as the states of the interventions are changed. It can be seen that the best result is produced when the intervention "Drill agrowells?" is implemented and that desilting the reservoir has no impact.



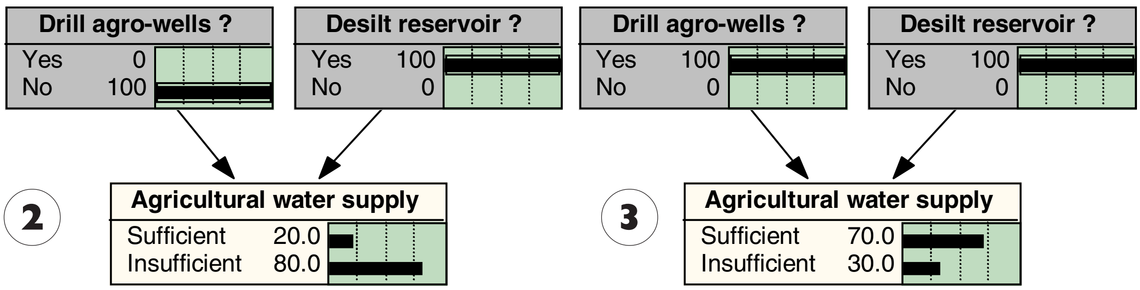


Figure 4.16

Example 11.2

Figure 4.17 shows a network with one intervention ("Build wells?") and one controlling factor ("Rainfall"). Obviously, it is hard to predict exactly whether the rainfall will be high or low but, based on past rainfall records, the probability that it will be high can be estimated at 30%.

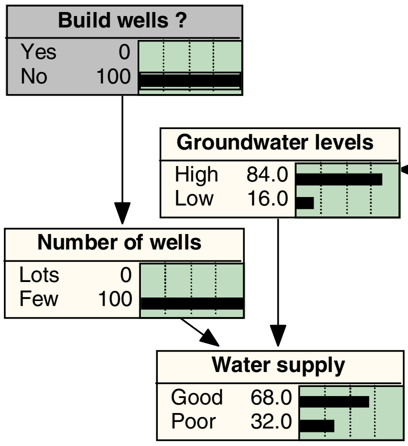


Figure 4.17

Example 11.3

For the BN shown in Figure 4.16, the following list can be compiled:

Objective: Agricultural water supply	**Change in probability that "Agricultural water supply" is "Sufficient"**
"Drill agrowells" is "Yes"	Up 50%
"Desilt reservoir" is "Yes"	0 %

Example 11.4

Consider the BN in Figure 4.18. The intervention "Subsidise inputs?" has a much more direct impact on the objective "Agricultural productivity" than the intervention "Soil conservation?" has. There is only one intermediate node between "Subsidise inputs?" and "Agricultural productivity" while there are three between "Soil conservation?" and "Agricultural productivity". You should ask yourself whether subsidising inputs will affect agricultural productivity more directly than soil conservation measures in reality.

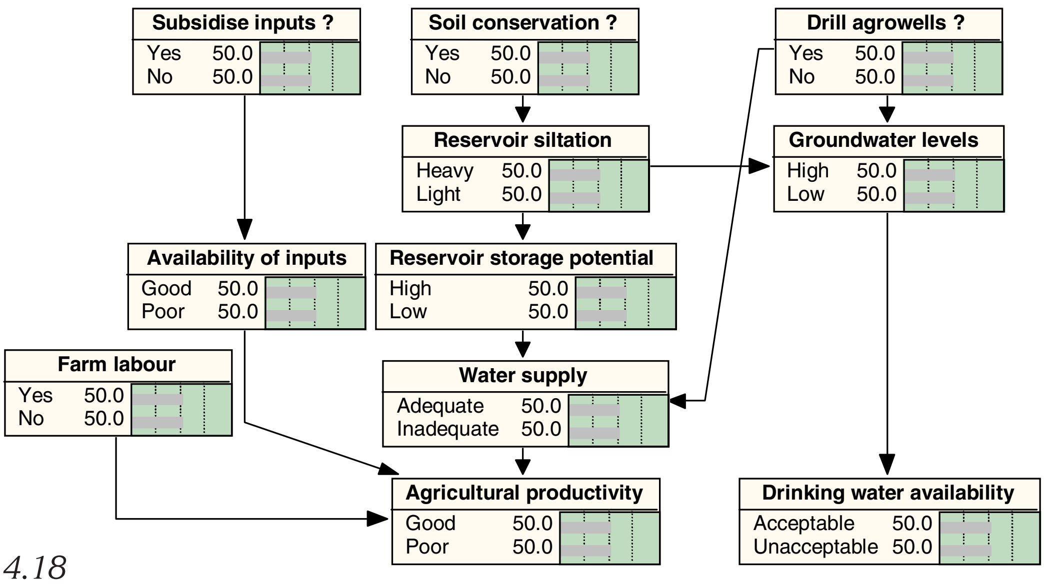


Figure 4.18
(repeated from Figure 4.7)

Example 11.5

Consider the BN in Figure 4.18. Among other things, this BN captures the way in which soil conservation affects agricultural productivity. Changing the value of the node "Soil conservation?" will produce changes in the value of the node "Agricultural productivity" based on the information entered into the CPTs. However, it is important to understand how these changes arise or, in other words, how the information in the CPTs leads to the changes observed. This will help you make sure that the BN properly represents what you think and will also help to develop your understanding of the environmental system.

To do this, start by changing the value of the node "Soil conservation?": first, set it to "Yes" and then set it to "No". Examine the effect this has on the node "Reservoir siltation" and check that you understand the changes you see. Next, change the value of "Reservoir siltation" directly and examine the affect this has on the node "Reservoir storage potential". Again, make sure you are clear as to why these changes are taking place.

The next step is slightly more complicated. You now want to make sure you understand the reasons why the node "Water supply" changes. As you can see from Figure 4.18, this node has an additional parent to the one you have just considered: "Drill agrowells?". Set "Drill agrowells?" to "Yes" and change the state of "Reservoir storage potential" from "High" to "Low" and back again, until you are sure you have understood the changes. Then set "Drill agrowells?" to "No" and change "Reservoir storage potential" again.

Understanding the changes in "Agricultural productivity" is more complicated still as it has three parents. The procedure is the same, however. Fix "Farm labour" to "Yes" and "Availability of inputs" to "Good" and then change "Water supply" between "Adequate" and "Inadequate". Then fix "Farm labour" to "No", keeping "Availability of inputs" as "Good" and examine the affect of changing "Water supply". Continue like this until you have examined all combinations of the states of "Farm labour", "Availability of inputs" and "Water supply".

Example 11.6

Table 1 shows the CPT for the node "Agricultural productivity" shown in Figure 4.18. Assume you have been carrying out the process described in Example 11.5 and have fixed "Farm labour" as "No" and "Availability of inputs" as "Good". When you start to change "Water supply" from "Adequate" to "Inadequate" you notice that the chance of good agricultural productivity is greater when the water supply is inadequate. Clearly, this makes no sense. To find out why this is happening you should look at the two rows highlighted in Table 1 which relate to the state combinations of concern. The probabilities shown here are the same as the ones you have seen for the node "Agricultural productivity" in the network itself. Examining them suggests that the most likely cause of the problem is that the probability for "Good" has been accidentally swapped with the probability for "Poor" in row 3. However, this should be checked with the source of the original data to make sure that this is a mistake and not a real reflection of how the environmental system works.

Table 1

Water supply	Farm labour	Availability of inputs	Agricultural productivity Good	Agricultural productivity Poor
Adequate	Yes	Good	0.90	0.10
Adequate	Yes	Poor	0.65	0.35
Adequate	***No***	***Good***	***0.25***	***0.75***
Adequate	No	Poor	0.40	0.60
Inadequate	Yes	Good	0.70	0.30
Inadequate	Yes	Poor	0.45	0.55
Inadequate	***No***	***Good***	***0.55***	***0.45***
Inadequate	No	Poor	0.20	0.80

Example 11.7

Figure 4.19 shows the same BN with the node "Awareness programme?" in two different states. You can see that the BN suggests that implementing an awareness programme will only reduce the chance of high soil erosion by a negligible amount. It is important to understand why this is the case.

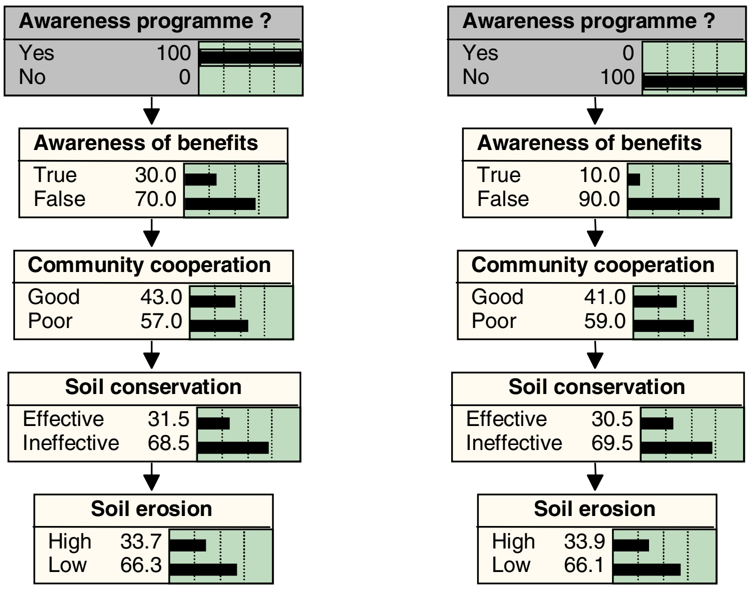


Figure 4.19

Figures 4.20 to 4.22 show the results of carrying out the procedure explained in Example 11.5. They are best examined in reverse order. Figure 4.22 shows a 20% difference in the chance that soil erosion will be high depending on the state of "Soil conservation". Figure 4.21 shows a 10% difference, depending on the state of "Community cooperation", and Figure 4.20 shows 1% difference, depending on the state of "Awareness of benefits". While you should expect nodes that are further away from the objective to have less effect, the major reduction in effect when "Awareness of benefits" is changed instead of "Community cooperation" suggests that implementing an awareness programme is probably not the best way of achieving a reduction in soil erosion. The BN suggests that this is because improving the awareness of the benefits has only a minimal impact on community cooperation. A better potential management strategy might look at alternative ways of promoting community cooperation.

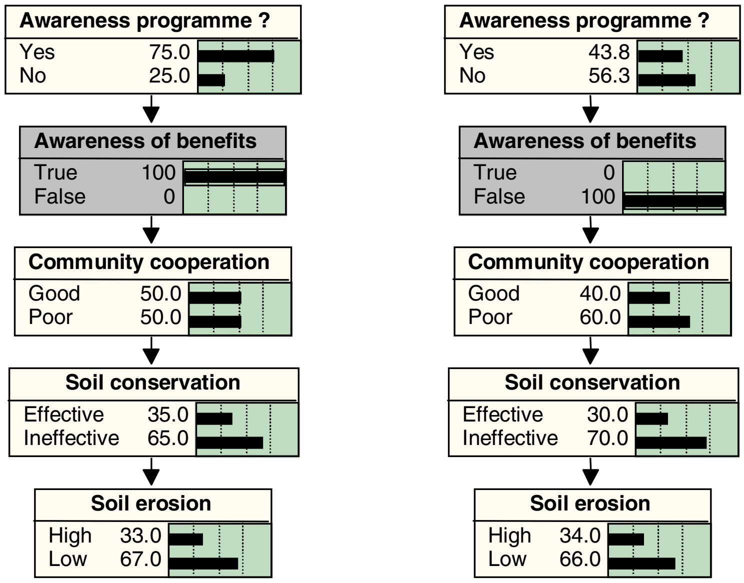


Figure 4.20

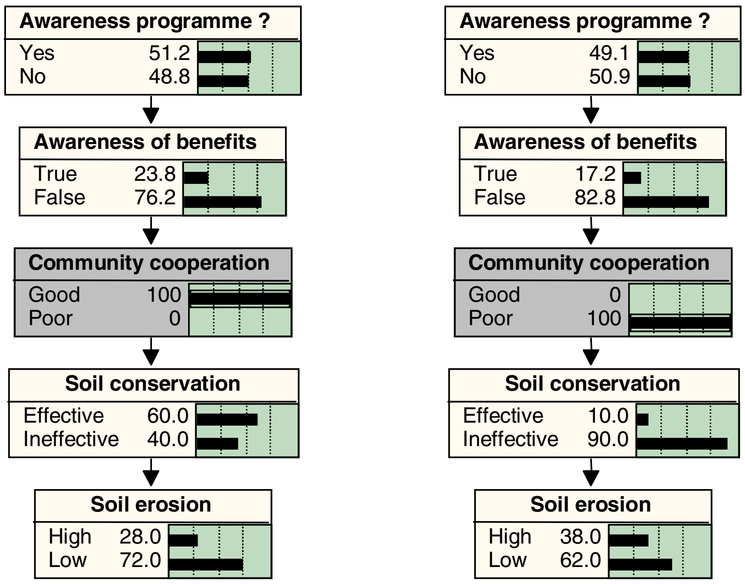


Figure 4.21

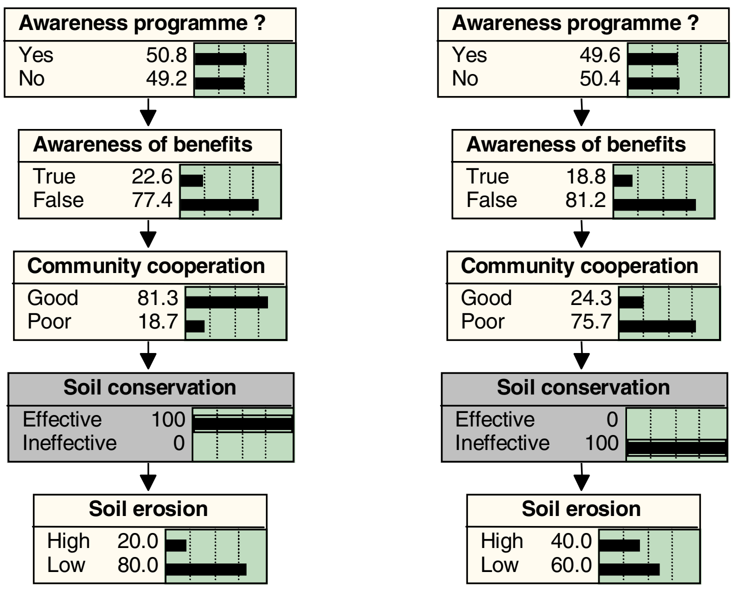


Figure 4.22

Example 11.8

Figure 4.23 shows the same BN with the node "Water application" changed from "at 50%" to "at 100%". It can be seen that when the soil is poor, the increase in water leads to only a small increase in the chance that higher yields will be obtained. Figure 4.24, however, shows that when the soil is good, the impact of improving the water supply is much greater. This suggests that an intervention aimed at improving the soil quality will boost the benefits gained from improving the water supply.

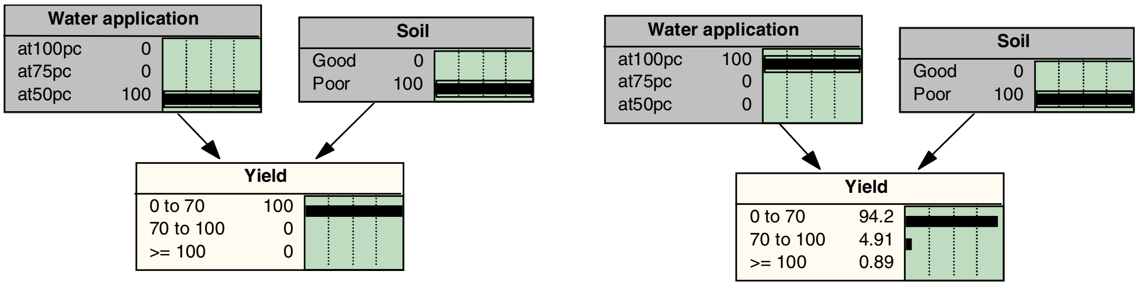


Figure 4.23

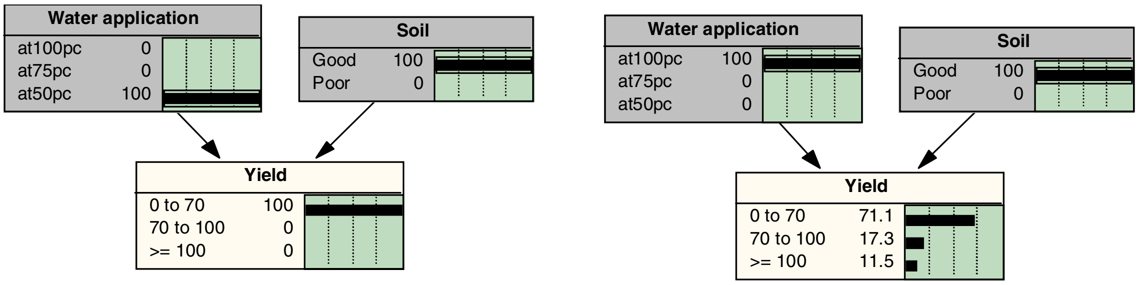


Figure 4.24

Example 11.9

Figure 4.25 shows the same BN repeated twice for two different states of the intervention "Crop". It can be seen that, although cotton has a better chance of providing a high yield, wheat is not too far behind. As both states of the intervention can be implemented, this suggests that it may be a good idea to plant both crops (perhaps with a small bias towards cotton) in order to minimise risk.

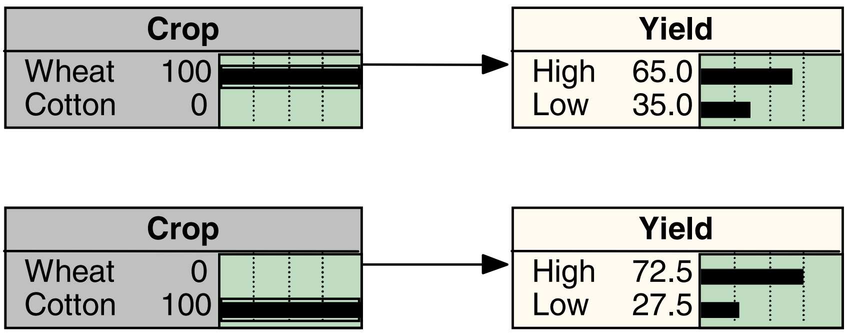


Figure 4.25

This example (and all the guidelines for Step 11) uses the BN in "predictive mode". In other words, the variables at the base of the arrows (the interventions) are changed to examine the impacts on the variables at the head of the arrows (the management objectives). However, it is sometimes useful to use a BN in "diagnostic mode" whereby the management objectives are changed and the impacts on the interventions are examined. When used like this, the BN shows the most probable configurations of the intervention nodes. **This should always be done with care, as the results can be hard to interpret.** This is illustrated below.

Figure 4.26 shows the same network as in Figure 4.25 but being used in diagnostic mode. The **incorrect** conclusion to reach from this result is that a "High" yield will be obtained by planting 47.3 % of the fields with wheat and 52.7% of the fields with cotton. This is not the case, as is shown by Figure 4.27 where the probabilities of "Crop" have been fixed at these values.

