TOWARDS THE USE OF BAYESIAN NETWORKS FOR DECISION SUPPORT IN WATERSHED MANAGEMENT: A CASE STUDY FROM THE PERU SIERRA

Arancibia, A.* and Moriarty, P.**

*National University of Engineering. Calle Michel Fort 353 – Dpto 302 Urb.Ingenieria. LIMA 31, PERU. E-mail: aranc@ije.nl
**IRC International Water and Sanitation Centre. PO Box 2869, Delft, 2601. The Netherlands. E-mail:moriarty@irc.nl

ABSTRACT
Decision making in watershed management, as in other approaches that seek sustainable development through improved stakeholder led natural resource management, requires managers to make decisions based on a wide range of different factors ranging from the physical to the social and economic. These decisions are frequently hampered by lack of data, and great uncertainty related to the interplay of the different factors.

To support them in this task, managers and decision makers are increasingly turning to computer based decision support tools to aid them in these difficult conditions. One approach that is promising, particularly for areas where data is scarce and uncertainty high are Bayesian Networks.

This paper reports an exercise in using a Bayesian network approach to analyse a watershed management programme in Peru’s Sierra mountains, and draws conclusions about both the utility of the tool, and the approach being adopted to watershed management.

KEY WORDS
Bayesian Networks, Decision support systems, Peru, Watershed management.

INTRODUCTION
Watershed management (WM), along with other area based management approaches such as integrated catchment management (ICM), is becoming increasingly widely adopted all inclusive strategy for land and water management and sustainable rural development. In Peru, watershed management is still relatively new and is practiced principally in the mountainous regions called the Sierra. It is focused primarily on erosion control and poverty alleviation, under the hypothesis that good natural resource management is a pre-requisite for increased productivity and consequently a sustainable decrease in poverty. In this way watershed management projects in Peru are trying to emulate experiences gained over the last 25 years in poverty alleviation and natural resources conservation in many other developing countries.

To be effective, and use often limited funds and human capacity to their full potential, watershed management professionals have to integrate a wide range of social, economic, hydrological, and ecological factors and to make decisions about where to invest time, money and effort. Often the available data on which to make these decisions is either totally lacking, or has a high degree of uncertainty attached to it. In addition, the interaction of different factors – for example rain intensity, hill slope, soil characteristics and erosion - may only be poorly understood, thus introduction yet more uncertainty. Add to this the very complex relationships between physical and non physical factors, for instance soil, rainfall, grass production leading to (animal) fibre production and the use of this to raise income, and the complexity and uncertainty threaten to become overwhelming. Faced by this reality, watershed management in many cases remains more of an art than a science, with much of what is done best characterized as (hopefully) inspired guess work.

To face the challenge of making better and more objective decisions under such difficult conditions, natural resource managers are increasingly turning to decision support systems of one sort or another. Bayesian Networks are one such tool, and one that has shown itself to be particularly
suited to use in situations characterized by great uncertainty, and where it is necessary to link a wide range of multi-disciplinary data and knowledge to support decision making.

This paper describes a watershed management programme in Peru, through a study of one watershed. The tool used to carry out the analysis of the project was a Bayesian network and the study has allowed the suitability of the Bayesian network approach to be tested, as well as to carry out a critical evaluation of the approach to watershed management adopted. The work reported here is a summary of a more detailed study carried out as part of the first authors MSc thesis (Arancibia, 2003).

BACKGROUND
Description of the study area
The Peru-Sierra natural resource management and poverty alleviation programme, is being implemented by the ministry of agriculture’s national program of watershed management and soil conservation (PRONAMACHCS), with the objective of helping to alleviate the poverty of the people of the rural sierra by promoting a productive and sustainable use of natural resources, in a participatory manner (World Bank, 1996). This paper takes the Ayas watershed, one of more than 122 being treated in a world bank funded project, as a case to see how the approach is currently being applied, and how it could be strengthened.

Ayas watershed is located within the larger Muylo watershed, in the Sierra region of Peru. It has an area of approximately 3,200 ha, and a population of 122 people. It lies at an elevation of between 3,650 and 4,250 m aol, and has an average annual rainfall of 600mm/year. The population of Ayas are a Campesino community, that is to say, one consisting of indigenous Andean people who live under a system of shared land ownership and a, among other things, practice a tradition of community labour contributions.

