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

Comparison of two modelling approaches for an integrated crop economic model

Samranpong, C. [1] **and C. Pollino** [1]

[1] *Integrated Catchment Assessment and Management (iCAM) Centre, The Fenner School of Environment and Society, The Australian National University, Australian Capital Territory*
Email: chalermpol.samranpong@anu.edu.au

Abstract: Integrated crop models with biophysical and economic component models were developed in order to support decisions on agricultural land productivity assessments in northern Thailand. Over the past few decades, efforts to produce higher crop yield focussed on extending crop land to increase crop yield per capita. This process included considerations of land and water quality, improved land and water use efficiency and greater involvement of farmer communities in the planning process. In order to support the planning process, decision makers needed a tool to assist in assessing the complex production system, focusing on the beginning of the cropping production process to leaving the farm. Such a tool should also support the dynamic assessment of economic land suitability for the 19 major economic crop types used in northern Thailand.

The Land Development Department (LDD) in Thailand framework provides a guide to the suitability of different crop types to a range of land quality attributes. Two modelling approaches were used to develop an integrated crop-economic model based on the framework of the LDD. The first model type is a mechanistic model, which estimates crop yields using soil and climate information and estimates economic returns. To introduce uncertainty into the model, fuzzy sets were used. The second approach used a Bayesian network model. The Bayesian network was used to estimate the probability of achieving a crop yield given climate and soil input data. Economic returns in the Bayesian network were estimated using utilities. The LDD framework is used in the models to estimate crop yield and economic returns using available biophysical and economic information.

This paper will introduce both models and assess the constraints that influenced the construction of the models. A set of criteria will be used to evaluate the models in order to examine their usefulness, representativeness and robustness. The comparison of models will focus on the: data type, technique, and model outcomes. The results of this comparison will help in evaluating the strengths and weaknesses of each of the modelling approaches, and based on these outcomes, recommendations on methods for building crop-economic models will be made.

Keywords: *Crop economic model, Decision Support System (DSS), Fuzzy sets, Bayesian Networks (BNs), northern Thailand*

1. INTRODUCTION

Integrated crop models, with biophysical and economic component models, were developed in order to support decisions on agricultural land productivity assessments in northern Thailand. For a land productivity assessment, it is important to simulate the regional crop yields and economic returns from a range of available information and knowledge. This assessment needs to include the wide range of crop varieties within the region and to assess the entire production system, focusing on the beginning of the cropping production process to leaving the farm. Uncertain parameters are present in both the crop yield and economic models. Crop yield is an uncertain parameter due to environmental factors. Economic return is highly uncertain given the rapid fluctuation of economic situation, such as crop price and production cost. Such uncertainties need to be communicated to decision makers in order to better inform decisions.

For northern Thailand, there is the need for a crop model that can support a wide range of crop varieties, can assess complex production systems and can consider the farming process from the beginning of the cropping production process to leaving the farm. Crop modelling requires detailed knowledge on many variables including climate, soil, land cover, crop requirements, water availability and on-farm management practices. Much of this data is available as spatial information, which can be accessed using a Geographic Information System (GIS). Modern GIS technology enables fast and efficient integration of knowledge and easier interaction of that information with models and decision support tools. The non-spatial information such as crop management, crop price, and production cost, can be prepared separately in database management systems (DBMS) and linked to spatial information.

Although GIS technology has advanced considerably in recent years, limitations include errors in location and features associated geometry (detail inFoody & Atkinson, 2002). Spatial variation in geographic information is one of the main sources of uncertainty in integrated crop economic models. Spatial soil property data (provided by Land Development Department (LDD) in Thailand) is a source of uncertainty for predicting crop yields. Uncertainties are found in both the boundary and attribute information. Consequently, when using GIS databases in modelling, it is important to incorporate uncertainties from GIS data, as well as modelling uncertainties (e.g. data, knowledge, and representation) in the assessment.