Figure 4.26

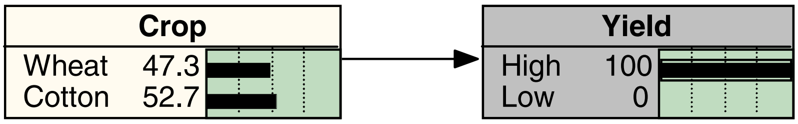
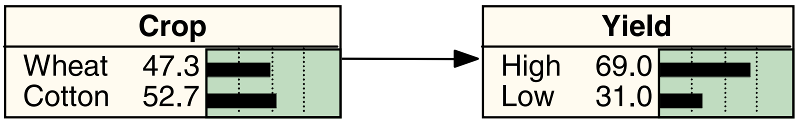


Figure 4.27



The **correct** conclusion is as follows: if you had planted half the fields with wheat and half the fields with cotton (in Figure 4.26, the probabilities in the CPT of "Crop" are 0.5 each) and then went back at harvest time to check on the yields, then, out of those fields which had produced a high yield, 47.3 % of them would be wheat and 52.7% of them would be cotton. The important point to note is that the result applies only to those fields with a high yield and that it is dependent on the ratio of wheat to cotton with which you planted the fields in the first place. If you had planted 90% of the fields with wheat (instead of 50%) then you would have got a different answer as there would obviously have been more chance of a high yielding field being a wheat field (simply because there were more of them).

In this case, the result is useful as it shows that there is little difference in the performance of cotton over wheat in terms of producing a high yield (but it only does this because the probabilities in the CPT of "Crop" were set to 50% each, as has been done in Figure 4.26, which, before any other information is entered into the BN, gives wheat and cotton an equal chance of achieving a high yield). This is, of course, the same conclusion as can be reached using the BN in predictive mode (as in Figure 4.25) but provides a different way of looking at it, which may help to improve your understanding of how the system is working.

Step 12: Hold a second joint stakeholder workshop to discuss your decisions

Having chosen the interventions that you think are best, it is important to go back to the stakeholders to tell them about the decisions you have made. In doing this, you should be prepared to change these decisions depending on the reaction of stakeholders to them. Therefore, it is important to stress that there is still an opportunity for discussion before the interventions are implemented and that changes are still possible.

Ideally, you want to obtain as much stakeholder support a possible for the interventions that you will implement. Of course, it is likely that there will be some stakeholders who disagree with the intervention you are proposing. If this happens, you should listen to why they disagree and be sure that you have considered their point of view in reaching the decision you have made. You should also explain how you arrived at your decision. To help with this, you may want to show them the master BN you used. However, this will only be appropriate if you are confident that all the stakeholders present will understand it. When this is not the case, you should prepare an explanation of the important points beforehand in a format that will be easily understood by all present.

If a significant proportion of the stakeholders object to the decision then it is likely that the interventions you are proposing will be difficult to implement. In this case, it is recommended that you select other interventions that have more stakeholder support.

Chapter 5:

Managing resources in the Poya Ganga — a hypothetical case study

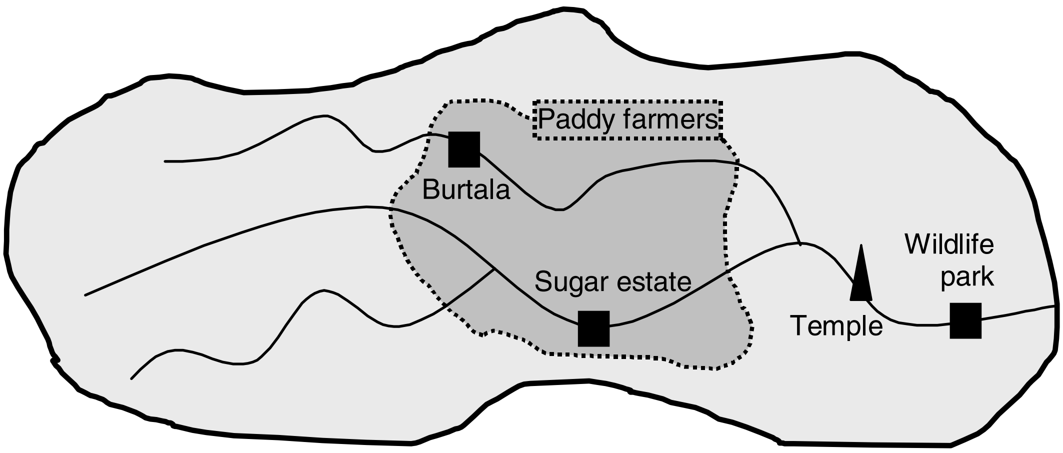
The scenario

The Poya Ganga River flows into the sea on the south-east coast of Sri Lanka. It is 100 km long and has a basin area of 2000 km^2. The mean annual rainfall shows a variation between 1200 mm and 2700 mm with a drought period from June until the middle of September.

There are five principal groups of water users in the basin:

- ✦ Agricultural water users – mostly paddy farmers.
- ✦ A large sugar estate
- ✦ Domestic water users – centred around the town of Burtala
- ✦ Devigama Temple
- ✦ A national wildlife park

The relative positions of the various water users along the river are shown below.



From the map, it can be seen that the paddy farmers draw their water from both main branches of the river while water use by the sugar estate is largely unaffected by that of Burtala (and vice-versa).

For a number of years now, the farmers and the sugar estate have been steadily increasing their water use in an effort to increase production. At the same time, local builders are digging sand from the riverbed to use in house construction (known as sand mining). This is decreasing the bed level at the rate of around 0.5 m per year and, consequently, reducing the area that can be irrigated by a gravity-fed system. Moreover, the population of Burtala has been growing. Furthermore, although actual water use by the temple itself is insignificant, a minimum flow is required in the river to permit ritual bathing.

You have been appointed by the Chairman of the District Water Resource Committee to make recommendations as to a future water management strategy in the basin. You have decided to use Bayesian networks to help you to formulate this strategy and you have obtained a copy of these Guidelines to help you.

Step 1: Be clear about what you want to use the BN for

After thinking about the situation, you list your management objectives as:

1. Maintain the minimum required flow through the temple (this will also satisfy the requirements of the wildlife park);
2. Meet minimum domestic requirements for the Burtala area;
3. After meeting these two objectives, maximise rice and sugar production.

At this stage, you also list a number of management interventions whose potential you would like to investigate:

1. Control of sand mining to prevent reduction in irrigable area;
2. Increasing surface storage to encourage more efficient use of water by the farmers;
3. Introduction of drip irrigation to sugar estate;
4. Encourage rainwater harvesting to augment domestic water supply;
5. A well rehabilitation and installation programme to increase groundwater use.

You realise that, if you are to develop an integrated plan for the basin as a whole, then you need to take the entire basin as the geographical area that you are considering. Moreover, you decide to concentrate on water use in the dry season, as this is the crucial time of year – during the rainy season there is far less pressure on resources. This does not quite define your time scale, however, as you are interested in the overall effects your proposed interventions (and any others) might have. You realise that few, if any, of your interventions will have an impact within a single dry season, so you decide to fix a planning period of three years.

Next you think about which people you need to involve in the decision making process. Although there are other water users in the basin, there are not many of them so you decide to focus on the main users only: paddy farmers, the sugar estate, domestic users, the temple and the wildlife park. These are the people you hope to affect by your management strategy. In addition to these groups, there are numerous government departments whose collaboration you will need if the plan is to succeed. These include the Department of Irrigation, the Water Supply and Drainage Board, the Department of Agrarian Services and the Department of Agriculture. Together with the user groups and the District Water Resource Committee, these groups constitute the stakeholders you will need to consult.

Step 2: Establish contacts with stakeholders

You approach each of your stakeholder groups in turn.

Paddy farmers

The department of Agrarian Services advises you that the farmers are a fairly homogeneous group, who have formed organisations to represent themselves, consisting of all the farmers who draw water from the same canal. You ask to consult with the chairmen from three of the farmers' organisations, one from the head of an irrigation scheme, one from the middle and one from the tail.

The sugar estate

The estate manager agrees to be involved in the consultation process.

Domestic users

The domestic users are quite a diverse group so, with the help of local staff from the Water Supply and Drainage Board, you ask three people to represent them. You loosely describe these people as a poor urban water user, a rich urban water user and a rural water user (urban water users are connected to a piped supply, rural users are not). On the advice of the Water Supply and Drainage Board, all the people asked are women, as they are the primary users of domestic water.

The temple and the wildlife park

As the needs of these users are straightforward and incorporated into the management objectives you defined in Step 1, you decide that they do not need to be directly involved in the consultation process.

Government departments

After approving your provisional list of management interventions, the relevant government departments decide that they do not need to be directly involved in the consultation process either. They ask, however, that you keep them informed of progress and remind you that any interventions would require their agreement as well as funding.

Step 3: Initial stakeholder group consultations

Initial consultation with domestic users

You meet the three representatives of the domestic users together and begin by explaining to them that you are developing a water management strategy for the basin and wish to know their opinions on the issues it should be addressing. Next you tell them your management objectives, explaining how increasing water use by paddy farmers may affect domestic supplies and how everybody's water use affects the river flow at the temple and through the wildlife park. You also point out to them that if use increases much further, the river will stop running during the dry season before it reaches the temple.

The stakeholders agree that all of these management objectives are important: they visit the temple for major festivals and their families are either dependent on rice production to make a living or work on the sugar estate. They point out, however, that domestic requirements are rarely met all year round: in rural areas wells dry up, while the piped supply to the town often becomes contaminated. You ask why each of these problems occurs. In the case of the wells, the women believe it is because they are not deep enough, while in the case of the town supply it is because people and animals defecate near to the river off-take. Not unreasonably, they think your management plan should stop both of these things happening.

You take this as your cue to ask them what they think should be done to stop these things, in practical terms. They reply that the people who originally installed the wells (a foreign-based NGO) should come back and drill them deeper. For the town supply, they think that fencing off the area round the river off-take would help, but that people, particularly children, also need to be educated about the consequences of contamination.

Next you show them your list of potential management interventions and ask their opinion about them. They immediately notice the plan to encourage rainwater harvesting and start to laugh. When you ask why they are laughing, they explain that this has been tried before without any success. Apparently, the problem is that the water collected off house roofs is stored in a large tank, but that this is often left uncovered and consequently it becomes dirty as well as providing a breeding ground for mosquitoes. Most people pulled their water collection systems down within a month of it being installed! When you ask whether a different system design might help, they look sceptical but say that they would give it a go.

They have little to say about the other interventions with the exception of the well rehabilitation programme (which they agree with) and the sand mining. Nobody seems to like the sand miners as they are blamed for causing riverbank collapse, which is threatening some farmers' land. The women point out, however, that if the sand miners were not sand mining they would have to earn a living in some other way, or face unemployment.

Initial consultation with paddy farmers

You begin your discussion with the paddy farmers in much the same way as you did with the domestic users. Again they agree with all your management objectives, although they point out that they should have priority over the sugar estate, as more people are dependent on rice production. You ask them if they think there should be any other objectives, but they cannot think of any. You go on to ask them to describe how rice production might be maximised. They explain that there are two main problems: the water supply and the availability of crop inputs. They describe these in detail.

The currently poor water supply arises from a combination of a number of factors. Firstly, the river off-takes to the irrigation scheme, controlled by the Department of Irrigation, are in a very poor state of repair. Secondly, farmers who are illegally settled on land by the river off-take extract water from the canal, to which they are not entitled. Most important, however, is the fact that there are very few reservoirs available, so excess water in the rainy season cannot be kept for later use.

The crop inputs the farmers are concerned about are mainly seed and fertiliser. Although farmers usually have the money to buy them, they are rarely available and, when they are available, they are of poor quality.

You then ask them how they think their problems should be solved. There are quite a few ideas: Transferring control of the river off-take to the farmers; enforcement of legislation prohibiting farming near the off-take; construction of reservoirs; better provision of seeds and fertiliser.

On presenting your potential management interventions to them, they are reminded of how much they dislike sand miners. Steps to control sand mining would clearly be popular, at least with the farmers. They fully agree with your plan to build new reservoirs and increase groundwater use and repeat the domestic users' reservations about rainwater harvesting.

Initial consultation with the sugar estate

On meeting the estate manager you, once again, explain what you are doing. The estate manager looks suspicious and tells you that you cannot reduce the water supply to the estate as too many people's jobs rely on it. He makes this point very forcibly. You suspect that he thinks that a decision has already been made and that your visit is just for show, so you try to reassure him that this is not the case. You point out that there *is* a potential problem with water resources in the basin and that if something isn't done then everyone risks losing out. Fortunately, he appears to give you the benefit of the doubt.

He agrees with your management objectives but argues that the estate is more important than the paddy farmers, as more people are dependent on it. This is directly contradictory to farmers' claims so you make a mental note to check up on it. He argues further that if the estate continues to be successful, then it will create more jobs and less people will have to rely on farming to make a living. Again, you make a mental note to check with the farmers whether this would be desirable outcome, if it were to become true. He has no further management objectives to add to your list.

When you ask him how these management objectives can be met, he immediately responds that the factory would be able to use water more efficiently if it could increase its storage capacity. When you question him further, he says that sufficient storage capacity may be achieved by renovating the reservoir that they already have. He is not certain about this, though, and would like the option to build a further reservoir. He thinks that they would probably have done this by now apart from the resistance of local politicians concerned about the temple supply.

He also explains that the estate uses water for three separate purposes. Firstly, it is used to irrigate the nursery sugar cane. Secondly, it is used in the process that extracts the sugar from the cane, and thirdly, it is used to supply the arrack distillery run by the estate. Although it is already done to a certain degree, he believes that with the necessary investment, there may be greater opportunities to recycle water between these three separate uses.

When you show him your list of possible interventions, he naturally picks out your idea about drip irrigation. He says that they have considered this, but have yet to adopt the new technology because of the investment in equipment and training it would require. He does agree, however, that use of drip irrigation probably would improve water use efficiency. He also wonders whether there would be enough water to fill all of the new reservoirs you are proposing to build, even in the wet season. You agree that is something that needs to be checked and make a mental note to do so.

Step 4: Construct preliminary BNs

Preliminary BN representing perspective of domestic users

You review your notes on your discussion with the domestic users and listen to the tape recording you made. You come up with the following variables, which you think capture what the stakeholders were saying to you:

Variable	Captures what?
Temple river flow	Management objective number 1 (agreed by this stakeholder group)
Rural water supply	Management objective number 2 (agreed by this stakeholder group) with reference to rural water users
Urban water supply	Management objective number 2 (agreed by this stakeholder group) with reference to urban water users
Rice yield	Management objective number 3 (agreed by this stakeholder group)
Sugar estate productivity	Management objective number 3 (agreed by this stakeholder group)
Faecal pollution	Represents defecation near river off-take
Well depth	Captures whether the wells are deep enough to yield in the dry season
Well deepening	An intervention to deepen wells, as suggested by the stakeholders
New well drilling	Intervention number 5 (agreed by this stakeholder group)
Well rehabilitation	Intervention number 5 (agreed by this stakeholder group)
Rainwater harvesting	Intervention number 4 (agreed by this stakeholder group)
Health education	An intervention to prevent faecal pollution, as suggested by stakeholders
Fence off-take	An intervention to prevent faecal pollution, as suggested by stakeholders
Drip irrigation	Intervention number 3 (agreed by this stakeholder group)
New reservoirs	Intervention number 2 (agreed by this stakeholder group)
Surface storage	Captures the impact new reservoirs would have on overall storage
Sand mining control	Intervention number 1 (agreed by this stakeholder group)
Irrigable area	Captures the impact of sand mining control
Unemployment	Captures the impact of sand mining control
River extraction (3 types)	Captures how much is being taken out of the river and the consequent downstream effect

You then categorise these variable according to the general network structure, as follows:

Category	Variables
Management objectives	Temple river flow, Rural water supply, Urban water supply, Rice yield, Sugar estate productivity
Interventions	Well deepening, New well drilling, Well rehabilitation, Rainwater harvesting, Health education, Fence off-take, Drip irrigation, New reservoirs, Sand mining control
Intermediate factors	Faecal pollution, Well depth, Surface storage, Irrigable area, River extraction (three types)
Controlling factors	None
Implementation factors	None
Additional impacts	Unemployment

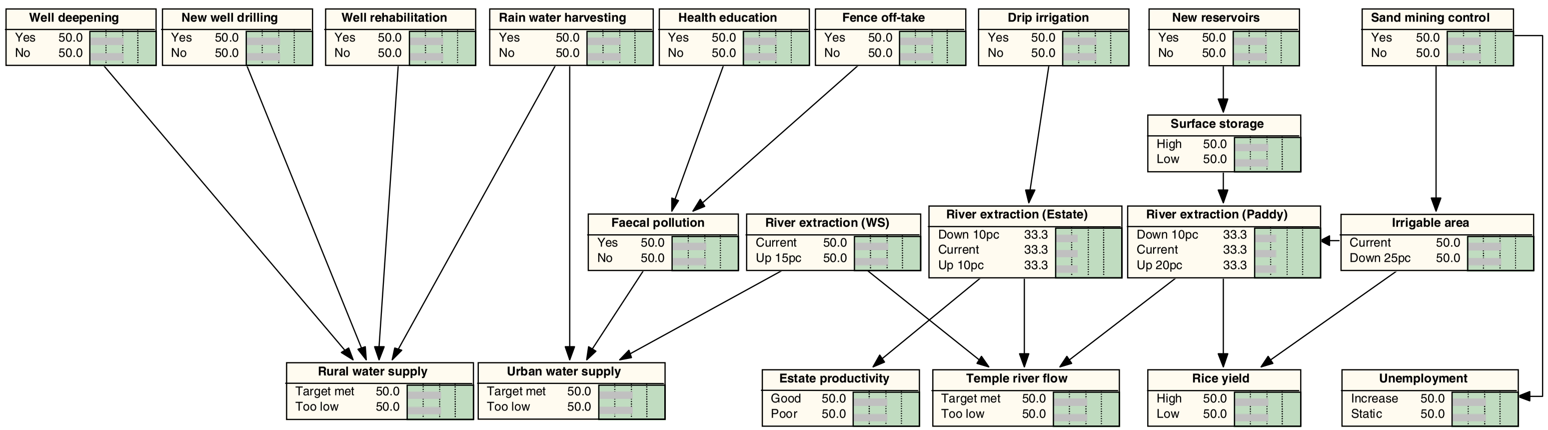
You note that you now have three new interventions to consider.

Based on these lists, you construct a BN diagram, as shown, choosing provisional states that seem appropriate to the logic expressed by the stakeholder group. Most of them are fairly easy — interventions are either implemented (Yes) or not (No) while some of the management objectives have targets (the minimum required flow at the temple, for example). You give the other management objectives and some intermediate nodes, such as Surface storage, qualitative state names, confident that the stakeholders will be able to quantify them the next time you meet.

However, a number of other intermediate nodes are slightly trickier. You realise that river extractions are the key variables to investigate in drawing up your management plan. Balancing these will be the key to meeting most of your management objectives, so you need to be able to examine what will happen if extraction goes up as well as down. You decide to specify likely changes given the interventions suggested (Up 10% and down 10%, for example) but note that you will need to check these with the stakeholders later and also verify them using more objective means. Remembering that it is important to minimise the number states, you only give two states to River extraction (WS) as you don't think the demand for the domestic water supply will ever decrease.

On performing the recommended checks, you notice that, although it has no parents in your BN diagram, you have classed River extraction (WS) as an intermediate node. You decide that this should, instead, be a controlling factor, as the states you have given it reflect an increase in population. While such an increase will obviously have a major impact on whether urban water supply and the temple flow is sufficient, you have no direct control over it.

You also notice that a single node called Well depth cannot sensibly describe the several wells in the Poya Ganga river basin, all of which will have different depths, and it is obviously not possible to have a node for each well in the area. You consider redefining it so that it does represent the wider area, but then decide that it is probably not needed at all, so you delete it.