The campesinos livelihoods are based primarily on productive activities, in particular rainfed agriculture and livestock production. Of these, rainfed agriculture (primarily production of potatoes and beans) was estimated to contribute approximately 50% of total income, which livestock in the form of milk and other dairy products from cattle, and fibre production from alpacas and vicunas contributed roughly 30% and 17% respectively (Arancibia, 2003).

Watershed management activities
Watershed management in Ayas is being carried out by Peru’s national program of watershed management and soil conservation (PRONAMACHCS) as part of a World Bank funded project (World Bank, 1996). Their interventions, consist of a bundle of physical interventions, carried out with the help of the community, and community capacity development in the form of technical training.

In Ayas PRONAMACHCS has implemented a mix of furrow infiltration in the higher parts of the catchment, to reduce runoff and to control erosion, with native grass being sown to further consolidate the process of land reclamation. In addition terraces have been improved, and lower down the catchment, a small reservoir captured fed by spring, and used to feed an irrigation system to grow grass for cattle production. Fencing has been implemented in the treated parts of the catchment to protect the grass from livestock.

The impacts of the interventions are believed to include: reduced erosion and greater and more reliable stream flow through increased infiltration. The activities have also allowed the Campesinos to add dairy farming to their livelihoods activities, and it is intended to lead to greater fibre
production by increasing the quality of grass available to vicunas and alpacas. The rainfed farming of the lower catchment has not been affected.

Bayesian Networks and Decision Support systems
Power (2002) defines decision support systems (DSS) as a ‘specific class of computerized information system that supports decision making activities. DSS are interactive computer based systems intended to help decision makers use data documents, knowledge and/or models to identify and solve problems and make decisions’

Bayesian Networks are ‘a graphical tool for building decision support systems to help make decisions under uncertain conditions, and to show uncertainty in a way that can be clearly understood’ (Cain, 2001). They are composed of three principal elements:

- A set of nodes representing key factors within a system (i.e. ‘Income’ or ‘Rainfall’). Each node can have a number of states associated with them (i.e. High, Low, 100mm, 500mm etc).
- Nodes are joined together by links representing the flow of causality between the nodes. Together the nodes and links make up a ‘causal diagram’.
- A set of conditional probabilities describes the relationships between the links. These probabilities are contained within the conditional probability tables (CPT) which lie behind each node, and which define the probability that it will be in any given state according to the combined probability of its parent nodes.

For further introduction to Bayesian Networks and the theory behind them see (www.norsys.com or www.hugin.de). Jensen (1996) also provides a solid background to both Bayesian networks, and the Bayesian probability theory that underlies them.

As has already been mentioned the use of Bayesian Networks has diverse advantages, which Batchelor and Cain (1999) summarise as follows:

- A graphical nature which, permits and supports formal discussion about system structure among stakeholders from a wide variety of backgrounds and so encourages interdisciplinary discussion and participation. This in turn facilitates the formal identification of the system variables and interactions.
- A single tool that can help to understand, the relationship between physical, social and economic factors within a system, with the relationships between variables from different contexts specified in terms that make fully transparent the underlying uncertainty associated with them.
- The approach enables expert knowledge to be incorporated into the model on the same basis as more objectively derived data.
- It is relatively simple to adapt BNs to new situations or new knowledge or data.

Taken together, these features allow the creation of a system model capable of drawing not only on mathematically expressed physical process relationships, but also subjective elements corresponding to the experience of the people who are, in many cases, an integral part of the system being modelled.

One fact about the use of BNs is that relevant mentioned is “BNs do not make decisions. They simply show the impact of any particularly action on all factors linked to it, with all the attendant uncertainties; it is left to the planner or manager to make the final choice. But with the network, the whole process of decision – making is made to be much more inclusive and transparent” Bromley (2002).
Despite these benefits the BNs have important limitations including:

- because the approach is probability based, processes cannot be modelled directly.
- uncertainty is explicit in the use of conditional probabilities. This uncertainty reflects the limitations of understanding and information but also means that important trends and differences can be obscured.
- because much of the information represented in the networks is subjective, the outputs should be viewed only as relative trends among alternatives rather than absolute numbers or true probabilities (Rieman et al, 2001).
- potentially time consuming, because it requires stakeholder consultation and data collection and collation, although this is of course true of other approaches as well.