To deal with the uncertainties associated with agricultural land productivity assessment models, two types of approaches are trialled and discussed in this paper. The first approach uses fuzzy logic, the second uses Bayesian networks. Fuzzy logic is a mathematical approach for dealing with complex systems where only approximate information on components and connections are available (Berkes & Berkes, 2009). Fuzzy sets are used to represent the imprecise nature of information and convert this into simple mathematical expressions that can be manipulated to make mathematical inferences. A simple crop model using fuzzy approach requiring soil properties and weather data during crop growing period was developed previously (Samranpong & Ekasingh, 2003).

Bayesian networks (BNs) are graphical models that use probabilities to represent the strength of causal linkages. Uncertainties between variables are expressed using probability distributions. BNs can be used to assess optimal decisions and are often used to promote an improved understanding of environmental systems. As BNs are diagrammatically based, it is easy to build the model, understand the outputs provided, and to adapt models to other systems.

The two modelling approaches were used to develop an integrated crop-economic model based on the framework of the LDD. The framework provides a guide to the suitability of different crop types to a range of land quality attributes. The models aim to estimate crop yield and economic returns using available biophysical and economic information.

2. THE MEA THA CATCHMENT

The Mae Tha catchment (Figure 1) is located in the east of the Ping River basin, Thailand. The catchment has an area of 1,198 km^2. This area is settled by 98 villages consisting of 14,500 farmer households.

Approximately 84% of the Mae Tha catchment area is in natural forest zones with 74% forest area remaining. In agricultural lands, orchards occupy approximately 9.4% of the catchment area, followed by rice paddies at 6.2%. Rice is predominantly used for subsistence purposes. The remaining area is used for cash crops, urban and water sources.

Agricultural land has been expanding along with illegal deforestation, reflecting the land scarcity in upstream areas and the expansion of urban areas. Due to increasing trends towards monoculture and ongoing fluctuations in the economy, farmers are increasingly considered to be at greater economic risk.

Figure 1. The Mae Tha catchment and its location in Thailand

3. MODEL DEVELOPMENT

3.1. Overview

Two types of crop models, a mechanistic-fuzzy logic system and a Bayesian network, were developed to simulate crop yield and economic return. The model types were selected by the authors based on three criteria: knowledge integration, uncertainty representation, and model-GIS linkage. The first model (Figure 2) estimates crop yields using soil and climate information, and estimates economic returns to the farmer considering the whole production cycle. To introduce uncertainty into the model, fuzzy sets were used. The second model type (Figure 3) was used to estimate the probability of achieving a crop yield given climate and soil input data. Economic returns in the Bayesian network were estimated using utilities. Both models were constructed using and modifying model parameters from Tansiri and Saifak (1999).

In this paper, a polygon for land mapping unit (LMU) is used. An LMU is a unique geographic unit, which considers areas that contain similar patterns of soils, climate, water resources, land use and also political boundaries. Models simulate crop yield and economic return using the provided information of each LMU and assign the result to it one by one. Therefore the outcomes of both models can be displayed as a map. The two model types are introduced in more detail below.