Preliminary BN for domestic users

Preliminary BN representing perspective of paddy farmers

You note that, as the paddy farmers also agreed with your management objectives and interventions, many of the variables in this BN will be the same as those in the domestic users BN. For this stakeholder group, your list is as follows:

Variable	Captures what ?
Temple river flow	Management objective number 1 (agreed by this stakeholder group)
Rural water supply reference to	Management objective number 2 (agreed by this stakeholder group) with rural water users
Urban water supply reference to	Management objective number 2 (agreed by this stakeholder group) with urban water users
Rice yield	Management objective number 3 (agreed by this stakeholder group)
Sugar estate productivity	Management objective number 3 (agreed by this stakeholder group)
New well drilling	Intervention number 5 (agreed by this stakeholder group)
Well rehabilitation	Intervention number 5 (agreed by this stakeholder group)
Rainwater harvesting	Intervention number 4 (agreed by this stakeholder group)
Drip irrigation	Intervention number 3 (agreed by this stakeholder group)
New reservoirs	Intervention number 2 (agreed by this stakeholder group)
Surface storage	Captures the impact new reservoirs would have on overall storage
Sand mining control	Intervention number 1 (agreed by this stakeholder group)
Irrigable area	Captures the impact of sand mining control
Input availability	Captures the availability of inputs
Input quality	Captures the quality of inputs
Illegal extractions	Captures extractions by farmers illegally settled on the land near the river off-take
Law enforcement	An intervention to prevent illegal extractions, as suggested by the stakeholders
Off-take condition	Captures the poor state of repair of the river off-take
Farmer control	An intervention aimed at handing control of the river off-take over to the farmers, as suggested by the stakeholders
River extraction (3 types)	Captures how much is being taken out the river and the consequent downstream effect

You then categorise these variables as:

Category	Variables
Management objectives	Temple river flow, Rural water supply, Urban water supply, Rice yield, Sugar estate productivity
Interventions	New well drilling, Well rehabilitation, Rainwater harvesting, Farmer control, Drip irrigation, New reservoirs, Sand mining control
Intermediate factors	Surface storage, Irrigable area, River extraction (Estate and Paddy), Law enforcement, Illegal extractions, Off-take condition, Input availability, Input quality, Inputs
Controlling factors	River extraction (WS)
Implementation factors	None
Additional impacts	None

You decide Law enforcement is an intermediate factor and not an intervention, as its implementation is not straightforward. You make a note to ask the farmers what interventions they think will lead to law enforcement. You decide you also need to ask them for specific interventions that will improve input availability and input quality. You also note that this stakeholder group has currently suggested one extra intervention.

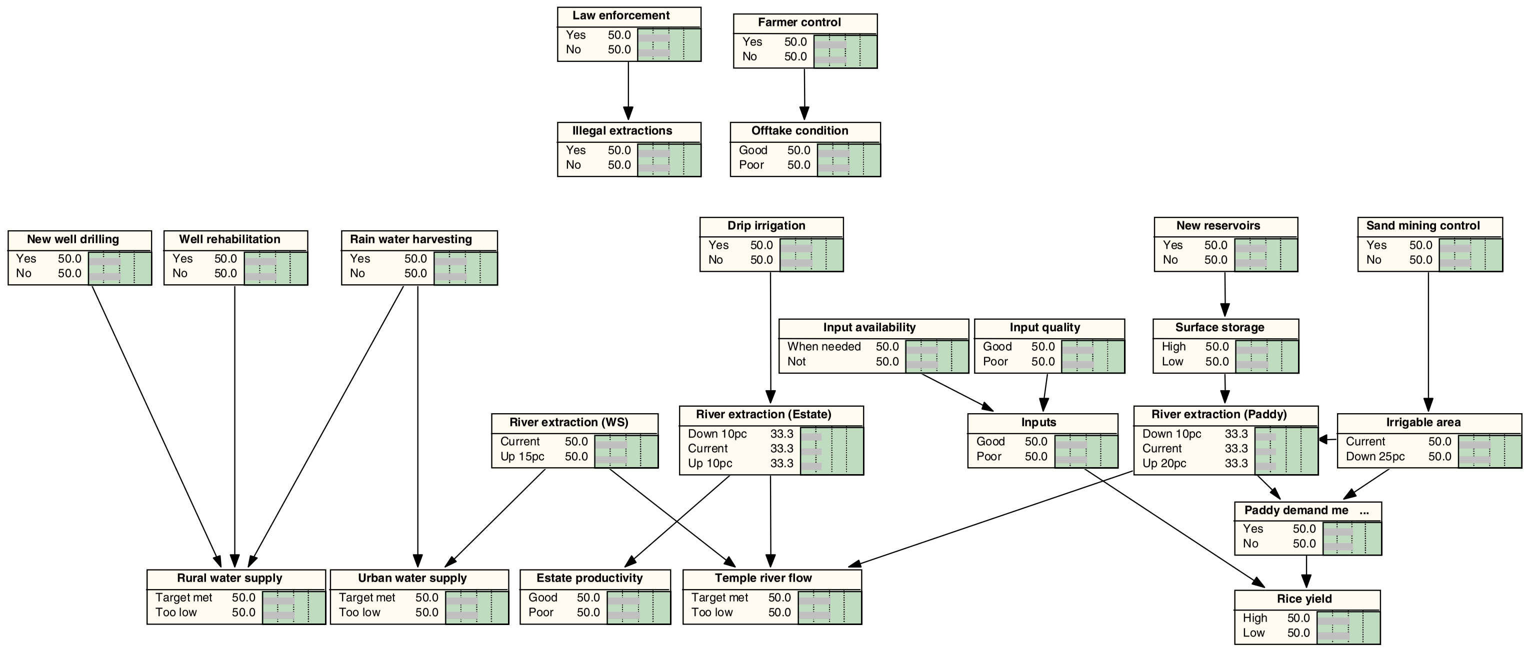
Based on these lists you construct a BN diagram. While doing so, you note that you can group together Input availability and Input quality and make them parents of a single variable called Inputs. You also think you can group River extraction (Paddy) and Irrigable area as parents of a single variable called "Paddy demand met?" This is because comparing the water extracted from the river to the irrigable area will determine whether enough water is available to grow rice on the whole area.

As you work on the BN diagram, you become aware that your questioning did not reveal all the information you needed. You note two further questions that you will need to ask the stakeholder in order to complete the diagram:

1. What effect does the condition of the river off-take have on water extraction? Does it reduce the amount of water that can be taken from the river or does it cause a portion of the water extracted to be lost before it reaches the irrigation canal?
2. What is affected by the illegal extractions? Do they affect the amount of water extracted or the proportion of extracted water that reaches the farmers?

Without the answers to these questions, you are not sure how to connect some of the variables to the network. Consequently, you leave them "floating" with the intention of connecting them to the BN diagram once the farmers have answered your questions.

This time, there are no new problems with state definitions. Also, with the exception of the "floating" variables, the double-checks confirm that you are happy with the logic the BN expresses. When you have done this, you suddenly remember that the farmers were very keen for you to understand that rice yield was a more important management objective than the productivity of the sugar estate. You wonder how you can include this in your BN diagram. In the end, you decide that you don't need to, as this is a value judgment that is implicit to you as the decision maker. If it is not possible to meet both management objectives, you will have to decide which is the most important and select interventions that promote improvement in your favoured management objective.



Preliminary BN for paddy farmers

Preliminary BN representing perspective of estate manager

You note that, as the estate manager also agreed with your management objectives and interventions, many of the variables in this BN will also be the same as in the others. For the estate, therefore, your list is as follows:

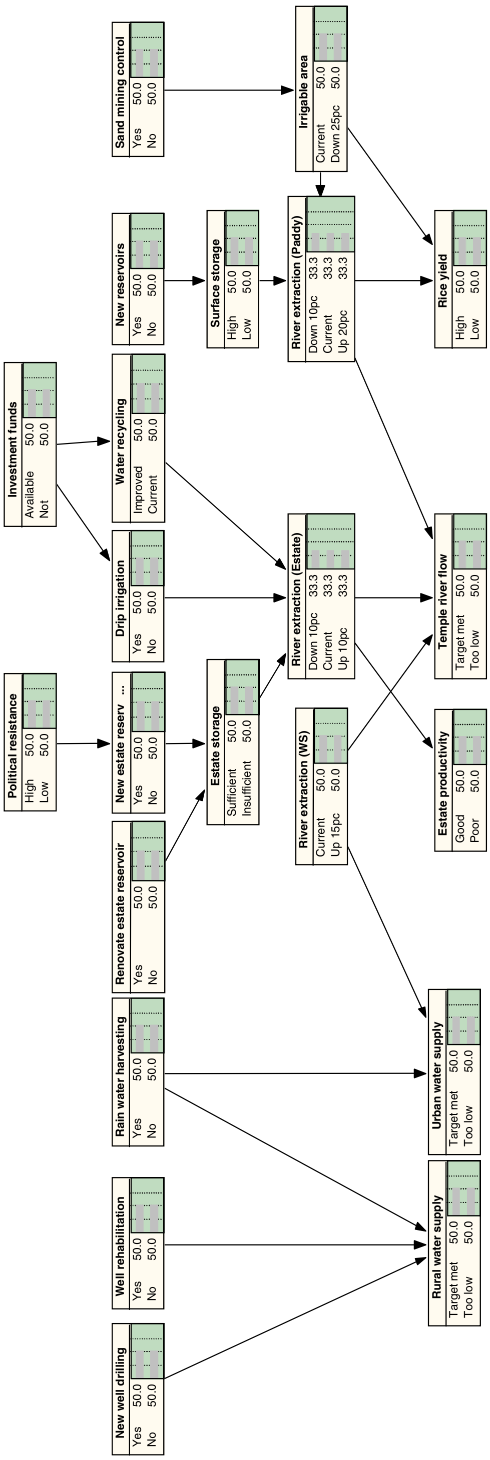
Variable	Captures what?
Temple river flow	Management objective number 1 (agreed by this stakeholder group)
Rural water supply reference to	Management objective number 2 (agreed by this stakeholder group) with rural water users
Urban water supply reference to	Management objective number 2 (agreed by this stakeholder group) with urban water users
Rice yield	Management objective number 3 (agreed by this stakeholder group)
Sugar estate productivity	Management objective number 3 (agreed by this stakeholder group)
New well drilling	Intervention number 5 (agreed by this stakeholder group)
Well rehabilitation	Intervention number 5 (agreed by this stakeholder group)
Rainwater harvesting	Intervention number 4 (agreed by this stakeholder group)
Drip irrigation	Intervention number 3 (agreed by this stakeholder group)
New reservoirs	Intervention number 2 (agreed by this stakeholder group)
New estate reservoir	An intervention to increase estate reservoir storage, as suggested by the stakeholders
Renovate estate reservoir	An intervention to increase estate reservoir storage, as suggested by the stakeholders
Estate storage	To capture the water storage available to the estate
Political resistance	To capture the political resistance that prevent the construction of a new reservoir
Water recycling	An intervention capturing a possible improvement if water recycling on the sugar estate, as suggested by the stakeholders
Investment funds	Captures the necessary conditions for two of the interventions
Surface storage	Captures the impact new reservoirs would have on overall storage
Sand mining control	Intervention number 1 (agreed by this stakeholder group)
Irrigable area	Captures the impact of sand mining control
River extraction (3 types)	Captures how much is being taken out the river and the consequent downstream effect

You then categorise the variables as:

Category	Variables
Management objectives	Temple river flow, Rural water supply, Urban water supply, Rice yield, Sugar estate productivity
Interventions	New well drilling, Well rehabilitation, Rainwater harvesting, Renovate estate reservoir, New estate reservoir, Drip irrigation, Water recycling, New reservoirs, Sand mining control
Intermediate factors	Estate storage, Surface storage, Irrigable area, River extraction (Estate and Paddy)
Controlling factors	River extraction (WS)
Implementation factors	Investment funds, Political resistance
Additional impacts	None

You note that you now have another three new interventions.

Constructing the BN presents no new problems and giving preliminary definitions to the states is straightforward. The double-checks reveal no problem with the logic. However, you do wonder whether resistance from local politicians is relevant to any changes other than the construction of a new reservoir on the sugar estate. You also realise that new storage reservoirs should only be built if they can be filled.



Preliminary BN for sugar estate

Step 5: Further stakeholder group consultations

Further consultation with domestic users

You decide that the domestic users will not find it easy to understand the logic represented by the BN diagram so you draw up a series of lists to allow them to validate the BN structure, as shown below:

List 1

Rural water supply
New well drilling
Well rehabilitation
Rainwater harvesting

List 2

Urban water supply
Rainwater harvesting
Faecal pollution
River extraction (WS)

List 3

Estate productivity
River extraction (Estate)

List 4

Temple river flow
River extraction (WS)
River extraction (Estate)
River extraction (Paddy)

List 5

Rice Yield
River extraction (Paddy)
Irrigable area

List 6

Unemployment
Sand mining control

List 7

Faecal pollution
Health education
Fence off-take

List 8

River extraction (Estate)
Drip irrigation

List 9

River extraction (Paddy)
Surface storage

List 10

Irrigable area
Sand mining control

List 11

Surface storage
New reservoirs

You begin by explaining that the lists should only include the information discussed at your previous meeting. You tell them that if there is anything they don't understand or don’t agree with they must say so. You then show list 1 to the group and make sure they understand what all the names mean. You explain that the things listed under the Rural water supply heading are meant to be the things which affect the rural water supply. First, you ask them if this is true and then you ask them if there is anything else that they would like to add to the list. You remind them that you are only really interested in things that are quite important. The group agrees with list 1 and doesn’t want to add anything so you move on to the next list.

Much to your surprise, the group agrees with all the lists; so, happy that the logic of the BN diagram is a fair representation of the group’s logic, you move on to talking about the states. To save time, you decide you do not need to worry about the states of Unemployment and all the nodes with Yes/No states, since these are obvious, so you begin with the quantitative states you defined in the previous step. You start with the node Irrigable area and explain that the state “Down 25%” is meant to be a rough guess at how farm land would be lost if sand mining were to continue in its present rate for the next three years. They are not really sure, but they think back over the last three years and seem to agree that your estimate is not too far from the truth. You have similar discussions for all three River extraction variables but they are even less sure about these. However, you decide that you can move on, as you will be able to get better information from the farmers and the estate manager.

For the same reasons, you decide to skip over the Estate productivity, Temple river flow and Rice yield nodes and move onto the Urban water supply node. You explain that "Target met" is meant to refer to the management objective of meeting minimum domestic requirements. You then ask them what they think those minimum requirements are. The two women who are connected to the urban supply think through a typical day and guess that they use between 20 and 30 buckets a day (you know that buckets have a capacity of about 10 litres) and they seem quite happy with this. Moving onto the Rural water supply node, the woman who relies on the rural supply thinks she uses a bit less than the urban people — somewhere between 10 and 20 buckets a day — but she says that she would use more if it didn't take so long to get it from the well.

Finally, you review the BN with them by looking at all the nodes that have two or more children. You begin with Rural water supply and ask the following questions, one at a time:

1. Imagine how a programme of well deepening will help meet the rural water supply target. Will the impact of this be altered if a drilling and rehabilitation programme is carried out at the same time? Will the impact be altered if rainwater harvesting is improved?
2. Imagine how a programme of new well drilling will help meet the rural water supply target. Will the impact of this be altered if a deepening and rehabilitation programme is carried out at the same time? Will the impact be altered if rainwater harvesting is improved?
3. Imagine how a programme of well rehabilitation will help meet the rural water supply target. Will the impact of this be altered if a drilling and deepening programme is carried out at the same time? Will the impact be altered if rainwater harvesting is improved?
4. Imagine an increase in rainwater harvesting will help meet the rural water supply target. Will the impact of this be altered if a well deepening and rehabilitation programme is carried out, or if new wells are drilled?

The group is confident that the answer to question 4 is no but they have to think about the other questions carefully. In the end, they decide that well deepening will change the effect that well rehabilitation has on rural water supply and vice-versa. They explain that this is because the two interventions together have a greater impact than the sum of each of their separate effects – well rehabilitation will increase the amount of water drawn from a well but will increase it even more if the well is deep enough so that it always has water in it. You accept their explanation and note Well deepening as a modifying parent.

You repeat this procedure for each of the following nodes: Urban water supply, Temple river flow, Rice yield and Faecal pollution. You don't identify any further modifying parents.

Further consultation with paddy farmers

Before beginning a general discussion with the farmers, you ask the farmers the questions that you came up with during Steps 3 and 4:

1. How important is it to your family that there are good opportunities for employment at the sugar estate?

The farmers reply that few families rely on farming alone and that many are dependent on wages earned on the sugar estate.

2. What is the effect that the condition of the river off-take has on water extraction? Does it reduce the amount of water that can be taken from the river or does it cause a portion of the water extracted to be lost before it reaches the irrigation canal?

The farmers reply that it reduces the width of the off-take opening onto the river and, since the Irrigation Department only allows them to leave it open for a fixed length of time, this reduces the amount of water they can extract.

3. What is the effect of the illegal extraction? Does it affect the amount of water extracted or does it affect the proportion of extracted water that reaches the farmers?

The farmers reply that it is both. The illegal extractions cause bank collapse, which further reduces the width of the off-take, but clearly not all of the water extracted reaches the farmers, as some is used on the illegally occupied land.

You also ask the farmers to suggest specific interventions to improve law enforcement, input availability and input quality. They suggest the following:

- Law enforcement is just a case of the authorities being prepared to do something about it. The illegally settled farmers who are currently there should be forcibly removed and the same should happen to any new settlers. Your stakeholder group is unanimous that the legal farmers would be quite happy to report offenders to the authorities as long as they were confident that something would be done about it.
- Currently inputs are largely provided by the government. The farmers think that if the government would allow a farmer co-operative to supply inputs then both availability and quality would improve.

You decide that farmers will also not find it easy to understand the BN diagram directly, so you draw up lists similar to those you used for the domestic stakeholders. Happily, the farmers approve your logic. You also discuss states in a similar way. In particular, you focus on the variables related to rice production. The farmers are happy with your qualitative descriptions of these states but, encouraged by you, give them quantitative definitions as follows:

Off-take condition	**Good**:	The entrance to the off-take is fully open
	Bad:	The entrance to the off-take is partially obstructed
Input availability	**When needed**:	Inputs are available whenever they are needed
	Not:	Inputs are sometimes unavailable when needed
Input quality	**Good**:	Quality of inputs satisfies farmers
	Poor:	Quality of inputs does not satisfy farmers
Inputs	**Good**:	Inputs are of good quality and available
	Poor:	Inputs are either unavailable or of poor quality or both
Surface storage	**High**:	Surface storage is 30% greater than the current level
	Low:	Surface storage is less than 30% greater than the current level
Rice yield	**High**:	More than 4 tonnes/hectare
	Low:	Less than 4 tonnes/hectare

They think that your estimate of a 20% increase in use is a little high, assuming that the off-take is fixed and that irrigable area doesn't increase by much (currently, most land that can be irrigated is being irrigated). They suggest a 10% increase might be more realistic. Conversely, they think your suggested 10% decrease is a little low. If storage were improved, illegal extractions were stopped and control of the off-take were handed over to the farmers, they think water use might decrease by as much as 20%.

Having agreed the states, you review the network to check for modifying parents. As the group doesn't find any, you thank them and close the meeting. Afterwards, you re-draw the network to incorporate the changes suggested during this second consultation.

