Explorations of a Bayesian Belief Network for the Simultaneous Farming of Rice and Shrimp Crops

A. Lewis[1], M. Randall[2], B. Stewart-Koster[3], N. Dieu Anh[3], M. Burford[3], J. Condon[4], N. Van Qui[5], L. Huu Hiep[6], D. Van Bay[6], J. Sammut[7]

[1]*Institute for Integrated and Intelligent Systems, Griffith University, Queensland, Australia*
a.lewis@griffith.edu.au

[2]*Bond Business School, Bond University, Queensland, Australia*
mrandall@bond.edu.au

[3]*Australian Rivers Institute, Griffith University, Queensland, Australia*
b.stewart-koster@griffith.edu.au

[4] *Graham Centre for Agricultural Innovation, School of Agricultural & Wine Sciences, Charles Sturt University, Wagga Wagga, NSW 2650, Australia*

[5] *Department of Soil Science, College of Agriculture and Applied Biology, Can Tho University, Campus 2, 3/2 Street, Can Tho City, Vietnam*

[6] *Research Institute for Aquaculture No 2, 116 Nguyen Dinh Chieu, District 1, Ho Chi Minh City, Vietnam*

[7] *Centre for Ecosystem Science, School of Biological, Earth & Environmental Sciences, The University of New South Wales, Sydney, NSW 2052, Australia*

Abstract

Efficiencies in farming practice in many parts of South East Asia can make substantial, positive differences to villages and communities. The use of automated decision-assistance tools such as Bayesian Belief Networks (BBNs) can help to accomplish this. For the problem described herein, farmers attempt to grow both rice and shrimp crops in the same physical area. The motivation becomes one of finding a set of conditions that minimises the probabilities of crop failures. In this work, we explore an existing BBN and determine a range of likely environmental scenarios and the factors that farmers can control to help improve the likelihood of harvesting successful rice and shrimp crops.

1. Introduction

Farming any crop requires the efficient use of resources so as to ensure high yield and low risk of crop failure. This is especially so for subsistence farming across the developing world. It is often the case that farming practice is driven by experience and information handed down from previous generations or those in the community. Increasingly, however, this knowledge coupled with advances in agriculture, machine intelligence and data analytics means that far more precision can be embedded into traditional practice to ensure better and more consistent cropping outcomes. In this paper, we examine one such system, in which rice and shrimp are simultaneously farmed in the same physical area in the Mekong Delta, Vietnam. Using an existing belief system, encoded as a Bayesian Belief Network (BBN), we are able to derive sets of conditions with appropriate decision values that improve the probability of crop success. Additionally, we are able to pinpoint the reasons for these, which gives us greater insight into the causal connections in the network.

The remainder of this paper is organised as follows. Section 2 describes Bayesian Belief Networks (BBNs). Section 3 outlines how information was elicited and structured as a BBN applied to the rice/shrimp problem and analyses the resulting network to determine which variables define either the decision or scenario vectors. Section 4 explores the model outcomes, and shows how recommended actions can be deduced and delivered. Finally, Section 5 provides some concluding remarks and outlines future directions for this research.

2. Bayesian Belief Networks

The world consists of complex sets of interrelationships that are often non-linear and probabilistic in nature. A popular modelling technique that is able to capture relationships such as these, is referred to as Bayesian Belief Networks [1, 2]. Technically, a BBN is a directed, acyclic graph. Each node in the graph is called a Conditional

Probability Table (CPT) and has a set of mutually exclusive states that show the probability of achieving a particular outcome, given the condition of its parent nodes. In effect, any ultimate parent node is considered input to the problem, whereas an ultimate child node is an output. These networks may be of unlimited size and complexity in terms of the number of nodes and links.

Given the above characteristics, BBNs have been applied extensively to ecological and environmental systems [3]. There is a number of articles that demonstrate their effectiveness. A couple of the more relevant and related examples are briefly discussed here. The predecessor work to this current paper will be discussed more extensively in the next section.

Using a collaborative approach, between farmers and researchers, to knowledge elicitation, Smith, Russel and King [4] develop a BBN that models the interaction between rodents (rats in particular) and growing rice crops in Cambodia. The aim of this system is to find effective rodent control measures that minimise harm to the human population. To this end, two BBNs were developed. The first concentrated on maximising the effectiveness of the rodent traps while the second concerned the cost benefit ratio. From this, a better understanding of the system interactions was gained.

One of the main issues with BBNs are that they are based on human beliefs, and these are encoded into networks. In quite similar circumstances to the work described in this paper, Baran, Jantunen and Chheng [5] developed BBN-based decision support systems for water management in situations of conflicting, conjunctive water use in agriculture in the Mekong Delta. Pollino, Woodberry, Nicolson, Korb and Hart [6] endeavored to classify CPT nodes as either based on belief (via knowledge elicitation), or determine if they can be encoded with scientifically established data. They demonstrated this "parametisation" approach for the development of a BBN to study risks to native fish communities in the Goulburn Catchment area (Victoria, Australia). Given the improved reliability of the model, they were able to confidently rank the risks to assist management prioritisation efforts.

3. A BBN Applied to the Rice/Shrimp Problem

The work in this paper extends that of Stewart-Koster et al [7] by performing data analysis on their rice/shrimp BBN. In their paper they describe a combined aquiculture system in which rice and shrimp crops are grown in the same physical area. Essentially, this system consists of a pond which contains a raised soil platform surrounded by a narrow, but deep ditch. The rice is planted on the platform while the shrimp mainly inhabit the ditch, but often feed on algae associated with the rice plants during the night. Both crops, however, require quite different conditions in order to be successful. One of the main differences is that of salty/brackish water. While shrimp may prefer this, it is a major limitation to growing rice.