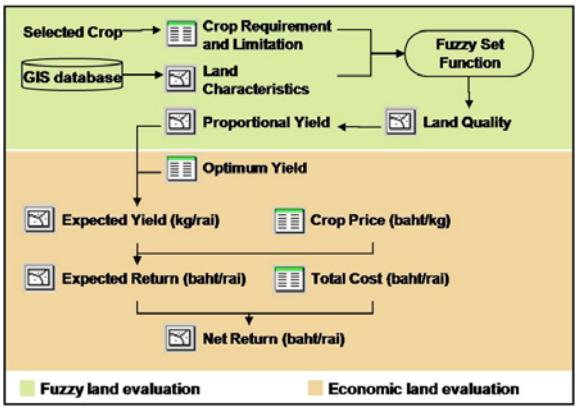

Figure 2. The fuzzy approach crop-economic model

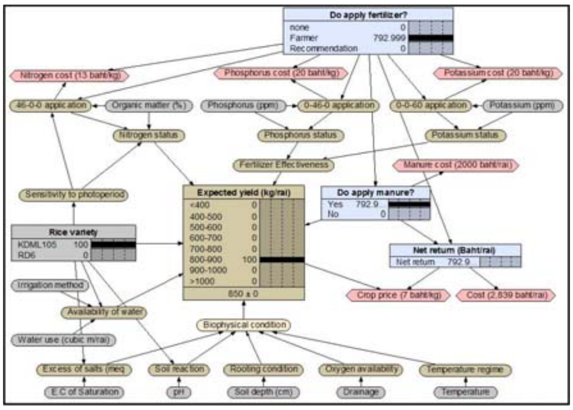

Figure 3. The probability approach crop-economic model

3.2. Fuzzy and economic land evaluation models

A mechanistic land evaluation model was developed, incorporating the fuzzy logic approach. The land evaluation model is based on the FAO framework (1976) which classifies land suitability into classes. This classification defines an exact boundary of land characteristics, while the fuzzy set permits flexibility in defining the boundary and consider a degree of closeness to the ideal point (Baja, Chapman, & Dragovich, 2002).

Eight soil properties were used as model parameters for the fuzzy land evaluation model. Some of these properties varied according to depth (e.g. pH), so it was necessary to consider soil attribute variations within soil layers. Therefore, the values of each soil attribute were calculated from the weighted value of the respective soil depth under consideration layers. An exception is made for classification data such as soil drainage, where only a topsoil value was considered.

In fuzzy models, a Standard Membership Function is used to express the membership function of a fuzzy subset whose parameters may be adjusted to fit a specific membership function in an appropriate situation. The characteristics of membership functions are being non-decreasing and having values inside 0 to 1.0 only within a bounded interval (Robinson, 2003; Sicat, Carranza, & Nidumolu, 2005). This function, which is robust for both quantitative and qualitative variables, is used to classify attributes within the land evaluation model according to vague concepts of membership. Land quality (LQ), was defined using appropriate fuzzy membership functions, which were based on literature reviews and expert recommendations and land characteristics. By using the fuzzy technique, the set of LQ and suitability indices were treated as continuous numbers range from 0 (least suitable) to 1.0 (most suitable). Membership function values of LQs considered were then combined using a convex combination function to calculate a Join Membership Function (JMF) (Baja et al., 2002).

Expected yield for a given crop was obtained by multiplying the optimum yield with proportion yield factor. The optimum yield is not a biological maximum but rather a realistically maximum attainable yield recorded in field surveys. The physical suitability index which directly affected crop yields was used as a proportional yield factor. Expected total revenues were estimated by multiplying expected yields by a given output price based on a user's scenarios. Once a crop yield and the production cost are calculated, total costs were then subtracted from total revenues giving expected net returns for each LMU. Full details of the models developed are given in Samranpong et al. (in prep).

3.3. Bayesian Networks agro-economic model

The BN agro-economic model uses probabilities to estimate crop yield by integrating knowledge and uncertainty parameters within spatial database. It is rarely possible to predict the exact yield of a crop on a given set of land characteristics and farming practices (e.g. fertilizer application); therefore, as the BN approach incorporates uncertainty, it can handle and represent such variability.

A BN model is composed of three elements; a set of nodes, links, and probabilities. In Figure 4, the nodes represent variables that are used to calculate crop yield. The links represent relationships between these nodes. Each link has direction, representing cause and effect. The probabilities, are used to calculate the belief that a node will be in a particular state given the combination of states from parent (up arrow) nodes. These are call conditional probability tables (CPTs).

Figure 4 shows an example of a BN diagram, constructed using expert knowledge, data and policy information. In the model, the phosphorus status is affected by available phosphorus in the soil (Phosphorus (ppm) node), and fertilizer application, (0-46-0 application node). The amount of fertilizer is controlled by the decision node, 'Do apply fertilizer?'. The model shows the outcome where the farmer decides to apply fertilizer at a recommended rate, as determined by the available phosphorus. The utility node, 'Phosphorus cost (20 baht/kg)', represents the economic cost of the fertiliser application. If the cost of fertilizer is 20 baht/kg, it means that phosphorus fertilizer will cost farmer 60 baht/rai (6.25 rai = 1 hectare).