Further consultation with the sugar estate

As he has a high level of education, you decide that the sugar estate manager will understand the logic of the BN diagram so you show it to him directly. As you expected, he has no problems understanding it, and agrees that the logic it represents captures what he believes. You start to discuss states and he provides the following definitions:

Political resistance	**High**:	Opposition from local politicians is sufficient to affect the productivity of the estate
	Low:	There is no opposition from local politicians, or it is not strong enough to affect the productivity of the estate
Water recycling	**Improved**:	Use of recycled water increases by 20%
	Current:	Use of recycled water remains at current levels
Estate storage	**Sufficient**:	Storage is increased by 50%
	Insufficient:	Storage is increased by less than this

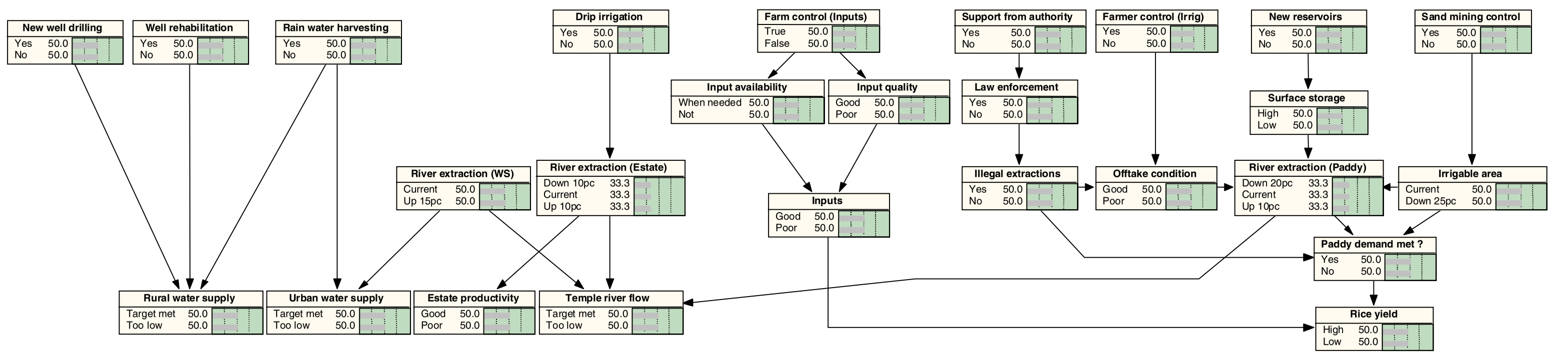
He thinks that the states you have chosen for River extraction (Estate) are not bad. He agrees that if nothing changes, they will start to extract around 10% more from the river as they intend to increase production at the distillery. He thinks, however that reductions greater than 10% can be achieved, particularly if water recycling can be improved. You agree to change this state to "Down 20%".

Finally, you review the network for modifying parents but don't find any.

Step 6: Draw conclusions from stakeholder consultation

Conclusions from consultation with domestic users

The variables have changed slightly since you originally categorised them in Step 4, so you re-draw your table to reflect this:



Adapted BN for paddy farmers

Category	Variables
Management objectives	Temple river flow, Rural water supply, Urban water supply, Rice yield, Sugar estate productivity
Interventions	Well deepening, New well drilling, Well rehabilitation, Rainwater harvesting, Health education, Fence off-take, Drip irrigation, New reservoirs, Sand mining control
Intermediate factors	Faecal pollution, Surface storage, Irrigable area, River extraction (Estate and Paddy)
Controlling factors	River extraction (WS)
Implementation factors	None
Additional impacts	Unemployment

You then draw up lists of the management objectives and the interventions that affect them:

<u>Rural water supply</u>
Well deepening
New well drilling
Well rehabilitation
Rain water harvesting

<u>Urban water supply</u>
Rainwater harvesting
Health education
Fence off-take

<u>Estate productivity</u>
Drip irrigation

<u>Temple river flow</u>
Drip irrigation
New reservoirs
Sand mining control

<u>Rice yield</u>
New reservoirs
Sand mining control

Under the Rice yield management objective, you highlight sand mining control, as it affects the management objective in more than one way. You also note that sand mining control affects the Temple river flow management objective.

Conclusions from consultation with paddy farmers

You draw up a similar table and lists for the paddy farmers:

Category	Variables
Management objectives	Temple river flow, Rural water supply, Urban water supply, Rice yield, Sugar estate productivity
Interventions	New well drilling, Well rehabilitation, Rainwater harvesting, Farmer control (Inputs), Farmer control (Irrig), Drip irrigation, New reservoirs, Sand mining control
Intermediate factors	Surface storage, Irrigable area, River extraction (Estate and Paddy), Law enforcement, Illegal extractions, Off-take condition, Input availability, Input quality, Inputs
Controlling factors	River extraction (WS), Support from authority
Implementation factors	None
Additional impacts	None

Rural water supply
New well drilling
Well rehabilitation
Rainwater harvesting

Urban water supply
Rainwater harvesting

Estate productivity
Drip irrigation

Temple river flow
New reservoirs
Drip irrigation
Sand mining control
Farmer control (Irrig)

Rice yield
New reservoirs
Sand mining control
Farmer control (Inputs)
Farmer control (Irrig)

Conclusions from consultation the sugar estate

…. and for the sugar estate:

Category	Variables
Management objectives	Temple river flow, Rural water supply, Urban water supply, Rice yield, Sugar estate productivity
Interventions	New well drilling, Well rehabilitation, Rainwater harvesting, Renovate estate reservoir, New estate reservoir, Drip irrigation, Water recycling, New reservoirs, Sand mining control
Intermediate factors	Estate storage, Surface storage, Irrigable area, River extraction (Estate and Paddy)
Controlling factors	River extraction (WS)
Implementation factors	Investment funds, Political resistance
Additional impacts	None

Rural water supply
New well drilling
Well rehabilitation
Rainwater harvesting

Urban water supply
Rainwater harvesting

Estate productivity
Drip irrigation
Renovate estate storage
New estate reservoir
Water recycling

Temple river flow
New reservoirs
Sand mining control
Drip irrigation
Renovate estate storage
New estate reservoir
Water recycling

Rice yield
New reservoirs
Sand mining control

Conclusions from comparison

Although your stakeholder consultation has not produced any new management objectives for your management strategy, it has uncovered a number of potential interventions that you had not previously considered. It has also served to involve stakeholder representatives in the decision process and to give them, you hope, some sense of ownership of the process. To some degree, you hope that this will facilitate implementation of the interventions you finally decide on.

Comparing the BN diagrams and the lists you've drawn up, suggests that there are no major differences in perspective between the stakeholder groups. You realise that this is largely because each group has concentrated on the management objectives that relate to it specifically, so there is no real potential for conflict. You note again, however, the estate managers concern that, if storage is increased too much, then there may not be sufficient water to fill it all, even in the wet season.

Steps 7 & 8: Hold joint stakeholder workshop to discuss different viewpoints; complete stakeholder BNs

As there are no major differences in stakeholder viewpoints, you decide that there is no need to hold a joint stakeholder workshop at this point.

Step 9: Construct 'master' BN diagrams

You begin constructing the "master" BN by combining the diagrams from each of the three user groups. You are happy that their BN diagrams represent your understanding of what the issues are and how they can be solved. You also decide that a single BN will adequately represent issues of equity, as there are management objectives that relate to each of the main user groups within the one network. However, you feel that neither sustainability nor the wider human and environmental consequences have been adequately accounted for.

To address the wider consequences of the proposed management interventions, you include variables to represent human health, soil erosion and water quality. You link these to existing variables in the network that you feel logically feed into them. Considering sustainability, you decide that there are three "feedbacks" which need to be accounted for:

1. Increases in groundwater use to support improved rural and urban supplies may exceed the sustainable yield of the aquifer;
2. Unemployment may affect the level of illegal extractions;
3. Changes in river water quality may affect rice yields.

You also note that some of the interventions need to be implemented continuously if they are to be maintained in the longer term. You list these interventions as:

- ✦ Rainwater harvesting
- ✦ Water recycling
- ✦ Drip irrigation
- ✦ Law enforcement
- ✦ Sand mining control

Following the guidelines suggested in "Key skills 4", you replicate the necessary elements of the BN to create a second step which allows you to examine both the feedbacks and the continuing implementation of the interventions listed above. You include a new variable called "Sustainable yield" to mediate the feedback created by changes in rural and urban water use.

To complete the master BN diagram, you delete the variables representing Political resistance, Investment funds and Support from authority. You decide that these factors are better considered implicitly once the most suitable interventions have been identified.

The completed master BN diagram is shown below. Having completed it, you consider whether it will allow you to investigate how the management objectives you identified in Step 1 can be met. You decide that it can and proceed to Step 10.

Step 10: Collect data and specify conditional probability tables (CPTs)

You consider each variable and decide to use the following information types to fill in the CPTs for each of them:

Faecal pollution	Information Type 2
Illegal extractions	Information Type 4
Off-take condition	Information Type 4 (you consulted the paddy farmers but felt they were overconfident and adjusted their estimates accordingly)
Surface storage	Information Type 4
Irrigable area	Information Type 4
River extraction (Paddy)	Information Type 2
Input availability	Information Type 4 (you consulted the paddy farmers but felt they were overconfident and adjusted their estimates accordingly)
Input quality	Information Type 4 (you consulted the paddy farmers but felt they were overconfident and adjusted their estimates accordingly)
Estate storage	Information Type 2
River extraction (Estate)	Information Type 2
Paddy demand met?	Information Type 4 (calculation based on crop water requirements)
Inputs	Information Type 2 (this represented a simple expression of the stakeholders preference of availability as opposed to quality)
Rural water supply	Information Type 4
Urban water supply	Information Type 4
Health	Information Type 4
Estate productivity	Information Type 2
Temple river flow	Information Type 3 (a deterministic hydrological model was used to run a Monte-Carlo sampling procedure)
Soil erosion	Information Type 4
Rice yield	Information Type 4
River water quality	Information Type 3 (a deterministic water quality model was used to run a Monte-Carlo sampling procedure)
Unemployment	Information Type 4

The states of all the intervention nodes plus River extraction WS (a controlling factor) are given equal probabilities, since they will be changed during Step 11 to examine the impact they have on the management objective variables.

Examples of some of the probabilities elicited from the stakeholders are shown below:

CPT for Faecal pollution variable

Health education	***Fence off-take***	P(Faecal pollution=Yes)	P(Faecal pollution=No)
Yes	Yes	0.10	0.90
Yes	No	0.60	0.40
No	Yes	0.25	0.75
No	No	1.00	0.00

CPT for Estate storage variable

Renovate estate reservoir	***New estate reservoir***	P(Estate storage =Sufficient)	P(Estate storage =Insufficient)
Yes	Yes	1.00	0.00
Yes	No	0.80	0.20
No	Yes	1.00	0.00
No	No	0.00	1.00

CPT for Unemployment variable

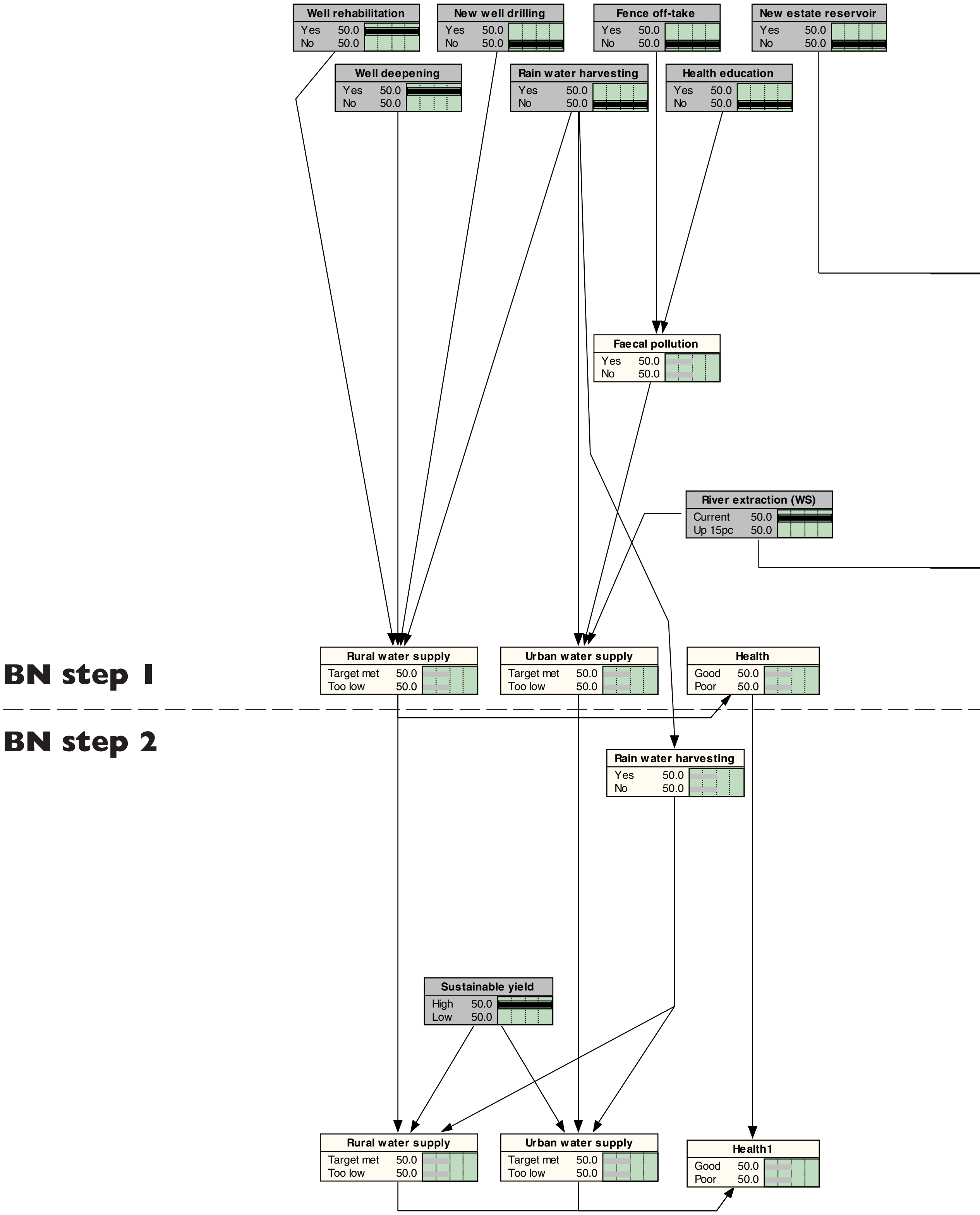
Sand mining control	***Estate productivity***	P(Unemployment =Increase)	P(Unemployment =Static)
Yes	Good	0.10	0.90
Yes	Poor	0.80	0.20
No	Good	0.00	1.00
No	Poor	0.70	0.30

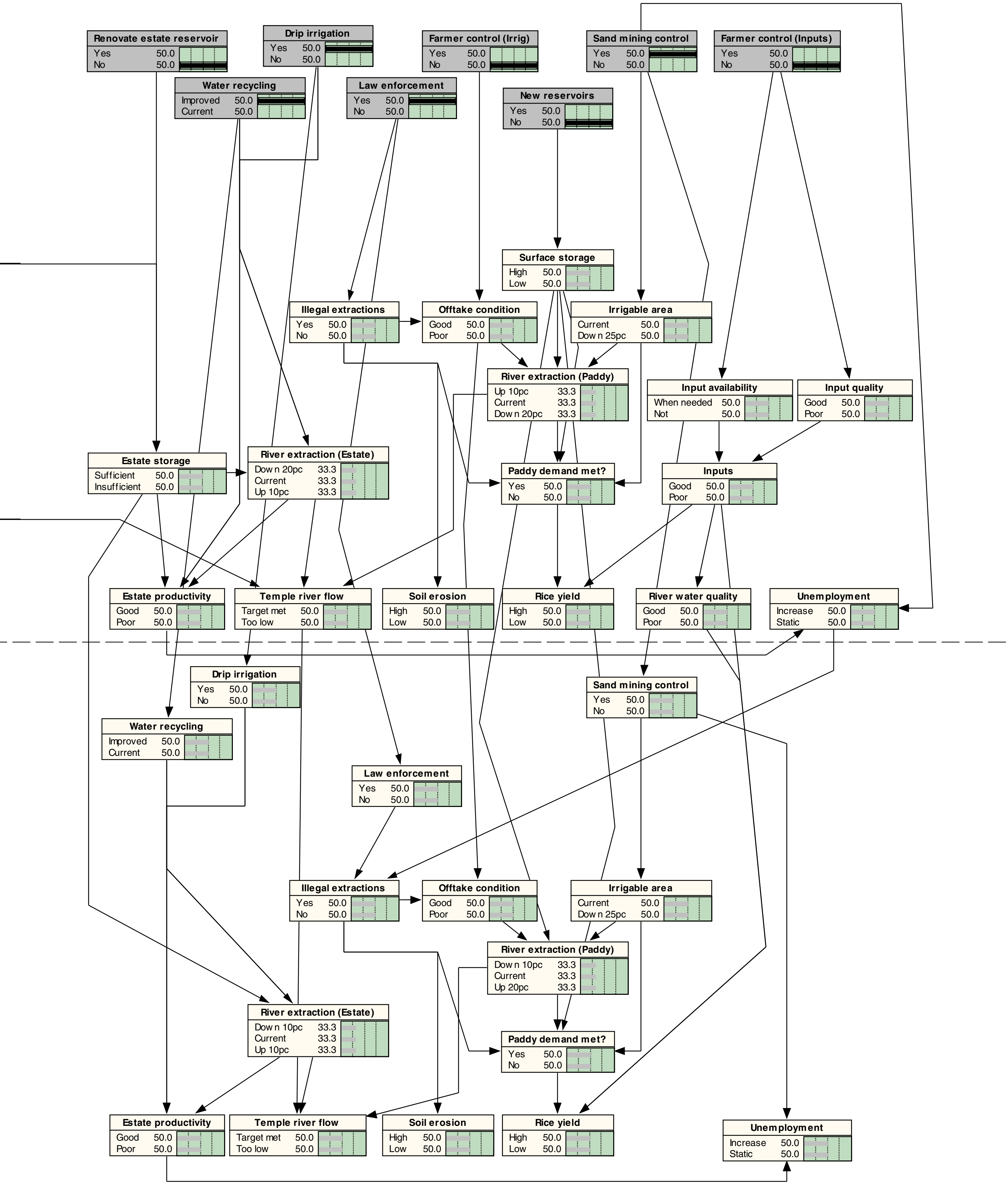
Step 11: Use master BN to make decision

Following the guidelines, you examine the changes that are produced in the management objective variables when the values of each intervention are changed. Some of the results surprise you. For example, improving Estate storage (which you believe should have a largely positive effect) *reduces* Estate productivity significantly. You consider the structure of the network and realise that this change is taking place because Estate storage is only linked to Estate productivity via River extraction (estate). The network suggests that an increase in Estate storage will reduce the estate's River extraction (because they need less water at peak times) which will, in turn, reduce the productivity of the estate (because it has less water). What the BN is not representing is the fact that if Estate storage is sufficient, then a decrease in River extraction will have only a minimal effect on Estate productivity.

To capture this interaction, you realise that you need to link Estate storage directly to Estate productivity as well as indirectly via River extraction. By doing this, you can enter different values in the Estate productivity CPT to represent the fact that Estate productivity will respond differently to River extraction, depending on the state of Estate storage.

Another surprising result is that controlling Sand mining tends to reduce the chance of an increase in Unemployment. Again, this is not what you expect. On examining the Unemployment CPT, you realise that you have entered the values provided by your stakeholders into it incorrectly. Correcting this error provides the result you expect.