Cain (2001) identifies two ways in which BNs can be used as a tool to build a DSS for watershed management:

- Providing a mathematically optimal decision on the basis of the information provided by the BN.
- Or, promoting and improving the understanding of the environmental system, leaving the decision makers to reach their own conclusions.

The later approach is felt to be more valid in situations where, due to lack of direct empirical data, heavy reliance is made on experience and expert opinion. In this context it is crucial to understand that the BN is nothing more than a formal structuring of the expert knowledge available to the decision maker. The BN built in this exercise represents firstly the authors beliefs and understanding based on studying the available documentation. Its development is described in the next section.

**METHODOLOGY - BUILDING THE BN**

The methodology used in the study was based on the general guidelines given by Cain (2001), and involved an iterative process of developing, testing, and refining. Firstly it was necessary to develop a model of the system following the generalised outline developed by Cain (see diagram), this was then tested against knowledge gained from the real system, followed by refinement and further testing.

The first step in the process is defining the objectives. The primary objective defined by the project is poverty alleviation, through productive and sustainable use of the natural resource base. Another objective of the project is the strengthening of the community with the longer term aim of achieving greater autonomy.

These objectives were represented as nodes in the network as follows (these nodes and all the others can be seen on the network diagram, figure 1):

- **Sustainability**, in the sense of project sustainability. This is related to the level of income that will allow the community to achieve greater autonomy. The node has two states, good and bad.
- **Income** was chosen as indicator of poverty alleviation, because is the only indicator for which data is available from project reports and assessments (PRONAMACHCS, 2002). The states are based on actual data and range from extreme poverty to the maximum income if a good development of economic activities is reached.
- **Autonomy of the community** in the sense of managerial, technical and financing capacity. Autonomy to manage the watershed without the support or guidance of an external Institution or organisation. This variable has two states: high and low.
• Management capacity, capability of the community to manage by themselves the watershed in the sense of technical and administrative capacity. The states are strong or weak. The parent nodes are participation and technical capacity.

What are the principal activities in the watershed, and how do these activities influence the identified objectives? The principal interventions and intermediate factors identified are:

- Productive activities: livestock (milk production and special fibre production and agriculture: own consumption and commercialisation.
- Community strengthening component: participation of the community, assistance to the trainings.

The controlling factors, factors outside the control of the project but that will nonetheless have an impact on project outcomes that were identified include:

- rainfall, annual rainfall, the ranges are decided according to the values shown in tables 3.7 and 3.9 referred as the average (600mm/year and maximum values more than 800mm/year).
- Labour contribution, defined as the unpaid labour in activities related with the watershed management. The minimum labour contribution is of one day per week, and the maximum of 3 days per week. Then the states are defined in three ranges: low (1 day/week), medium (2 days/week), and high (3 days/week).
- Auto financing capacity, defined as being able to access credits and loans. At the beginning the idea was to link this node with the node Income, but it was not possible because this will create a loop in the network, which is not allowed.
- Beneficiaries, defined as people that get direct or indirect benefits through the project activities.
- Attendance, defined as attendance to the training courses.
- Fibre/vicuna price, representing the range of likely prices for which fibre can be sold.

To make the analysis of performance of the watershed management, the following four scenarios were defined based on two key potential interventions (see table 1):

<table>
<thead>
<tr>
<th>Variables</th>
<th>Sc1</th>
<th>Sc2</th>
<th>Sc3</th>
<th>Sc4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appropriate Land Use</td>
<td>Actual</td>
<td>Actual</td>
<td>Future</td>
<td>Future</td>
</tr>
<tr>
<td>PRONAMACHCS</td>
<td>Intervention</td>
<td>Non-Intervention</td>
<td>Intervention</td>
<td>Non-intervention</td>
</tr>
</tbody>
</table>
Firstly, the impact of PRONAMACHCS. What will the status of the system be if a) PRONAMACHCS is involved in watershed management, and b) if they are not

Secondly, what would the impact of increasing the amount of land available for agricultural activities be?