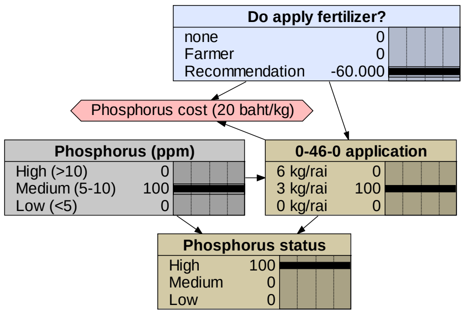

Figure 4. Bayesian network diagram of a phosphorus fertilizer application

Table 1. Probability of pH being in a certain range within a land mapping unit

Layer	Depth (cm)	pH	Probability
1	7	Strong acid	0.14
2	17	Medium acid	0.34
3	5	Neutral	0.10
4	21	Midly alkaline	0.42
Total	50		1.00

GIS variables are used as model input variables within the crop yield model structure. To consider soil attribute variation within a soil layer, GIS data is used to populate probabilities according to proportional depth (see Table 1). In the crop BN, decision nodes and utility nodes are used to calculate the expected total revenue and expected returns (Figure 4). In the complete crop-economic model (Figure 3), there is one decision node, ‘Net return (baht/rai)’ and two utility nodes, ‘Crop price (7 baht/kg)’ and ‘Cost (2,839 baht/rai)’. The model input nodes include spatial and non-spatial variables, and probabilities are assigned to states based on existing data. The spatial dataset was then used to generate a case file and this was incorporated in the crop model (EM learning function in Netica). The probability distribution for the ‘Expected yield (kg/rai)’ node and ‘Net return (Baht/rai)’ node were calculated for each case within the casefile.

4. COMPARISON OF MODELLING APPROACHES

Both crop models were designed for the simulation of crop yield and economic return. To demonstrate the use and behaviour of the model, we consider the scenario: ‘what is a net return that a farmer will receive if the rice price is 7 baht/kg and the production cost is 2,839 baht/rai’.

4.1. Crop yield estimation

Crop yield estimates from each model are shown as GIS outputs (Figures 5 and 6). The major difference between the models is that estimated yield from the fuzzy model is continuous, whereas in the BN, yields are discrete. Overall, the fuzzy approach (Figure 5) shows a lower estimated yield than the BN approach (Figure 6).

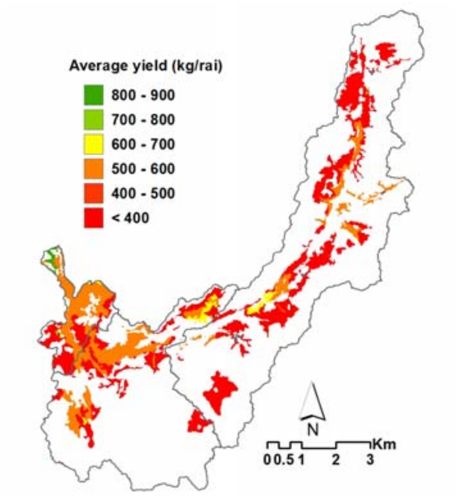

Figure 5. Estimated yield map simulated by fuzzy approach

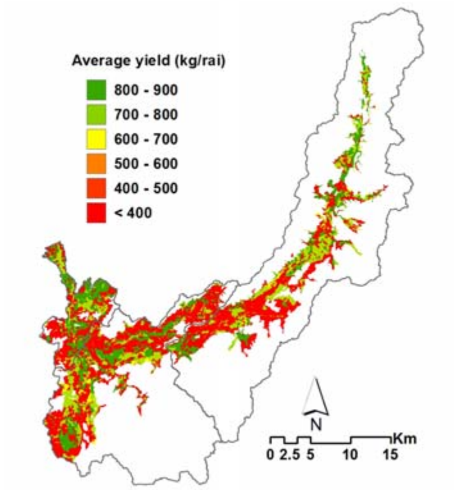

Figure 6. Estimated yield map simulated by BN approach

The comparison between model results and actual yield (average rice yield per capita (CDD, 2007) by major sub district), is presented in Table 2. The average yields estimated by BN approach are closer to the actual yield than those estimated by the fuzzy modelling approach.