Well rehabilitation
Yes 50.0
No 50.0
New well drilling
Yes 50.0
No 50.0
Fence off-take
Yes 50.0
No 50.0
New estate reservoir
Yes 50.0
No 50.0
Well deepening
Yes 50.0
No 50.0
Rain water harvesting
Yes 50.0
No 50.0
Health education
Yes 50.0
No 50.0
Faecal pollution
Yes 50.0
No 50.0
River extraction (WS)
Current 50.0
Up 15pc 50.0
BN step 1
Rural water supply
Target met 50.0
Too low 50.0
Urban water supply
Target met 50.0
Too low 50.0
Health
Good 50.0
Poor 50.0
BN step 2
Rain water harvesting
Yes 50.0
No 50.0
Sustainable yield
High 50.0
Low 50.0
Rural water supply
Target met 50.0
Too low 50.0
Urban water supply
Target met 50.0
Too low 50.0
Health1
Good 50.0
Poor 50.0


Renovate estate reservoir
Yes 50.0
No 50.0
Drip irrigation
Yes 50.0
No 50.0
Farmer control (Irrig)
Yes 50.0
No 50.0
Sand mining control
Yes 50.0
No 50.0
Farmer control (Inputs)
Yes 50.0
No 50.0
Water recycling
Improved 50.0
Current 50.0
Law enforcement
Yes 50.0
No 50.0
New reservoirs
Yes 50.0
No 50.0
Surface storage
High 50.0
Low 50.0
Illegal extractions
Yes 50.0
No 50.0
Offtake condition
Good 50.0
Poor 50.0
Irrigable area
Current 50.0
Down 25pc 50.0
River extraction (Paddy)
Up 10pc 33.3
Current 33.3
Down 20pc 33.3
Input availability
When needed 50.0
Not 50.0
Input quality
Good 50.0
Poor 50.0
Estate storage
Sufficient 50.0
Insufficient 50.0
River extraction (Estate)
Down 20pc 33.3
Current 33.3
Up 10pc 33.3
Paddy demand met?
Yes 50.0
No 50.0
Inputs
Good 50.0
Poor 50.0
Estate productivity
Good 50.0
Poor 50.0
Temple river flow
Target met 50.0
Too low 50.0
Soil erosion
High 50.0
Low 50.0
Rice yield
High 50.0
Low 50.0
River water quality
Good 50.0
Poor 50.0
Unemployment
Increase 50.0
Static 50.0
Drip irrigation
Yes 50.0
No 50.0
Sand mining control
Yes 50.0
No 50.0
Water recycling
Improved 50.0
Current 50.0
Law enforcement
Yes 50.0
No 50.0
Illegal extractions
Yes 50.0
No 50.0
Offtake condition
Good 50.0
Poor 50.0
Irrigable area
Current 50.0
Down 25pc 50.0
River extraction (Paddy)
Down 10pc 33.3
Current 33.3
Up 20pc 33.3
River extraction (Estate)
Down 10pc 33.3
Current 33.3
Up 10pc 33.3
Paddy demand met?
Yes 50.0
No 50.0
Estate productivity
Good 50.0
Poor 50.0
Temple river flow
Target met 50.0
Too low 50.0
Soil erosion
High 50.0
Low 50.0
Rice yield
High 50.0
Low 50.0
Unemployment
Increase 50.0
Static 50.0

Having satisfied yourself that the structure of the BN is correct and that the CPTs properly represent the interactions you expect, you draw up the following table, showing percentage changes in the probability of management objective variables being in their positive state, given implementation of various interventions:

	Management objectives							
Interventions	Rural water	Urban supply	Health	Estate prod.	Temple river flow	Soil erosion	Rice yield	River wat.qual.
Well rehabilitation	+20.0	0	+14.0	0	0	0	0	0
Well deepening	+10.0	0	+13.0	0	0	0	0	0
New well drilling	+40.0	0	+28.0	0	0	0	0	0
Rainwater harvesting	+20.0	+20.0	+20.0	0	0	0	0	0
Fence off-take	0	+22.5	+6.8	0	0	0	0	0
Health education	0	+12.0	+3.6	0	0	0	0	0
New estate reservoir	0	0	0	+4.0	+3.4	0	0	0
Renovate estate res.	0	0	0	+3.2	+2.7	0	0	0
Water recycling	0	0	0	+2.5	+1.0	0	0	0
Drip irrigation	0	0	0	+12.0	+1.7	0	0	0
Law enforcement	0	0	0	0	-0.3	+40.0	+3.6	0
Farmer control (irrig)	0	0	0	0	-1.5	0	+0.5	0
New reservoirs	0	0	0	0	+2.4	0	+11.2	0
Sand mining control	0	0	0	0	-2.2	0	-1.3	0
Farmer control (inputs)	0	0	0	0	0	0	+4.2	-0.5
Well rehabilitation + Well deepening	+50.0	0	+35.0	0	0	0	0	0
Well rehab. + New well drilling	+50.0	0	+35.0	0	0	0	0	0
Water recycling + Drip irrigation	0	0	0	+40.0	+2.4	0	0	0

You note that only Rainwater harvesting affects more than two management objectives positively. It is also clear that improvements in water supply management objectives are likely to be larger than the changes that may be possible in the other gaols. In terms of the BN structure, you realise that this is because the water supply management objectives are much closer to their interventions. However, you believe that this properly represents the reality of the situation in that rural and urban water supplies can be affected much more directly.

At first you are uncertain about the negative results produced in Temple river flow by Law enforcement, Farmer control (irrigation) and Sand mining control. Through studying the BN in more detail, you realise that this is because all of these interventions will enable the paddy farmers to extract more water for irrigation. In the case of Law enforcement and Farmer control (inputs), this has a positive effect on Rice yield, although there are also other interventions that have a more beneficial impact on yield without the negative impact on flow.

However, you find the fact that Sand mining control also has a negative impact on rice yield too slightly surprising. By examining the changes produced by the intervention in the intermediate variables between it and the Rice yield management objective, you realise that this is because a reduction in sand mining will halt the decrease in irrigable area. This will tend to decrease the water available to each farmer and so reduce yield. So sand mining may be beneficial in that it reduces irrigable area (as the river bed level is being decreased) and so increases the depth of water that can be applied over that area. However, this obviously has implications for the total amount of food produced so you conclude that achieving the correct balance between irrigable area and yield should be an important feature of your management plan.

Your use of the BN has highlighted a number of interesting combinations:

1. Well rehabilitation and Well deepening have a greater impact together than the sum of their separate impacts. This is not too surprising as this interaction was highlighted during your consultations with the domestic water users.

2. Well rehabilitation and New well drilling have a lesser impact when combined than the sum of their separate impacts. By investigating the values in the Rural water supply CPT, you realise that this is because a well rehabilitation programme will reduce the need for new wells. As this makes sense, you accept the results provided by the network. You also notice that this combination of interventions has the same effect as the Well rehabilitation and Well deepening combination. As you suspect this latter combination will be less costly, you rule out New well drilling as an intervention.

3. When Water recycling is combined with Drip irrigation, their combined effect is much greater on the Estate productivity but lesser (although still positive) on the Temple river flow. This suggests to you that, in terms of saving water, implementing both of these interventions may be excessive. However, in terms of productivity, additional benefits can be gained from implementing both. You also note that this combination has a much greater effect than alternatives such as improving estate reservoir storage.

Next you examine the changes in the state of the management objective variables between the first and second time steps. Although the feedbacks you included do lead to reductions in the chance of management objectives being in their positive states, their impact is fairly minimal and can be neglected at this stage as monitoring of these management objectives during implementation will reveal any problems not highlighted by the BN. However, the case of Rainwater harvesting provides an exception to this rule. Based on your experiences of implementing rainwater harvesting systems in the past, you do not believe that many people will continue to use these systems beyond the first year. As this experience has been captured in the BN, it indicates that although rainwater harvesting will have a beneficial impact in the first year, its subsequent impact is greatly reduced.

On the basis of this analysis, you decide to consider further the following interventions:

1. A combined programme of well rehabilitation and well deepening to address water supply needs. If funding allows, fencing the river off-take could also be implemented;
2. Implementation of water recycling and drip irrigation at the sugar estate;
3. Construction of small reservoirs for paddy farmers.

With the help of the Chairman of the District Water Resource Committee and the other government departments involved, you will consider the feasibility of these interventions in terms of:

1. The funds available for investment (if funds are limited you will need to ascertain whether targeting the sugar estate or targeting paddy farmers will be most beneficial for the basin as a whole);
2. The support for the interventions that can be expected from the relevant authorities;
3. Any political resistance to change that may be present;
4. Whether runoff is sufficient to fill the planned reservoirs to capacity.

In the next step, you will also discuss whether these interventions are acceptable to the stakeholder groups. You realise that securing their agreement is crucial if implementation is to succeed.

Step 12: Hold a second joint stakeholder workshop to discuss your decisions

You present your list of proposed interventions to all the stakeholders with whom you have consulted. You explain the reasons for rejecting some of their interventions and, as this is a *very* hypothetical case study, you are pleasantly surprised that they all agree with you! As full funding has been secured, all authorities are supportive and the Irrigation Department has confirmed that there is sufficient runoff, your recommendations will now be implemented. You realise that their success is not guaranteed but you are happy that you have made the best choice given the information available to you. Effective adaptive management will ensure that any unforeseen negative consequences can be avoided.

Appendix 1: Filling in conditional probability tables using information types 1 and 3

The general approach

Step 10 of the guidelines explains that it is necessary to fill in the conditional probability table (CPT) associated with each node in your network that is a child (almost all the nodes in your network will be children, although most of them will also be parents). Make sure you have read the guidelines in Step 10 in full before reading this Appendix, which will describe how to fill in a CPT if you wish to use information of Types 1 and 3. In Step 10, these were defined as:

Information Type 1:	Raw data collected by direct measurement (e.g. groundwater depth measured by piezometer, population measured by census, income measured by accounting)
Information Type 3:	Output from process-based models calibrated using raw data collected by direct measurement

If you are using information of types 1 and 3, then to fill in the CPT for the child node you are considering, you will need at least one piece of information relating to each of the possible combinations of states of the parent nodes (see Example 10.1 in Chapter 4). We will call each such piece of information a "case". It is important to understand that each case refers to a particular combination of the states of the parent nodes, as represented by a single row in a conditional probability table (or in the table below).

Referring to Example 10.1, a number of possible cases is shown below:

Groundwater levels	Number of wells	Water supply (quantitative)	Water supply (qualitative)
High	Lots	92 % of farmers	Good
High	Lots	97 % of farmers	Good
High	Lots	78 % of farmers	Poor
High	Lots	93 % of farmers	Good
High	Lots	92 % of farmers	Good
Low	Few	56 % of farmers	Poor
Low	Few	95 % of farmers	Good
Low	Few	43 % of farmers	Poor

Note that only two of the 4 possible parent state combinations noted in Example 10.1 are listed in the table above. Also note that case data is more likely to be available in quantitative terms. If this is the case, then you may need to convert it into the qualitative state names used by your BN (High, Lots, Good etc.). This is easily done using the definitions you have assigned to each state name (see Example 10.1. Note that both Netica and Hugin allow you to quantitatively define your qualitative state names for each node, allowing this conversion to take place automatically.

Clearly, each case gives us some knowledge about the state the child should be in when its parents have a particular combination of states. As the natural environment is highly variable and uncertain, it is rare that all the cases will agree on what state the child should be in for each particular combination of parent states. This is reflected in the table above. It is this uncertainty that we want to estimate and enter into the CPT.

We can do this by counting the number of cases that report a particular state for the child and comparing it to the total number of cases that refer to a particular parent state combination. This will give us the probability that the child is in that state given the parent state combination. This can then be entered into the appropriate CPT.

It is important to note that the more cases you have the more accurate the results of this procedure will be. As a rule of thumb, you should try to obtain at least 20 cases for each possible combination of parent states. If this is not available with Type 1 information you should consider using Type 3.

The general procedure described above is automated by software packages such as Netica and Hugin which have Bayesian learning algorithms built into them. It is recommended that you use these, although only after you have read and understood the documentation provided by the help files[1]. If the software package you are using does not have this feature, then follow the guidelines below. These provide a simple approximation to the learning algorithms used by both Netica and Hugin. Note, however, that using the approximation below together with the algorithms in Netica and Hugin can cause problems. This will be discussed in more detail below.

Approximate learning of case data

1. Note the number of states, n, the child node has.

For Example 10.1, the child node "Water supply" has two states, Good and Poor. Therefore, n = 2.

2. List the possible combinations of the states of the parent node.

For Example 10.1, these are:
Groundwater levels = High, Number of wells = Lots
Groundwater levels = High, Number of wells = Few
Groundwater levels = Low, Number of wells = Lots
Groundwater levels = Low, Number of wells = Few

[1] In Netica see "Learning from cases" in the Contents page of the Help menu. In Hugin see "Learning adaptation" and "Learning EM" in the Hugin Runtime manual.

3. Start with the first combination of parent states. If you wish, you can give an *initial* estimate of the probability that the child is in each of its states, for this parent state combination (remember that these probabilities must add up to 1). You should do this if you have good reason to think that the probabilities should be set at certain values, *prior* to looking at any case data you may have (which is why such probabilities are termed "prior probabilities"). Although these prior probabilities will be personal, subjective estimates, they will be based on your past experience. If you feel you do not have enough experience to choose prior probabilities, then you should give each child state the same probability (equal to 1/n). In the text that follows, prior probabilities will be denoted as $p^0(i)$, where the superscript 0 shows that it is a prior probability and *i* indicates the state of the child which is being referred to.

In the example above, i=1 refers to when "Water supply" is in the state Good and i=2 refers to when "Water supply" is in the state Poor. Based on your 20 years of experience in the water supply sector, you estimate the prior probabilities for the first parent state combination (when Groundwater levels = High, Number of wells= Lots) to be:

$$p^0(1) = 0.7,\ p^0(2) = 0.3$$

4. Express your uncertainty in these prior probabilities by saying that the extent of your knowledge is equivalent to you having observed a certain number of cases, N^0. If you are very certain that these probabilities are accurate then it may be appropriate to say that your certainty is equivalent to 100 cases. If you are fairly certain, then you might choose 50 cases. Refer to *"A note on the automatic learning algorithms in Netica and Hugin"* below for more guidance. If you did not assign any prior probabilities (i.e. you gave each child state a probability of 1/n because you weren't sure what else it might be) then let $N^0 = 0$.

In the example above, you think that the prior probabilities you have assigned to the first parent state combinations are good guesses but you wouldn't put a great deal of faith in them. Consequently, you choose $N^0=5$.

5. Next, look at the case data available to you. Count the number of cases which refer to the first combination of parent states. Call this number N.

Given the case examples above, for Groundwater levels = High, Number of wells = Lots, N = 5. Note that if there had been no cases (as for Groundwater levels = Low, Number of wells = Lots) then N = 0.

6. Count how many of N indicate the child to be in its first state. Then count how many of N indicate the child to be in its second state. Do this for each of the n child states. Call these counts N(i) where $2 \leq i \leq 0$ (assuming all variables have at least two states). Water supply has 2 states (n=2). There are N = 5 cases in total: 4 of them show Water supply = Good, one shows Water supply = Poor.

Counting how many times Water supply = Good gives N(1) = 4 and counting how many times Water supply = Poor gives N(2) = 1.

7. Update the prior probability for each of the child's states, using the formula:

$$p^1(i) = [N(1)+(p^0(i)N^0)] / N^1$$

where $p^1(i)$ is the updated probability estimate and $N^1 = N^0 + N$

In the example above, the probabilities would be calculated as:

$$p^1(1) = [4 + (0.7 \times 5)] / 10 = 0.75$$

$$p^1(2) = [1 + (0.3 \times 5)] / 10 = 0.25$$

8. Enter these probabilities into the appropriate row in the CPT (as indicated by the parent state combination).

9. Repeat instructions 3 to 5, above, for each parent state combination

Using information Type 3

Information Type 3 is produced from process-based models. If a model is to provide information relevant to a particular child's CPT, then it must take the parent node values as input (i.e. the states the parents are in) and output a value for the child node (i.e. the state the child is in). Depending on the type of model and the way in which it is run, this output will be in the form of a single value, many different values or a probability distribution across all the possible values that the child can take. The best way to fill in CPTs using each of these different forms of output will be discussed below.

Model output is in the form of a probability distribution across the values of the child node

Both Netica and Hugin can automatically fill in CPTs if the form of the distribution is known (e.g. normal or geometric) together with the statistics describing it (e.g. mean and variance)[2]. Otherwise, the same information can be used to calculate the probabilities associated with each child state manually. As the methods to do this are beyond the scope of these guidelines, it is recommended that you consult a statistician.

Model output is in the form of many different values for the same parent state combination

It is likely that such output will have been produced by Monte-Carlo sampling (see below) and you will have many cases for each parent state combination, as described in the general approach above. These should be fed into the automatic learning algorithms supplied by Netica or Hugin or the approximate learning method, described above, can be used.

Model output is in the form of a single value

Unless the relationship described by the variables in the CPT is truly deterministic (i.e. given the same parent state combination, then the child will always be in the same state) then you will need to generate some estimates of the uncertainty associated with the single value output. There are two ways to do this:

- Analytical uncertainty estimation

Both exact and approximation statistical techniques exist to calculate the uncertainty in a variable that is a function of other variables whose uncertainties are known. So, for example, if the uncertainty associated with

[2] In Netica see "Equations" in the Contents page of the Help menu. In Hugin see "Expressions" in the Hugin Runtime manual.

the input to a model is known together with the uncertainty in the model parameters, then the uncertainty in the model output can be calculated. Unfortunately, unless the model is very simple, it is very difficult to solve these equations accurately. However, if analytical equations can be used to produce uncertainty estimates in model outputs, then they will produce probability distributions across the states of the child. These can then be fed into CPTs, as described above.

- Empirical Monte-Carlo uncertainty estimation

Monte-Carlo sampling is where a model is repeatedly run with different sets of input variables and model parameters. These different sets are produced by randomly varying the input variables and parameters within defined limits. Each run will produce a single case, which can be fed into the CPT using either the automatic learning algorithms or the approximate learning method. As a detailed description of the Monte-Carlo sampling is beyond the scope of these guidelines, it is recommended that you consult a statistician.

Uncertainty about uncertainty

If the probabilities you have calculated for a CPT are based on only a few cases, then you should not place great confidence on their accuracy. For example, imagine that you only have the cases listed in the table at the beginning of this Appendix. Using the approximate learning method, and assuming you didn't specify any prior probabilities for the last three parent state combinations, you would calculate the probabilities for the CPT as follows:

Groundwater levels	**Number of wells**	***P (Water supply = Good)***	***P (Water supply = Poor)***
High	Lots	0.75	0.25
High	Few	0.50	0.50
Low	Lots	0.50	0.50
Low	Few	0.33	0.67

Now imagine that you obtain a new case, as follows:

Groundwater levels = High, Number of wells = Lots, Water supply = Poor

The probabilities in the first row of your CPT (for Groundwater levels = High, Number of wells = Lots) would now change to:

$p^2(1) = [0 + (0.75\times 10)] / 11 = 0.68$

$p^2(2) = [1 + (0.25 \times 10)] /11 = 0.32$

This represents a change of about 10% in the probability for p(1) and may have important consequences on the probabilities that the BN will calculate for the success of your management objectives (although this will depend on the values given to other variables in your network).

Obviously, the problem is even greater where no cases were previously available. Imagine a second new case, as follows:

Groundwater levels = High, Number of wells = Few, Water supply = Poor

The probabilities in the second row of your CPT (for Groundwater levels = High, Number of wells = Few) would now change to:

$p^1(1) = [0 + (0.50 \times 0)] / 1 = 0$

$p^1(2) = [1 + (0.50 \times 0)] / 1 = 1$

This is a change of 100%!