The analysis was made three different stages:

**Stage 1** the likely outcome (success in meeting objectives) under the four scenarios with all other nodes left in their current (actual) states

**Stage 2** as for stage 1, but with the key nodes of sustainability, autonomy, and management capacity set to their ‘optimal’ states

**Stage 3** with all the key objective nodes optimised. This used the ability of BNs to reason ‘backwards’ along the chain of cause and affect to identify the most likely causes (most important factors) in achieving that success

**RESULTS AND DISCUSSIONS**

Some typical outputs of the different stages of testing are shown in figures 2 and 3. Based on the analysis of Ayas watershed using the network, the following principal results were obtained

![Figure 2. Results from stage 1](image1)

![Figure 3. Results from stage 2.](image2)

**Recommendations for improved watershed management interventions**

The analysis of the Ayas case study using the Bayesian Network allowed a number of key observations and recommendations regarding the process of watershed management in Peru to be made. These include:

- **Impacts on income.** Increasing the land treated using furrow infiltration and irrigated grass will produce direct increments in the income level, with the maximum levels of income being reached only if there is enough rainfall. Income is very dependent of rainfall; more so than on the area of land used for the productive activities, or the interventions of PRONAMACHCS. The implication of this is that water harvesting activities through soil and water harvesting structures is the key activity in watershed management in order to increase the income level.

- **The importance of achieving autonomy.** Autonomy can only be achieved if the community manages to sustainable increase incomes, and hence improve their capacity to auto finance key interventions. Yet, current management capacity is heavily reliant on PRONAMACHCS, particularly now that added technical complexity has been introduced in the form of the sprinkler irrigation system. The technical training provided by PRONAMACHCS is critical, but the...
danger is that once the project ends access to technical support will disappear, and the capacity of the community be diminished

- **An improved monitoring system to underpin more effective interventions.** Currently the PRONAMACHCS project, like many others, monitors almost exclusively on expenditure and activities (how many workshops held, how many hectares of land treated), and not on impact (income levels, erosion before and after the project etc). This is a major flaw, which means that the focus of the project is fatally shifted from poverty reduction to efficient project implementation. In addition, there is a lack of clarity as to the understanding of the watershed systems that lies behind decisions as to where to prioritise investments and activities, with the links between activities to reduce environmental degradation and poverty not clearly spelled out.

**CONCLUSIONS**

In general, the Bayesian Network software was a simple and effective tool for analysing a watershed management programme. Even this relatively simple and conceptual level of analysis permitted important insights to be gained into how watershed management is being approached in the Sierra, and to identify some of the potential threats and opportunities related to it.

However, to use Bayesian networks as a true decision support system tool in watershed management would require a monitoring system more attuned to collecting critical information, as well as a desire on the behalf of project implementers to start making decisions in a more transparent manner. In particular critical indicators of effective implementation (qualitative and quantitative) would need to be collected, such as income levels, degree of satisfaction with the project, intention to maintain structures etc.

Bayesian Networks can improve decision making in complex natural resource management projects. They can act as a platform for stakeholder involvement by allowing the views and beliefs of the stakeholders to be included and modelled. They allow uncertainty to be recognised and dealt with explicitly, and for those involved in decision making to work out their chain of reasoning – how does reduced erosion turn into higher incomes. However, for BNs or any other DSS tools to be taken up in a wider way, a fundamental shift in the approach to implementing NR projects will first be required. One that puts effectiveness over efficiency, and that values transparency and stakeholder input to decision makings. Until this change is made, until that is there is a genuine desire in implementing agencies to make better decisions the role for DSS tools will remain marginal.

**REFERENCES**


PRONAMACHCS. (2002). Evaluación de casos exitosos en el ámbito del Proyecto de Manejo de Recursos Naturales para el Alivio de la Pobreza en la Sierra BIRF – 4130PE. MINAG, Lima, Peru.