The Fuzzy approach uses a mechanistic land evaluation model, which does not incorporate farming practices such as fertilizer application and disease control. In contrast, the BN model is able to consider the influences of farm management. However, the constraint in incorporating management in the model is obtaining local information on farm management practices.

Table 2. The comparison between an average rice yield in 2007 and the model predictions of the Fuzzy and Bayesian network (BNs) models for sub-districts in northern Thailand

Sub district	Average yield (kg/rai)		
	Actual yield	Fuzzy	BNs
Mueang Chi	482	423	444
Tha Sop Sao	467	445	471
Tha Kat	609	416	427
Tha Khum Ngoen	618	386	441
Tha Thung Luang	670	410	503

4.2. Economic return estimation

The models were also used to estimate the net return (Figure 7 and 8), based on predicted crop yields (Figure 5 and 6). The fuzzy approach model predicts continuous net returns using an equation, where as the BN uses decision (Net return (Baht/rai)) and utility nodes (Cost (2,839 baht/rai), Crop price (7 baht/kg)). Model predictions differed based on differences in: predicted crop yield and production cost.

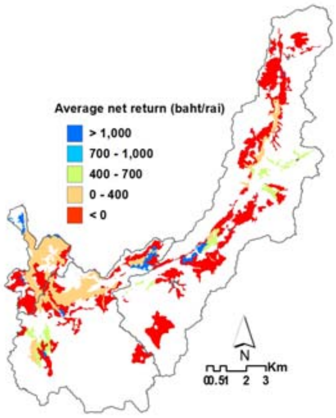

Figure 7. Estimated net return map simulated by fuzzy approach

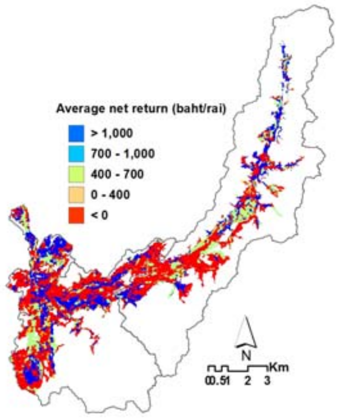

Figure 8. Estimated net return map simulated by BN approach

5. EVALUATION

In Thailand, the spatial soil data is vector-based, with many soil attributes being represented as crisp set terms such as 'very poorly drained', 'somewhat poorly drained', 'moderately well drained', and 'excessively drained'. The fuzzy approach represents both continuous data (e.g. rainfall) and ordinal data (e.g. drainage from 'very poorly drained' to 'excessively drained') and converts these attribute values into the membership grades (from 0 to 1.0), according to the class limits specified by the requirement of considering crops. In the BN, the predictions represent the likelihood of a considered soil property being a member of a boolean set.

Both models show a direct correlation between the predicted cropping areas and present paddy areas (Onpraphai et al., 2007). Given that the outcomes of the models are not only specific to a paddy field, they can be used to provide information on identifying appropriate areas for extending rice growing areas.

With the fuzzy set methodology, a set of selected land characteristics were converted to membership values individually, then integrated by using a multiplication function to produce a land suitability index. This approach requires adequate knowledge of the mechanisms relating crop responses to land characteristics. In using a multiplication function, the selection of criterion weights must be taken carefully because the weights can have a major effect on results. It is difficult to assign a sensible weight for an individual factor. Likewise, this model does not have the ability to simulate crop yield with different crop management but in reality various cultural practises and water management practices are applied by farmers in the study area.

The key advantage of the probabilistic Bayesian network approach is the ability to integrate inadequate knowledge or understanding of system processes to provide a mathematically optimal outcome (Cain, 2001). The BN can combine field data and expert knowledge from a combination of sources. The results are affected by model construction which requires stakeholder consultation and data collection and collation. However, a limitation of the BN approach is the inability to deal with continuous data efficiently (Ticehurst, Letcher, & Rissik, 2008; Uusitalo, 2007), such as where the expected yield is represented as being within discrete states.