So, the fewer cases you have to calculate the probabilities in the CPT (which represent your uncertainty about the relationship between a child and its parents), the more uncertain you are about what those values are. This is uncertainty about uncertainty!

There are several ways to deal with this. This problem should only arise with Type 1 information — with a working model it should be possible to produce as much Type 3 information as you need. Therefore, the most straight-forward approach to dealing with a shortage of Type 1 information may be to develop a simple model, calibrate it with the data you have and then run it in a Monte-Carlo procedure to produce multiple cases.

Alternatively, you should examine the sensitivity of your network to changes in the probabilities contained in the CPT. This can be done "by hand" by opening the CPT, changing the value about which you are unsure, running the BN again and examining the affect the change has on your management objective. If the change in the management objective is not great, then you need not worry. If it is, then you should note the degree of the change and consider it when making any final decisions using the BN.

Using Example 10.1 in Chapter 4 again, imagine that we have calculated the probabilities for the CPTs (as in the first table overleaf).

CPT for groundwater levels:

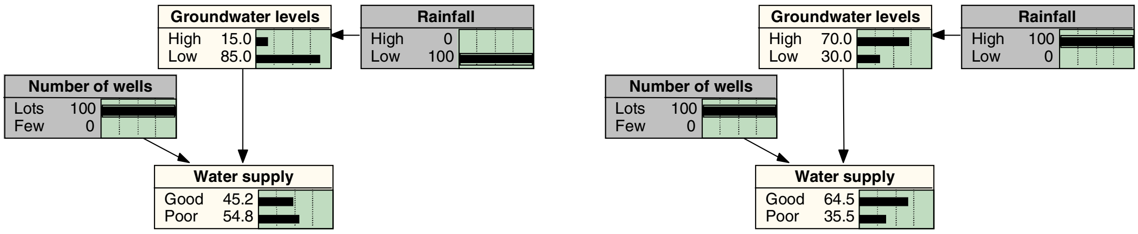
Rainfall	P (GW levels = High)	P (GW levels = Low)
High	0.70	0.30
Low	0.15	0.85

CPT for water supply:

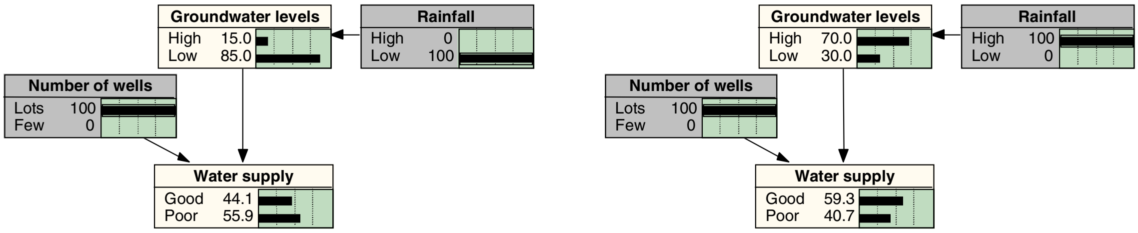
Groundwater levels	**Number of wells**	***P (Water supply = Good)***	***P (Water supply = Poor)***
High	Lots	0.75	0.25
High	Few	0.30	0.70
Low	Lots	0.40	0.60
Low	Few	0.10	0.90

Assume we have more than 20 cases for all parent state combinations in both CPTs, with the exception of Groundwater levels = High, Number of wells = Lots in the CPT for Water supply. For this parent state combination, we only have the five cases contained in the table, and therefore we are uncertain that the probabilities we have calculated (0.75 and 0.25) are accurate. To investigate how this might affect the management objective (Water supply, itself) we can change the probabilities and watch what happens. The results of doing this are shown below:

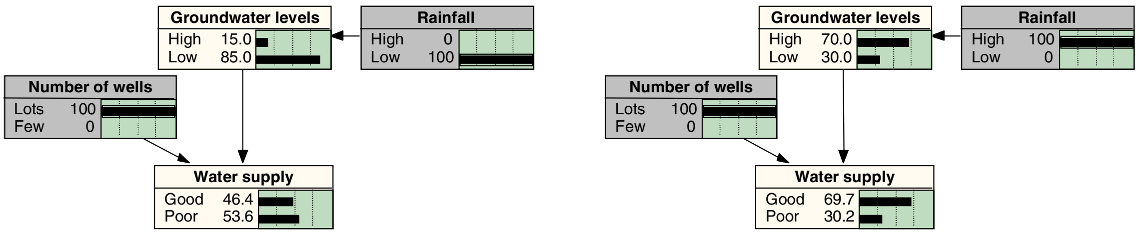
P (Water supply = Good) = 0.75, P (Water supply = Poor) = 0.25



P (Water supply = Good) = 0.675, P (Water supply = Poor) = 0.325 (10% worse)

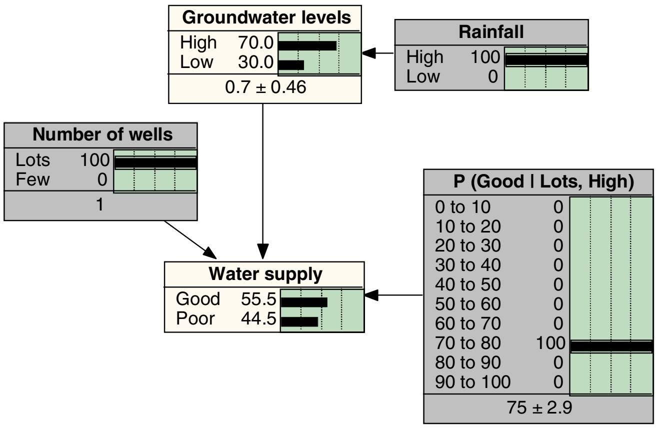


P (Water supply = Good) = 0.825, P (Water supply = Poor) = 0.175 (10% better)



Note that for Low Rainfall (the BNs on the left), the change in water supply is negligible. However, for High Rainfall (the BNs on the right) the change is fairly large. Potentially, your uncertainty about the uncertainty will affect your management decision.

The procedure outlined above (of changing the CPTs by hand) can be automated by creating a node representing the probability that should be in the CPT. As this is quite a technical procedure, full details are not provided here. However, the BN structure to do this for the network above is shown below together with the equations you will need to specify the Water supply CPT. Note that the equations shown are those used by Netica. Hugin uses different syntax. Consult the help file in either package for further details (see footnote 2 on page 94).



Code for Netica

Note that the equations use system names as follows:

- WattSupp = Water supply
- NoWells = Number of wells
- GWLevels = Groundwater levels
- P_Good = P(Good | Lots, High)

```
P (WatSupp | NoWells, GWLevels, P_good) =
(NoWells == 1 && GWLevels == 1 && WatSupp == 1) ? P_Good:
(NoWells == 1 && GWLevels == 1 && WatSupp == 0) ? 100 - P_Good:
(NoWells == 1 && GWLevels == 0 && WatSupp == 1) ? 10:
(NoWells == 1 && GWLevels == 0 && WatSupp == 0) ? 90:
(NoWells == 0 && GWLevels == 1 && WatSupp == 1) ? 20:
(NoWells == 0 && GWLevels == 1 && WatSupp == 0) ? 80:
(NoWells == 0 && GWLevels == 0 && WatSupp == 1) ? 30:
(NoWells == 0 && GWLevels == 0 && WatSupp == 0) ? 70:
1
```

A note on the automatic learning algorithms in Netica and Hugin

The learning algorithms in both Netica and Hugin make use of the idea of "experience". In simple terms, the programs keep count of the number of cases that have been used to calculate the probabilities in the CPTs. The more cases that have been counted, the more 'experienced' the probabilities in the CPT are. The practical effect of this is to reduce the change brought about in the CPTs for each new case learnt, as the number of previously learnt cases increases. If only a few cases have been learnt previously (the CPT has a low experience), then a new one will have a large effect on the probabilities. If many new cases have been learnt previously (the CPT has a high experience), then a new one will have only a minimal effect.

This reflects our ideas about experience in the real world (see the box opposite).

Imagine an experienced stationmaster. He knows that a train always passes through his station at midday every day. He also knows, however, that for a couple of days in every month the train is late. The fact that the train turns up late on one day does not change his expectation about when the train will turn up the next day – he's still fairly sure that it will arrive at midday. Now imagine the experienced stationmaster retires and is replaced by a young man fresh out of station master school. He hasn't had a chance to talk to his predecessor so only the only information he has is the timetable, which says the train comes through at midday. On his first day, however, the train is late and this happens again on the second day. Now the new station master is beginning to believe that the train is always late so when, on the third day, a customer asks what time the train is arriving he replies that it should arrive at midday but is usually late. (*As an aside, if the same question had been asked of the experienced station master he would have been more likely to tell the customer that the train is usually on time and he would have been right. It is obvious that we should have more confidence in the information given by the experienced stationmaster.*) On the third day, however, to the surprise of the stationmaster, the train turns up on time. While the new stationmaster had been fairly confident this morning that the train would be late, he's no longer sure what to think. When the train turns up on time every day for the next two weeks, the new stationmaster becomes increasingly confident that this is what usually happens. So when the train is late again, it doesn't really affect his opinion about when it usually turns up. The point of the example is that, while on the third day his beliefs about the train were radically altered by a new case (the train turning up on time), after two weeks a new case (the train being late again) didn't really change his opinion. This was because he was more experienced after the end of two weeks. Netica and Hugin both work exactly in this way.

The concept of experience reflects the level of confidence you should have in the values contained in the CPT. The more experience that the probabilities in a CPT have, the more confidence you should place in them and the less likely you are to change them as new cases arise.

The table below shows this. The numbers in it were generated using Netica by learning each one of a number of identical nodes with a different number of cases. Once learning had been completed, all the nodes displayed equal probability distributions across their two states (i.e. 50% and 50%). Then each node was learned with one further case and the change in probability was noted as in the table.

Number of cases learned (experience)	Change produced by one more case (from 50 % probability)
0	16.7
20	2.2
40	1.2
80	0.6
120	0.4
160	0.3
200	0.2

It is important that you understand this, particularly if you are intending to use an automatic learning algorithm after you have entered numbers into a CPT using the approximate learning method. This is because entering probabilities produced using the approximate method directly into a CPT does not alter the experience associated with those probabilities. You may have used a large number of cases to calculate a probability using the approximate method, but the software will not know this (unless you tell it) and will assign a low experience value to your CPT. Any subsequent learning using an automatic algorithm will give much more weight to the new data and lead to incorrect results.

Consider the table above. Imagine that you had used 120 cases to calculate a probability distribution of (50%, 50%) across a two-state variable and then entered it directly into the CPT. A single new case learned by the automatic algorithm would change that distribution to (66.7%, 33.3%) because the BN assumes that there is no experience associated with that probability. Clearly, however, the probability should have an experience of 120 associated with it and so should only change to (50.4%, 49.6%) when a new case is learned. There is a big difference between 66.7% and 50.4% and it may well affect the probability assigned to your management objective.

- **As it is not currently possible to alter directly the experience values stored by either Netica or Hugin, then mixing the approximate learning method with the automatic method should be avoided**

This is also true of any values entered directly into the CPT, whether they have been generated by the approximate method or otherwise (see Appendix 2).

Appendix 2: Filling in conditional probability tables using information types 2 and 4

Step 10 of the guidelines explains that it is necessary to fill in the conditional probability table (CPT) associated with each node in your network that is a child (almost all the nodes in your network will be children, although most of them will also be parents). Make sure you have read the guidelines in Step 10 in full before reading this Appendix, which will describe how to fill in a CPT if you wish to use information of types 2 and 4. In Step 10, these were defined as:

Information Type 2:	Raw data collected through stakeholder elicitation (e.g. stakeholder perceptions of groundwater depth, population and income)
Information Type 4:	Academic "expert" opinion based on theoretical calculation or best judgement.

Both of these information types are subjective and are obtained through consultation with an "expert" be that expert an academic or a stakeholder.

You will need to ask your experts to make estimates of the chance that a variable will be in each of its states depending on the states of its parents. Referring to Example 10.1 in Chapter 4, you will ask them to specify directly the probabilities that water supply will be good or poor, depending on the groundwater levels and the number of wells.

Obviously, the way in which you do this will depend on the ability of the stakeholders and the size of the CPT. If the CPTs to be filled in are small *and* you believe the stakeholders will understand the logic underlying them, then you can help them to fill the CPTs in directly. To decide whether this is appropriate, try filling in some of the CPTs yourself. If you do decide that this is an appropriate approach, then you can skip most of this appendix.

If you consider such direct elicitation inappropriate, then it is important to minimise the questions you will need to ask to elicit the necessary information. This will make it possible to complete the CPTs within the time you are likely to have available in practice. It is also important that you consider the best way to ask the questions to each stakeholder as it is important that they understand properly what you are asking them. General guidance on this is provided below, however, if you are unsure how best to proceed, it is recommended that you consult with someone who has local experience of facilitating stakeholder participation.

How to minimise the questions you need to ask stakeholders

This may look quite complicated but don't be put off: with a little practice, it is fairly straightforward. This is particularly true if you have managed to limit the number of states each node has to 2. With more practice, you will also learn how to structure a BN to make this step as easy as possible.

A meeting should be arranged with each stakeholder in isolation from his or her stakeholder group. This is because you need to get an idea of any differences of opinion there may be within the stakeholder group. Be sure to use the same master BN diagram for each member of that group (you may have different master BN diagrams, depending on how you are dealing with issues of equity). Later, you will combine the different answers provided by each stakeholder to give a measure of the differences of opinion between them.

You should prepare questionnaires before the meeting by constructing a table similar to the template shown in Table 1 *for each child node* in the BN diagram, and using it as the focus for the questions listed below. This will be referred to as an "elicited probability table" (EPT) to distinguish it from a CPT. Each EPT will be developed into a CPT but, since EPTs are designed to limit the number of probabilities elicited from the stakeholder, they do not contain all the probabilities required by the CPT. The other probabilities in each CPT will be calculated from their corresponding EPT later.

A general approach to constructing an EPT will be described first, but, depending on the structure of the stakeholder BN diagram, a number of variations may be necessary. These will be covered separately. The general approach should be studied, in any case, as it includes advice that will be useful in all cases. The procedure described for averaging the probabilities in an EPT applies to all variations.

The general approach

Completing an EPT

Take each node in your network one at a time. There is no need to elicit probabilities for those nodes that have no parents. If the BN has been structured correctly, then these nodes will represent either factors that you hope to control directly through management, or factors that you have no control over at all. When you come to use the BN to make a decision, you will change the states of these nodes to see how this affects your management objectives.

In the general approach, it is assumed that the child we are considering has three parents, none of whom affect the degree of change that the others have on the child (i.e. they are NOT modifying parents: see Step 5 in the guidelines). It is also assumed that the child node has two discrete states, one of which is more desirable than the other. In Table 1, this is called the "success" state. Each parent is also assumed to have two discrete[1] states,

[1] A *discrete* variable is one with a well-defined, finite set of states (e.g. number of wells): a *continuous* variable can take a value between any other two values (e.g. rainfall depth).

one of which is more likely to give rise to the success state of the child node. In Table 1, this state will be called the "positive" state. When these assumptions are correct, draw a table like that shown in Table 1.

Table 1: Elicited probability table (EPT) for the general approach

	Non-modifying [2] parent 1 (NMP1)	*Non-modifying parent 2 (NMP2)*	*Non-modifying parent 3(NMP3)*	*Child state and score out of ten*
Question 1:	Positive state of NMP1	Positive state of NMP2	Positive state of NMP3	**Elicited state & score out of ten**
Question 2:	Negative state of NMP1	Negative state of NMP2	Negative state of NMP3	**Elicited state & score out of ten**
Question 3:	Negative state of NMP1	Positive state of NMP2	Positive state of NMP3	**Elicited state & score out of ten**
Question 4:	Positive state of NMP1	Negative state of NMP2	Positive state of NMP3	**Elicited state & score out of ten**
Question 5:	Positive state of NMP1	Positive state of NMP2	Negative state of NMP3	**Elicited state & score out of ten**

The EPT is formally structured, as follows. The first line (question 1) is such that the parents are all in their positive states. The second line (question 2) is such that the parents are all in their negative states. For all the other lines, each parent is "switched" from its positive state to its negative state. This is done one parent at a time so that, after the first two, each line only ever has one state that is negative.

Before beginning the interview, you should run through, with the stakeholder, the sort of questions you are going to ask. Explain that you will ask the questions in sequences with each sequence being linked to the same child node.

For question 1, ask each stakeholder to imagine a situation such as that described by the states of the parents on the first row. If it is appropriate, given the nature of the child you are considering, ask them to imagine this situation as it directly affects them. For example, if the child is 'Crop yield', ask them to think about their own fields. Also, get them to describe their fields, as this will help them to 'picture the scene', but may also provide information that will be useful later (see the next section).

Now ask whether they believe that this situation will cause the child node to be in the success state or in its 'failure' state and note down the answer in the box marked "Elicited state and score out of ten" for question 1 in the table. Although this gives you an answer to the question, you also need to know how certain the stakeholder is about this answer. The following questions should help you estimate this.

Ask them to imagine ten years in which the situation described by the states of the parents existed. If they can, ask them to think about ten *actual* years, directly from their own experience. If they can't do this (perhaps because

[2] A *non-modifying parent* is one whose effect on its children is *independent* of the states of any other parents of those children (see Step 5 in Chapter 4 for more details).

there have only been four years when this situation occurred), then ask them to imagine the situation occurring for ten years. Having done this, ask them to estimate for how many of the ten years that they are imagining would the child be in the state that they say. Write this into your table too. If a stakeholder does not think that the child will be in the same state for all ten years, then you should find out the reasons why he or she thinks this, if time permits.

Question 2 can then be asked in exactly the same way, although, this time, you should ask the stakeholder to imagine the different situation described by the states of the parents on the second row (i.e. all parents are negative). By asking questions 1 and 2 in Table 1, you should have established the highest and lowest probabilities of the child being in the success state. Questions 3, 4 and 5 can then be asked in the same way but you should remind the stakeholder of the range of probabilities from which they should, logically, choose (as defined by questions 1 and 2). If the stakeholders decide that they have to choose a probability outside of the range specified by questions 1 and 2, then you should ask them to re-evaluate their original answers to these questions.

When you have completed the EPT, review it with the stakeholder to check that the answers are logical. If they are not logical, then you should try to understand why the stakeholder has chosen the seemingly illogical values. If, after consideration, the stakeholders would like to change some of the values they have chosen, then you should allow them to do so.

Averaging EPT probabilities across stakeholder groups

If you have a number of master BN diagrams (so you can investigate whether the distribution of benefits between groups is equitable), then, for each BN diagram, you will need to average the probabilities elicited from the stakeholders whose group it represents. Use the probabilities you elicited from each stakeholder within the group to form a single EPT for the group. Obviously, you will need to repeat this procedure for each stakeholder group, separately.

Take each child node in turn. Using the scores obtained from each individual stakeholder, calculate the scores for all the states of the child node, as follows. In cases where the child has more than two states (discrete or continuous sub-ranges) you will know this already as you will have asked the stakeholder directly. In cases where the child node has two states, simply subtract the score given by the stakeholder from 10 and the remainder is the score for the other state. Multiply all the scores by 10 to convert them into percentage probabilities.

For each child node, compare the probabilities calculated from each member of the group. For each of the situations listed in the EPTs for that node, calculate the average of the probabilities given by each stakeholder for each state of the child node. An example is shown below for a case where the child has two states:

	Probability of success (state 1 of the child node)	*Probability of failure* (state 2 of the child node)
Stakeholder 1:	30.0 %	70.0 %
Stakeholder 2:	80.0 %	20.0 %
Stakeholder 3:	0.0 %	100.0 %
Average:	36.6 %	63.4 %

In some cases, it may be more appropriate to weight the elicited scores before averaging them. This may be the case when the child for whom the value has been elicited represents something that is related to an area or to a certain number of people. For example, If you have followed the guidelines in the previous section, a probability elicited for a child called "Crop yield" should relate to a particular farmer's field. In this case, you should weight each farmer's probabilities by the field areas, as shown below:

	Area of field (hectares)	*Basic probability of success*	*Weighted average probability of success*
Stakeholder 1:	10	30.0 %	30.0 × (10/50) +
Stakeholder 2:	35	80.0 %	80.0 × (35/50) +
Stakeholder 3:	5	0.0 %	0 × (5/50)
Total area:	50		= 62%

Another example might be a child called "Effectiveness of farmer organisation". Different stakeholders you are questioning may belong to different farmer organisations, which may have different numbers of members. If the number of members in an organisation is relevant to your management problem, then you should weight the scores from each farmer by the number of members in their organisation.

Calculating interpolation factors from the EPT

As discussed above, EPTs are designed to limit the number of probabilities elicited from the stakeholder. As a result, they do not contain all the probabilities required by the CPT and it is now necessary to calculate these. The first stage in doing this is to calculate what we will call "interpolation factors". The way in which these are calculated depends, as above, on the structure of the BN diagram. The general case will be dealt with here while variations are covered in subsequent sections.

Note that it is not necessary to calculate interpolation factors for all of the parents – one of the parents (but only one) can be missed out. This is true no matter what the structure of the BN diagram. The reason for this is purely mathematical and you will understand why when you come to use the interpolation factors to calculate all the probabilities in the CPT. By convention, you should not calculate an interpolation factor for the parent listed in the left-most column of the CPT.

Once you have averaged the scores from each stakeholder, you should have a table similar to Table 2. In the table, probabilities P_{1X} and P_{1Y} refer to the probability you elicited for question 1, in the general approach, which you have since averaged over all stakeholders in a particular group. Similarly,

Table 2: Elicited probability table (EPT), following averaging, for general case

Non-modifying parent 1 (NMP1)	*Non-modifying parent 2 (NMP2)*	*Non-modifying parent 3 (NMP3)*	*Probability of: success*	*failure*
Positive state of NMP1	Positive state of NMP2	Positive state of NMP3	$\mathbf{P_{1X}}$	$\mathbf{P_{1Y}}$
Negative state of NMP1	Negative state of NMP2	Negative state of NMP3	$\mathbf{P_{8X}}$	$\mathbf{P_{8Y}}$
Negative state of NMP1	Positive state of NMP2	Positive state of NMP3	$\mathbf{P_{5Y}}$	$\mathbf{P_{5Y}}$
Positive state of NMP1	Negative state of NMP2	Positive state of NMP3	$\mathbf{P_{3X}}$	$\mathbf{P_{3Y}}$
Positive state of NMP1	Positive state of NMP2	Negative state of NMP3	$\mathbf{P_{2X}}$	$\mathbf{P_{2Y}}$

probabilities P_{8X} and P_{8Y} refer to the probability you elicited for question 2 and so on (look at the parent state combinations to work out which row in Table 2 refers to each question in Table 1). Don't worry about the numbering being out of order: it refers to the final position of the elicited probability in the CPT. This will be shown later (see Table 3, for general case).

The interpolation factors are calculated for each 'switch' in the state of a parent from positive to negative. They are all calculated in relation to the difference between the highest probability for the success state (when all parents are positive) and the lowest one (when all parents are negative). Expressed mathematically this is $P_{1X} – P_{8X}$. When one of the parents changes from a positive to a negative state, the probability of the child being in its success state is reduced. The interpolation factor simply quantifies this reduction, for each parent, as a proportion of $P_{1X} – P_{8X}$. Mathematically, this is expressed as:

Interpolation factor for NMP 3: $IF_3 = (P_{2X} – P_{8X}) / (P_{1X} – P_{8X})$	Equation 1
Interpolation factor for NMP 2: $IF_2 = (P_{3X} – P_{8X}) / (P_{1X} – P_{8X})$	Equation 2

You do not need to calculate interpolation factors for the probabilities of failure. Once the interpolation factors have been used to calculate the probability of the success state for each parent state combination in the CPT, the probability of the failure state is simply 1 minus that value.

Filling in the CPTs using the interpolation factors

Table 3 shows the layout of a CPT for child node with 2 states (X and Y), which has 3 parent nodes which all have 2 states (positive and negative - this is the same general case shown in Table 1). As you can see, it lists every possible combination of the parent node states. You will not have elicited probabilities for all of these combinations so you will have to calculate them, using the interpolation factors you have calculated above.

In Table 3, the probability for state combination 3 has been given by the stakeholders, while the probability for state combination 4 has not. However, the only difference between state combinations 3 and 4 is that parent 3 has switched from a positive to a negative state.

In the previous section, we said that the interpolation factors calculate the way the probability of a child changes when a parent node switches from a positive to a negative state. Therefore, to calculate P_{4X}, all we need to do is multiply P_{3X} (scaled by the lowest probability, P_{8X}) by the interpolation factor

Table 3: Conditional probability table (CPT) for the general case

State combination number	State of parent 1	2	3	Probability that child is in state X	Y
1	Positive	Positive	Positive	P_{1X}, elicited from stakeholders	P_{1Y}, elicited from stakeholders
2	Positive	Positive	Negative	P_{2X}, elicited from stakeholders	P_{2Y}, elicited from stakeholders
3	Positive	Negative	Positive	P_{3X}, elicited from stakeholders	P_{3Y}, elicited from stakeholders
4	Positive	Negative	Negative	**P_{4X}, not elicited**	**P_{4Y}, not elicited**
5	Negative	Positive	Positive	P_{5X}, elicited from stakeholders	P_{5Y}, elicited from stakeholders
6	Negative	Positive	Negative	**P_{6X}, not elicited**	**P_{6Y}, not elicited**
7	Negative	Negative	Positive	**P_{7X}, not elicited**	**P_{7Y}, not elicited**
8	Negative	Negative	Negative	P_{8X}, elicited from stakeholders	P_{8Y}, elicited from stakeholders

associated with parent 3 switching from positive to negative (IF_3 in Equation 1, above). This is what Equation 3 does.

Similarly, Equation 4 calculates P_{6X} from P_{5X} in the same way. Equation 5 calculates the probability of state combination 7 from the probability for state combination 5. These state combinations are only different in that the state of parent 2 has changed from positive to negative, so Equation 5 uses IF_2 from Equation 2.

$$P_{4X} = [(P_{3X} - P_{8X}) \times IF_3] + P_{8X} \quad \text{and} \quad P_{4Y} = 100 - P_{4X} \qquad \text{Equation 3}$$

$$P_{6X} = [(P_{5X} - P_{8X}) \times IF_3] + P_{8X} \quad \text{and} \quad P_{6Y} = 100 - P_{6X} \qquad \text{Equation 4}$$

$$P_{7X} = [(P_{5X} - P_{8X}) \times IF_2] + P_{8X} \quad \text{and} \quad P_{7Y} = 100 - P_{7X} \qquad \text{Equation 5}$$

A warning!

Once you have entered the probabilities elicited from your stakeholders into a CPT you should take care in subsequently using the automatic learning algorithms supplied by Netica and Hugin to improve them. For more details see the relevant section in Appendix 1.

Extensions of the general approach — completing an EPT

When a child has more than three non-modifying parents

When the BN is structured in this way, simply extend Table 1 by including further columns on the left. You will also need to include further rows so you can show the situation described by each of the new parents being in its negative state when all the other parents are in their positive states. Consequently, you will have one additional question for each new parent added to the general approach described above.

When the parents have more than two discrete states

Table 4 shows the case where one of the parents has three discrete states while all the others still have two. Similar tables can be constructed for cases where the other parents have different numbers of states. The important thing is to make sure you ask a question for each situation described by a parent node being in each of its states while the other parents remain in their positive states.

NMP1 has been given states called A, B and C. It is important that state A is the state most likely to cause the child node to be in its success state and state C is the least likely. Therefore, state A can still be thought of as the positive state.

Ask questions 1 to 6 in the same way as in the general approach.

Table 4: Elicited probability table (EPT) when a parent has more than three discrete states

	Non-modifying parent 1 (NMP1)	Non-modifying parent 2 (NMP2)	Non-modifying parent 3 (NMP3)	Child state and score out of ten
Question 1:	Positive state	Positive state	State A	**To be elicited**
Question 2:	Negative state	Negative state	State C	**To be elicited**
Question 3:	Negative state	Positive state	State A	**To be elicited**
Question 4:	Positive state	Negative state	State A	**To be elicited**
Question 5:	Positive state	Positive state	State B	**To be elicited**
Question 6:	Positive state	Positive state	State C	**To be elicited**

When the child has more than two discrete states

If the general case is true with the exception that the child node has more than two discrete states, then Table 1 can be used. In this case, however, some additional questions must be asked.

For question 1, ask each stakeholder to imagine a situation such as that described by the states of the parents on the first row. Then ask them which state they believe the child will be in given this situation. Noting down their answer in the box marked “State and elicited probability 1”, in Table 1, ask them to imagine that the situation described by the states of the parents were to last for ten years. Having done this, ask them to estimate for how many of the ten years that they are imagining would the child be in the state they say. Write this into your table too. If a stakeholder does not think that the child will be in the same state for all ten years, then you should find out the reasons why he or she thinks this, if time permits.

So far this is much the same as for the general case. If the stakeholders have said they believe that the child will be in a particular state for all ten years, then there is no need to continue. If this is not the case, then ask for how many of the remaining years the child will be in its other states. Again you should try to find out the stakeholder’s reasoning.

Questions 2 to 5 can then be asked in exactly the same way, the difference being that a new situation (to be imagined by the stakeholder) is described each time by the states of the parents.

When the child has continuous states

When this is the case, you should split the full range of values the child node can take into two or three sub-ranges (you will have agreed what this range should be, with the stakeholders, in Step 5 of Chapter 4). Do this before meeting the stakeholders but make sure you use the same sub-ranges for all stakeholders. Decide on the sub-range boundaries by thinking about threshold values of the child node that have practical relevance to you. Earning above a particular level, for example, may allow a farmer to buy fertiliser.

Each sub-range can now be thought of as being a discrete state, so if you have two sub-ranges you can follow the guidelines for the general situation and, if you have three, you can follow the guidelines for “When the child has more than two discrete states”. However, the questions should be asked in a slightly different way. This is because it is best to influence the answer given to you by the stakeholder as little as possible. So do not tell the stakeholder the sub-range boundaries you have chosen. Instead, for each situation you are asking them to imagine (following the general approach), ask them to choose a single value from anywhere within the full range the node. Note which of your sub-ranges this falls into, but do not tell the stakeholder. In the discussion below, this will be called the “original” value.

Next, as with the general approach, ask each stakeholder to imagine the situation lasting for ten years and find out if they expect the original value hold for all ten years. If they think it will change, then ask for how many years this child node will have a greater value and the greatest value that the child will attain in those ten years. If the greatest value is in a higher sub-range than the original value, you will need to find out how many years the stakeholder believes the child node will be in that higher sub-range. You will have to do this by asking them to give a value for every year in which they believe the child node will have a value greater than first one.

When you have done this, repeat this procedure for the years when the stakeholder thinks the value will be smaller than the original one. When you have finished, you should have a count of the number of years, out of the 10, for which the stakeholder believes the child node will be in each of the sub-ranges that you have given it.

When some of the parents have continuous states

In this case, you should draw up a table similar to Table 1, if the continuous parent you are considering has two sub-ranges, or Table 4, if the continuous parent you are considering has three sub-ranges. In the two sub-range case, the positive state in Table 1 should be replaced by the sub-range of the parent that is most likely to lead to the success state in the child. In the three sub-range case, state A should be replaced by the sub-range of the parent which is most likely to lead to the success state in the child and state C should be replace by that which is least likely to lead to the success state.

Of course, this means that you will now have to tell the stakeholder the sub-ranges you have chosen for the parent. This should not influence them unduly, since you should already have elicited the probabilities for the parent’s EPT.

When there is one modifying parent [3]

In this case, draw up a table like Table 5, which assumes that there are two non-modifying parents. If there are more, simply follow the instructions in "When a child has more than three non-modifying parents". It also assumes that the modifying parent has two states. If it has more than this, simply add an extra column to the right of the table — no extra rows will be required.

Table 5: Elicited probability table (EPT) when there is one modifying parent

	Non-modifying parent 1 (NMP1)	Non-modifying parent 2 (NMP2)	Child state and score out of ten whenmodifying parent is in State A	State B
Question 1:	Positive state	Positive state	**Elicited state, 1a**	**Elicited state, 1b**
Question 2:	Negative state	Negative state	**Elicited state, 1a**	**Elicited state, 1b**
Question 3:	Negative state	Positive state	**Elicited state, 1a**	**Elicited state, 1b**
Question 4:	Positive state	Negative state	**Elicited state, 1a**	**Elicited state, 1b**

The form of questioning is exactly the same as that used in the general approach, except that the whole procedure needs to be carried out twice, once for when the modifying parent is in state A and once for when it is in state B. Ask questions 1a to 4a first and then ask questions 1b to 4b.

Effectively, there are now two different scenarios for each question. Taking question 1 as an example, the first scenario side is described by NMP1 in its positive state, NMP2 in its positive state and the modifying parent in state A. The second scenario is described by NMP1 in its positive state, NMP2 in its positive state and the modifying parent in state B.

If the modifying parent only affects one of the non-modifying parents, then if the state of that non-modifying parent is positive, the states and probabilities elicited for the child should be the same no matter what state the modifying parent is in. For example, if the modifying parent only affects NMP1, then the states and probabilities elicited should be the same for question 4 no matter what state the modifying parent is in.

When there are two modifying parents

In this case, draw up a table similar to Table 6, where MP1 is the first modifying parent and MP2 is the second. As in the case above, if the modifying parents have more than two states or if there are more than two non-modifying parents, you can adapt the table accordingly.

Ask questions as described in "When there is one modifying parent" except that it will now be necessary to carry out the whole procedure 4 times. Ask questions 1aa to 4aa before going on to questions 1ab to 4ab.

Combinations of the above

Combinations of the above cases may arise. For example, you may have a case when there are four discrete parents, two of which have three states. When this happens, simply combine the instructions above.

[3] A *modifying parent* is one whose effect on (some of) its children is *dependent* on the states of other parents of those children (see Step 5 in Chapter 4 for more details).

Table 6: Elicited probability table (EPT) when there are two modifying parents (MPs)

	NMP1	NMP2	Child state and score out of ten when the states of MP1 and MP2 (respectively) are: A & A	A & B	B & A	B & B
Question 1:	Positive state	Positive state	**1aa**	**1ab**	**1ba**	**1bb**
Question 2:	Negative state	Negative state	**2aa**	**2ab**	**2ba**	**2bb**
Question 3:	Negative state	Positive state	**3aa**	**3ab**	**3ba**	**3bb**
Question 4:	Positive state	Negative state	**4aa**	**4ab**	**4ba**	**4bb**

Extensions of the general approach — calculating interpolation factors

For when parents have more than two states (discrete or continuous sub-ranges)

Following averaging, you will have a table similar to that shown in Table 7.

Table 7: Elicited probability table (EPT), following averaging, when one parent has three states

Non-modifying parent 1 (NMP1)	*Non-modifying parent 2 (NMP2)*	*Non-modifying parent 3 (NMP3)*	*Probability of success*	*failure*
Positive state of NMP1	Positive state of NMP2	State A of NMP3	P_{1X}	P_{1Y}
Negative state of NMP1	Negative state of NMP2	State C of NMP3	P_{12X}	P_{12Y}
Negative state of NMP1	Positive state of NMP2	State A of NMP3	P_{7X}	P_{7Y}
Positive state of NMP1	Negative state of NMP2	State A of NMP3	P_{4X}	P_{4Y}
Positive state of NMP1	Positive state of NMP2	State B of NMP3	P_{2X}	P_{2Y}
Positive state of NMP1	Positive state of NMP2	State C of NMP3	P_{3X}	P_{3Y}

The numbering refers to the CPT shown in Table 10.

The logic used to calculate the interpolation factors is the same as for the general case, except that two factors are needed for NMP3: the first accounts for the reduction in the probability of success when NMP3 changes from state A to state B; the second accounts for the change from B to C.

Interpolation factors for NMP 3:

$$IF_{3AB} = (P_{2X} - P_{12X}) / (P_{1X} - P_{12X}) \quad \text{Equation 6}$$

$$IF_{3BC} = (P_{3X} - P_{12X}) / (P_{1X} - P_{12X}) \quad \text{Equation 7}$$

Interpolation factor for NMP 2:

$$IF_{2} = (P_{4X} - P_{12X}) / (P_{1X} - P_{12X}) \quad \text{Equation 8}$$

Again, you do not need to calculate interpolation factors for the probabilities of failure.

For when children have three states (discrete or continuous sub-ranges)

After completing the averaging, you should have a table similar to Table 8.

Table 8: Elicited probability table (EPT), following averaging, for when child has more than 2 discrete states

Non-modifying parent 1 (NMP1)	*Non-modifying parent 2 (NMP2)*	*Non-modifying parent 3 (NMP3)*	*Probability that child is in state X*	*state Y*	*state Z*
Positive state	Positive state	Positive state	$\mathbf{P_{1X}}$	$\mathbf{P_{1Y}}$	$\mathbf{P_{1Z}}$
Negative state	Negative state	Negative state	$\mathbf{P_{8X}}$	$\mathbf{P_{8Y}}$	$\mathbf{P_{8Z}}$
Negative state	Positive state	Positive state	$\mathbf{P_{5Y}}$	$\mathbf{P_{5Y}}$	$\mathbf{P_{5Z}}$
Positive state	Negative state	Positive state	$\mathbf{P_{3X}}$	$\mathbf{P_{3Y}}$	$\mathbf{P_{3Z}}$
Positive state	Positive state	Negative state	$\mathbf{P_{2X}}$	$\mathbf{P_{2Y}}$	$\mathbf{P_{2Z}}$

The numbering refers to the CPT shown in Table 11.

The following scheme will only work when $P_{1X} > P_{8X}$ and $P_{8Z} > P_{1Z}$. This should be the case if the scores elicited from the stakeholders are logical. If it is not the case, you should review the stakeholder scores. Check in particular that P_{1X} and P_{1Z} are the probabilities for the parent state combination most likely to give rise to state X in the child and that P_{8X} and P_{8Z} are least likely.

In this case, interpolation factors are calculated for the first state (Ch = X) and the last state (Ch = Z) of the child.

Interpolation factor for NMP 3, for child in state X:

$$IF_{3,\ Ch = X} = (P_{2X} - P_{8X}) \ / \ (P_{1X} - P_{8X}) \qquad \text{Equation 9}$$

Interpolation factor for NMP 3, for child in state Z:

$$IF_{3,\ Ch = Z} = (P_{2Z} - P_{8Z}) \ / \ (P_{1Z} - P_{8Z}) \qquad \text{Equation 10}$$

Interpolation factor for NMP 2, for child in state X:

$$IF_{2,\ Ch = X} = (P_{3X} - P_{8X}) \ / \ (P_{1X} - P_{8X}) \qquad \text{Equation 11}$$

Interpolation factor for NMP 2, for child in state Z:

$$IF_{2,\ Ch = Z} = (P_{3Z} - P_{8Z}) \ / \ (P_{1Z} - P_{8Z}) \qquad \text{Equation 12}$$

For the same reasons as outlined in the general case, you do not need to calculate interpolation factors for state Y.