6. CONCLUSION AND DISCUSSION

Although various integrated crop economic models have been implemented in Thailand, there is an increased focus being placed upon the need for integrated assessments which incorporate research, field, and indigenous knowledge. There are few existing methods that meet this need. GIS is a useful tool for assisting planners in organising and analysing spatial data effectively and efficiently. Combined with suitable crop and economic models, GIS can assist in land evaluation assessments.

In comparing the fuzzy set and BN modelling approaches, we found that the BN modelling approach was more accurate in its estimation of crop yields and in assessing economic return. The justification for this is as follows: using probabilities to estimate crop yield by integrating knowledge and uncertainty parameters within the spatial database (in the BN modelling approach) uses the information directly, links it together, and then uses likelihoods to estimate yields; the probabilities and model structure can be adapted as new information is obtained; and a BN approach can integrate both scientific and farmers' knowledge into the assessment process. When the BNs are linked to GIS, the model can represent changed crop yields and returns spatially, for a given scenario. Fieldwork and workshops with local farmers are essential for further study because models estimated crop yield and economic return with a given scenario of general management while farmers have a variety of cultural practices such as soil improvement, the amount of fertilizer application, and irrigation management. Those factors might affect the outcomes of the model, which can be improved. Overall, we also found that the BN modelling approach predicted the crop yield and estimates of expected profit from agricultural land more usefully and reliably than does the fuzzy approach.

ACKNOWLEDGMENTS

The PhD research is funded by the Multiple Cropping Center (MCC), Chiang Mai University, Thailand. Thanks to the MCC staff and iCAM (the Australian National University, Australia) staff to provide good service, facilities, and environment for this research.

REFERENCES

Baja, S., Chapman, D.M., & Dragovich, D. (2002). A Conceptual Model for Defining and Assessing Land Management Units Using a Fuzzy Modeling Approach in GIS Environment Environmental Management, 29(5), 647-661.

Berkes, F., & Berkes, M.K. (2009). Ecological complexity, fuzzy logic, and holism in indigenous knowledge. Futures, 41(1), 6-12.

Cain, J. (2001). Planning improvements in natural resources management. Guidelines for using Bayesian networks to support the planning and management of development programmes in the water sector and beyond. Oxon, UK: Centre for Ecology and Hydrology.

CDD. (2007). The National Rural Development 2C.

FAO. (1976). A Framework for Land Evaluation, Soils Bulletin 32. Rome, Italy: FAO Publication.

Foody, G.M., & Atkinson, P.M. (Eds.). (2002). Uncertainty in Remote Sensing and GIS: John Wiley and Sons.

Onpraphai, T., Ekasingh, B., Sangchyoswat, C., Ekasingh, M., Samranpong, C., Veerajit, V., et al. (2007). Crop choice analysis for agricultural management in catchment (Progress Report I).

Robinson, V.B. (2003). A Perspective on the Fundamentals of Fuzzy Sets and their Use in Geographic Information Systems. Transactions in GIS, 7(1), 3-30.

Samranpong, C., Ekasingh, B., & Ekasingh, M. (in prep). EconSuit: Economic Land Evaluation for Agricultural Resource Management in Northern Thailand. Environmental Modelling and Software.

Samranpong, C., & Ekasingh, M. (2003). A fuzzy land evaluation using a geographic information system. In G.-I.a.S.T.D.A. (Public) (Ed.), Surveying and Mapping for Sustainable Development. Proceedings of the National Conference on Mapping and Geo-Informatics 2003 (pp. 144-155). Bangkok, Thailand: .

Sicat, R.S., Carranza, E.J.M., & Nidumolu, U.B. (2005). Fuzzy modeling of farmers' knowledge for land suitability classification. Agricultural Systems, 83(1), 49-75.

Tansiri, B., & Srifak, K. (1999). Land Evaluation Manual. Bangkok, Thailand: Land Development Department.

Ticehurst, J.L., Letcher, R.A., & Rissik, D. (2008). Integration modelling and decision support: A case study of the Coastal Lake Assessment and Management (CLAM) Tool. Mathematics and Computers in Simulation, 78(2-3), 435-449.

Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. Ecological Modelling, 203(3-4), 312-318.