For when child has more then three states

Follow the procedure above for the most successful child state (state X, above) and the least successful child state (state Z, above). For the same reasons as outlined in the general case, you do not need to calculate interpolation factors for the states in between these two.

For modifying parents

Following averaging, you will have a table similar to Table 9.

Table 9: Elicited probability table (EPT), following averaging, for when one parent is a modifying parent

Modifying parent (MP)	*Non-modifying parent 1 (NMP1)*	*Non-modifying parent 2 (NMP2)*	*Non-modifying parent 3 (NMP3)*	*Probability of success*	*failure*
State A	Positive state	Positive state	Positive state	P_{1X}	P_{1Y}
State A	Negative state	Negative state	Negative state	P_{8X}	P_{8Y}
State A	Negative state	Positive state	Positive state	P_{5X}	P_{5Y}
State A	Positive state	Negative state	Positive state	P_{3X}	P_{3Y}
State A	Positive state	Positive state	Negative state	P_{2X}	P_{2Y}
State B	Positive state	Positive state	Positive state	P_{9X}	P_{9Y}
State B	Negative state	Negative state	Negative state	P_{16X}	P_{16Y}
State B	Negative state	Positive state	Positive state	P_{13X}	P_{13Y}
State B	Positive state	Negative state	Positive state	P_{11X}	P_{11Y}
State B	Positive state	Positive state	Negative state	P_{10X}	P_{10Y}

The numbering refers to the CPT in Table 12.

In the equations below, note that there are different interpolaion factors and minimum probabilities, depending on whether the modifying parent is in state A or state B.

Interpolation factor for NMP 3, for MP state A:

$$IF_{3,\, MP = A} = (P_{2X} - P_{8X}) / (P_{1X} - P_{8X}) \qquad \text{Equation 13}$$

Interpolation factor for NMP 3, for MP in state B:

$$IF_{3,\, MP = B} = (P_{10X} - P_{16X}) / (P_{9X} - P_{16X}) \qquad \text{Equation 14}$$

Interpolation factor for NMP 2, for MP in state A:

$$IF_{2,\, MP = A} = (P_{3X} - P_{8X}) / (P_{1X} - P_{8X}) \qquad \text{Equation 15}$$

Interpolation factor for NMP 2, for MP in state B:

$$IF_{2,\, MP = B} = (P_{11X} - P_{16X}) / (P_{9X} - P_{16X}) \qquad \text{Equation 16}$$

As before, you do not need to calculate the interpolation factors for the probabilities of failure.

Extensions of the general approach — filling in the CPTs using the interpolation factors

When parents have more than two states (either discrete or continuous sub-ranges)

Table 10: Conditional probability table (CPT) when one parent has 3 states

State combination number	State of parent 1	State of parent 2	State of parent 3	Probability that child is in state X	state Y
1	Positive	Positive	A	P_{1X} *	P_{1Y} *
2	Positive	Positive	B	P_{2X} *	P_{2Y} *
3	Positive	Positive	C	P_{3X} *	P_{3Y} *
4	Positive	Negative	A	P_{4X} *	P_{4Y} *
5	Positive	Negative	B	**P_{5X} +**	**P_{5Y} +**
6	Positive	Negative	C	**P_{6X} +**	**P_{6Y} +**
7	Negative	Positive	A	P_{7X} *	P_{7Y} *
8	Negative	Positive	B	**P_{8X} +**	**P_{8Y} +**
9	Negative	Positive	C	**P_{9X} +**	**P_{9Y} +**
10	Negative	Negative	A	**P_{10X} +**	**P_{10Y} +**
11	Negative	Negative	B	**P_{11X} +**	**P_{11Y} +**
12	Negative	Negative	C	P_{12X} *	P_{12Y} *

*elicited from stakeholders **+not elicited**

The interpolation factors below refer to those in Equations 6, 7 and 8.

$P_{5X} = [(P_{4X} - P_{12X}) \times IF_{3AB}] + P_{12X}$ and $P_{5Y} = 100 - P_{5X}$ Equation 17

$P_{6X} = [(P_{4X} - P_{12X}) \times IF_{3BC}] + P_{12X}$ and $P_{6Y} = 100 - P_{6X}$ Equation 18

$P_{8X} = [(P_{7X} - P_{12X}) \times IF_{3AB}] + P_{12X}$ and $P_{8Y} = 100 - P_{8X}$ Equation 19

$P_{9X} = [(P_{7X} - P_{12X}) \times IF_{3BC}] + P_{12X}$ and $P_{9Y} = 100 - P_{9X}$ Equation 20

$P_{10X} = [(P_{7X} - P_{12X}) \times IF_{2}] + P_{12X}$ and $P_{10Y} = 100 - P_{10X}$ Equation 21

$P_{11X} = [(P_{10X} - P_{12X}) \times IF_{3AB}] + P_{12X}$ and $P_{11Y} = 100 - P_{11X}$ Equation 22

Note that P_{11X} is calculated directly from a previously derived value (P_{10X}) and not directly from a stakeholder elicited probability. This becomes more common the larger a CPT becomes.

When a child has three states

Table 11: Conditional probability table (CPT) for a child with three states

State combination number	*State of parent 1*	*State of parent 2*	*State of parent 3*	*Probability that child is in state X*	*state Y*	*state Z*
1	Positive	Positive	Positive	P_{1X}*	P_{1Y}*	P_{1Z}*
2	Positive	Positive	Negative	P_{2X}*	P_{2Y}*	P_{2Z}*
3	Positive	Negative	Positive	P_{3X}*	P_{3Y}*	P_{3Z}*
4	Positive	Negative	Negative	**P_{4X}+**	**P_{4Y}+**	**P_{4Z}+**
5	Negative	Positive	Positive	P_{5X}*	P_{5Y}*	P_{5Z}*
6	Negative	Positive	Negative	**P_{6X}+**	**P_{6Y}+**	**P_{6Z}+**
7	Negative	Negative	Positive	**P_{7X}+**	**P_{7Y}+**	**P_{7Z}+**
8	Negative	Negative	Negative	P_{8X}*	P_{8Y}*	P_{8Z}*

*elicited from stakeholders **+not elicited**

The interpolation factors below refer to those in Equations 9, 10, 11 and 12:

$P_{4X} = [(P_{3X} - P_{8X}) \times IF_{3,\ Ch = X}] + P_{8X}$ Equation 23

$P_{6X} = [(P_{5X} - P_{8X}) \times IF_{3,\ Ch = X}] + P_{8X}$ Equation 24

$P_{7X} = [(P_{5X} - P_{8X}) \times IF_{2,\ Ch = X}] + P_{8X}$ Equation 25

$P_{4Z} = [(P_{3Z} - P_{8Z}) \times IF_{3,\ Ch = Z}] + P_{8Z}$ Equation 26

$P_{6Z} = [(P_{5Z} - P_{8Z}) \times IF_{3,\ Ch = Z}] + P_{8Z}$ Equation 27

$P_{7Z} = [(P_{5Z} - P_{8Z}) \times IF_{2,\ Ch = Z}] + P_{8Z}$ Equation 28

$P_{4Y} = 100 - (P_{4X} + P_{4Z})$ Equation 29

$P_{6Y} = 100 - (P_{6X} + P_{6Z})$ Equation 30

$P_{7Y} = 100 - (P_{7X} + P_{7Z})$ Equation 31

When a child has more then three states

The same procedure as above should be applied to the first and last states of the child. However, there are no formal rules to determine the probabilities that should be given to the states in the middle. Instead, study the probabilities elicited from the stakeholders and try to replicate the trends they suggest in the state combinations which do not have elicited probabilities. Make sure that the probabilities given to the child states, for any parent state combination, add up to 100.

With modifying parents

The interpolation factors below refer to those in Equations 13, 14, 15 and 16:

$P_{4X} = [(P_{3X} - P_{8X}) \times IF_{3,\ MP = A}] + P_{8X}$ and $P_{4Y} = 1 - P_{4X}$ Equation 32

$P_{6X} = [(P_{5X} - P_{8X}) \times IF_{3,\ MP = A}] + P_{8X}$ and $P_{6Y} = 1 - P_{6X}$ Equation 33

$P_{7X} = [(P_{5X} - P_{8X}) \times IF_{2,\ MP = A}] + P_{8X}$ and $P_{7Y} = 1 - P_{7X}$ Equation 34

$$P_{12X} = [(P_{11X} - P_{16X}) \times IF_{3, MP = B}] + P_{16X} \text{ and } P_{14Y} = 1 - P_{14X} \quad \text{Equation 36}$$

$$P_{14X} = [(P_{13X} - P_{16X}) \times IF_{3, MP = B}] + P_{16X} \text{ and } P_{15Y} = 1 - P_{15X} \quad \text{Equation 38}$$

$$P_{15X} = [(P_{13X} - P_{16X}) \times IF_{2, MP = B}] + P_{16X} \text{ and } P_{15Y} = 1 - P_{15X} \quad \text{Equation 38}$$

Table 12: Conditional probability table (CPT) with one modifying parent

State combination number	*State of modifying parent (MP)*	*NMP1*	*NMP2*	*NMP3*	*Probability that child is in state X*	*state Y*
1	A	Positive	Positive	Positive	P_{1X}*	P_{1Y}*
2	A	Positive	Positive	Negative	P_{2X}*	P_{2Y}*
3	A	Positive	Negative	Positive	P_{3X}*	P_{3Y}*
4	A	Positive	Negative	Negative	**P_{4X}+**	**P_{4Y}+**
5	A	Negative	Positive	Positive	P_{5X}*	P_{5Y}*
6	A	Negative	Positive	Negative	**P_{6X}+**	**P_{6Y}+**
7	A	Negative	Negative	Positive	**P_{7X}+**	**P_{7Y}+**
8	A	Negative	Negative	Negative	P_{8X}*	P_{8Y}*
9	B	Positive	Positive	Positive	P_{9X}*	P_{9Y}*
10	B	Positive	Positive	Negative	P_{10X}*	P_{10Y}*
11	B	Positive	Negative	Positive	P_{11X}*	P_{11Y}*
12	B	Positive	Negative	Negative	**P_{12X}+**	**P_{12Y}+**
13	B	Negative	Positive	Positive	P_{13X}*	P_{13Y}*
14	B	Negative	Positive	Negative	**P_{14X}+**	**P_{14Y}+**
15	B	Negative	Negative	Positive	**P_{15X}+**	**P_{15Y}+**
16	B	Negative	Negative	Negative	P_{16X}*	P_{16Y}*

*elicited from stakeholders **+not elicited**

Appendix 3: Further reading

Books

The best introduction to Bayesian networks is:

F.V. Jensen, 1996. An introduction to Bayesian networks. UCL press, London, ISBN 185728332-5, 178pp.

For a more technical discussion on Bayesian networks:

R. G. Cowell, A. P. Dawid, S. L. Lauritzen and D. J. Spiegelhalter, 1999. Probabilistic Networks and Expert Systems. Springer-Verlag, New York, ISBN 0387987673, 370pp.

For a general introduction to Bayesian statistics:

E. Lloyd, 1984. Handbook of applicable mathematics. Vol. VI: Statistics. Part A. John Wiley, Chichester. ISBN 0471902721, 498pp.

For information on management science and decision support approaches:

G.M. Marakas, 1999. Decision support systems in the 21st century. Prentice Hall, New Jersey. ISBN 013744186, 506pp.

M. Pidd, 1996. Tools for thinking: modelling in management science. John Wiley and Sons, New York. ISBN 0471964557, 350pp.

Scientific journals

These papers relate to examples of the application of Bayesian networks to a range of decision problems:

Anderson, J. L. 1998. Embracing uncertainty: The interface of Bayesian statistics and cognitive psychology. *Conservation Ecology* **2** (1), 2. Available at: ***http://www.consecol.org/journal/vol2/iss1/art2***

Bacon, P.J., Cain, J.D. and Howard, D.C. 2001. Belief network models of landowner decisions and land-use change. *J. Environ. Manage.* (in press).

Batchelor, C.H. and Cain, J.D. 1998. Application of belief networks to water management studies. *Agric.Wat. Manage.* **40**, 51-57.

Cain, J.D., Jinapala, K., Makin, I.W., Somaratna, P.G., Ariyaratna, B.R. and Perera, L.R. 2001. Participatory decision support for agricultural management. A case study from Sri Lanka. *Agric. Systems* (in press).

Cain, J.D., Batchelor, C.H. and Waughray, D.K.N. 1999. Belief networks: a framework for the participatory development of natural resource management strategies. *Environ., Devel. & Sustainability* **1**, 123-133.

Cain, J.D., Moriarty, P.B. and Lovell, C.J. 1999. Holistic development of natural resources in Zimbabwe: Constructing a Bayesian Belief Network for integrated management. *2nd Inter-Regional Conf. on Environment-Water - Emerging Technologies for Sustainable Land Use and Water*

Management, Lausanne. Presses Polytech. et Universitaires Romandes.

Chong, H.G. and Walley, W.J. 1996. Rule-based versus probabilistic approaches to the diagnosis of faults in wastewater treatment processes. *Artificial Intelligence in Engineering* **1**, 265-273.

Spiegelhalter, D.J., Dawid, A.P., Lauritzen, S.L. and Cowell, R.G. 1993. Bayesian analysis in expert systems. *Statist. Sci.* **8**, 219-283.

Stassopoulou, A., Petrou, M. and Kittler, J. 1998. Application of a Bayesian network in a GIS-based decision making system. *Int. J. Geogr. Information Sci.* **12**, 23-45.

Varis, O. 1997. Bayesian decision analysis for environmental and resource management. *Environ. Modelling & Software* **12**, 177-185.

Varis, O. 1995. Belief networks for modelling and assessment of environmental change. *Environmetrics* **6**, 439-444.

Varis, O. and Kuikka, S. 1997. Joint use of multiple environmental assessment models by a Bayesian meta-model: the Baltic salmon case. *Ecol. Modelling* **102**, 341-351.

Varis, O. and Kuikka, S. 1997. BENE_EIA: A Bayesian approach to expert judgment elicitation with case studies on climate change impacts on surface waters. *Climatic change* **37**, 539-563.

Articles on Bayesian statistics

The Economist, 30th September 2000, p. 58: "In praise of Bayes".

Prospect, November 1998, p. 20: "Flukes and flaws".

Science, 19th November 1999, vol. 286, p. 1460: "Bayes offers a 'new' way to make sense of numbers".

Science, 24th December 1999, vol. 286, p. 2449: "The Bayesian way".

The Sunday Times, 18th April 1999, p. 6: "Fancy maths gives banks your number".

Useful web sites

http://www.nwl.ac.uk/research/Bayesnet/ for the web page associated with these guidelines

http://www.cs.berkeley.edu/~murphyk/Bayes/bayes.html for a good overview of Bayesian Networks

http://www.norsys.com/ for an introduction to BNs and Netica software

http://www.hugin.com/ for an introduction to BNs and Hugin software

http://b-course.cs.helsinki.fi/ for an interactive tutorial on Bayesian modelling

http://www.spiritone.com/~brucem/bbns.htm for guidelines on creating Bayesian network models in ecology

http://www.research.microsoft.com/dtas/ for the Microsoft web page on decision theory

http://dssresources.com/ for general information on decision support systems

Appendix 4: What's on the CD-ROM?

The CD included with this booklet contains:

- CPT calculator, a computer program to help generate conditional probability tables according to the method described in Appendix 2;
- A demonstration version of Hugin, a commercially available software package for the construction and analysis of Bayesian networks;
- A demonstration version of Netica, also a commercially available software package for the construction and analysis of Bayesian networks;
- An example Bayesian network, to be run in Netica;
- Documentation related to the above and including an electronic copy of the entire guidelines text in portable document format (pdf).

It is strongly recommended that you read the information in Appendix 2 before using CPT calculator. It is important that you understand how it generates conditional probabilities. Instructions for its use are provided below. CPT calculator has been written so that its outputs can be cut and pasted directly into Netica. This cannot be done in Hugin, although the values produced by CPT calculator can still be entered by hand.

How to use CPT calculator

1. To run the program, click on the Start button, select the Programs menu and then the CPT calculator program group. Left click on "CPT calculator".

2. You can only calculate one CPT at a time. Choose the variable in your BN whose CPT you wish to calculate then click in the child variable box in CPT calculator and type in its name. Click on the "Enter child name" button and then acknowledge the message box that appears (click on OK).

3. Click on the arrow to the right of the box labelled "Number of states" to reveal a pull-down menu. From this, select the number of states that your chosen variable has by clicking on it. *In this version of CPT calculator you may have only up to three states. This follows the advice given in the guidelines about limiting the number of states you give to a variable.*

4. Click on the "Edit state names" button and enter the names of the variable's states into the boxes provided. Make sure that the first name you enter is the state that you consider to be the most positive and that the last is the most negative (see Appendix 2). Click "OK" when you've finished.

5. Go back to your BN and count the number of parents that your chosen variable has. If the variable has 0 or 1 parent, you do not need to use CPT calculator. If it has more than 2 parents, then click on the "Add parent to end of table" button in CPT calculator until the correct number of parents are displayed in the first column of the table to your left (in yellow).

6. Select "Parent 1" from the pull-down menu underneath the label "Select parent to edit". Click on the "Edit parent" button and enter the requested information. See Appendix 2 for a definition of non-modifying and modifying parents. Again, make sure that the first state name you enter is the most positive and the last the most negative. Click "OK" when you've finished.

7. Repeat the above step for each parent variable, selecting each one in turn from the pull-down menu underneath the label "Select parent to edit".

8. When you have entered all the required information about the parent variables, click on the "Edit EPT" button. An elicited probability table will be displayed (see Appendix 2):

- The names of each of the non-modifying parents (NMP) will appear in grey boxes at the head of the columns to the left (one for each NMP). The states of the NMPs are listed under the names in particular combinations.
- Starting above the rightmost NMP column are some lines of text. The first line (starting at the bottom) contains the name of the child variable, then a colon, followed by its state names. This line is always present. Any lines above this contain the names of any modifying parents (MPs), followed by a colon, together with their state names. If there are no MPs, then there will only be one line of text (the one with the child variable details).
- Underneath the child variable details and to the right of the NMP columns, lies a grid of empty boxes, also arranged in columns. Each column relates to the combination of child variable and MP states (if any) indicated by the state names in the lines of text above each column.

9. Enter the probabilities you have elicited from your stakeholders into the empty boxes. The click on the "Calculate CPT" button. A message box will appear asking if you want to save the CPT. Usually you will answer "Yes".

10. If you answer "Yes" then a further message box will ask you whether you want to save the parent state names with the conditional probabilities. If you intend to look at the values produced by CPT calculator before copying it into Netica, then you should click "Yes". If you intend to copy the values produced by CPT calculator directly into Netica, then you should click "No". Enter a name for the file in the "File name" box and click on "Save". The outputs from CPT calculator are saved as a text file.

11. If you have saved the file with parent state names first but now wish to re-save it without them, then click on the "Save CPT" button and follow the instructions above.

12. If you wish to view the file you have just saved, click on the "View CPT file" button, enter the name of the file you wish to view in the "File name" box and click on the "Open" button. Click the "OK" button when you have finished looking at the file.

13. To copy the calculated CPT values into Netica, save the file without parent state names and view it as described above. Highlight all the values at one time (click to the left of the first value, hold down the left mouse button and drag the cursor to the right of the last value) and right-click on the mouse and select "Copy" from the menu that will appear. Next, go into Netica and open the CPT. Highlight the whole CPT by choosing "Select all" from the "Edit" menu and then paste in the values you copied from CPT calculator (choose "Paste" from the "Edit" menu).

14. When you have finished, return to the main form and calculate the next CPT in your BN.