Resolving issues like the above is not easy, and hence models of the system need to be constructed in order to properly understand it. These model designs require the collaboration of farmers, scientists, facilitators and even translators in a process the authors refer to as *participatory modelling*. The result of this is the production of a model in the form of a BBN which will allow farmers (the end users) to predict the probability of rice and shrimp crop failure under certain conditions such as expected rainfall volume, timing of the onset of the wet season, rice colour (a visual indicator of nutrient deficiency) and the quality of the shrimp stock. For reference, their BBN is reproduced in Figure 1 at the end of this paper.

The analysis conducted by Stewart-Koster et al. [7] revealed a number of interesting characteristics of the network. These are governed by the perceptions of the farmers, rather than necessarily historical or scientific fact. One of the major findings was that the systems were governed, more than anything else, by the timing of the wet season onset. This perception influences, for example, how much fertiliser they thought should be applied to the rice crop, or the quality of the water. On the soil side, better quality soil would influence the shrimp stocking density that the farmers could use. Overall better quality soil was associated with the probability of lower crop failure. However, overall, through the BBN elicitation process, it was found that farmers were generally pessimistic about the success of their rice crops, but more optimistic about shrimp harvests.

The work described in this paper focusses on extracting knowledge from the BBN. By inspecting Figure 1 it may be seen that two key nodes provide information about outcomes: "Risk of rice failure" and "Risk of shrimp failure". By testing inputs to other nodes, particularly those that define the farming environment or represent actions that farmers might take, it is possible to discover the sets of conditions that minimise the probability of rice and shrimp failure. This is not so much a process of *optimisation* as one of *data mining*, discovering relationships by inspection of data [8]. It is possible in this case since selecting one of the possible input conditions for a node sets the probability of that condition to 100%, and the possibility of other conditions to 0%. Hence, there are a finite and limited set of combinations to explore.

Further inspection of Figure 1 reveals a number of nodes that together define specific *scenarios.* Two – "Rainfall volume" and "Wet season onset" – define the governing climatic conditions for the season. Together with "Soil nutrient load", they set the environment in which farmers' actions must be taken. Farmers can then decide on a number of actions prior to planting. They can choose to apply liming agents to the soil, and can till the soil. Other factors that may affect the final outcomes will be what shrimp stocking density they choose, whether they decide to plant salt-tolerant rice, and what water colour management option they use.

At planting, it is possible for farmers to make a series of measurements and estimations: of "Soil salinity", "Platform soil pH", "Water temp[erature]", "Wet season water salinity" and "Soil nutrient load". In Figure 1, the corresponding nodes "cut-out" influences of nodes above, setting the scenario for the planting season. Farmers actions are now limited to whether to choose salt-tolerant rice, the stocking density of shrimp, and water colour management option.

After planting, it would be possible to lower the "cut-out" further, defining the scenario by "Platform soil quality" and "Wet season W[ater]Q[uality]", and what choices were made for rice salt-tolerance and shrimp stocking density. However, we have chosen to keep the "higher level" nodes from the planting season scenario because these are based on physical measurements, rather than subjective assessment: for example, whether wet season water quality was "good for shrimp". Farmers' actions are now limited to assessing rice colour (as part of the scenario), which will determine the level of fertiliser applied, and water colour management.

We have now divided the farming into three periods: pre-planting, planting, and post-planting. In each period, a number of conditions define a particular scenario, and farmers can make a number of choices, gradually decreasing in flexibility as the season progresses.

To investigate these different cases, we undertook three experiments, one for each of the three periods of interest. For each period, we repeatedly ran the BBN, entering "findings" for the nodes to define each scenario. Then each of the possible actions and choices were entered for each scenario. The corresponding "belief" of the probability of failure of the rice crop and shrimp crop was recorded for each combination of choices in each scenario.

It may be noted that Figure 1 includes a number of actions responding to shrimp stress and disease. While these actions significantly impact shrimp survival rates, they are critically dependent on a number of unknown factors, such as the timing of detection of shrimp stress. For this reason, these factors were not included in this particular study.

4. Outcomes of the Rice/Shrimp Model

Once the data had been gathered for each of the experiments, structured essentially as a complete enumeration of a decision tree [9], they were analysed to determine optimal practices within a given period and scenario. A separate dataset was delivered for each experiment, corresponding to one of the three periods (pre-planting, planting and post-planting. Three corresponding scripts in the R statistical analysis language [10] were developed to extract the data for a given scenario, reorder the data for monotonic increase in probability of shrimp failure and determine the set of conditions for mutually minimal probability of failure of rice and shrimp crops. The scripts output this information in text and graphical form.

An example of the output from the pre-planting analysis is shown in Figure 2. This shows the optimal actions for a scenario of early onset of the wet season with heavy rainfall and high soil nutrient load.

The "sweet spot" for best outcomes has automatically been circled on the graphs of probability of rice crop failure (dashed line, upper graph, overlaid with simple moving average smoothing) and shrimp crop failure (solid line, upper graph). The conditions to achieve the minimal probability of failure in each crop (presented at the top) are given at top left. (The scenario conditions are provided for reference at the top right.) Actions taken are graphed below. These allow exploration of "what if" variations; for example, it can be seen that if shrimp crop density is increased to “high” (labelled dotted line below the main graph, toward the right) there is a corresponding major increase in the probability of shrimp crop failure.

An example of output for the planting season is shown in Figure 3. The scenario shown is for low soil and water salinity, balanced soil pH, high soil nutrient load and low water temperature. As can be seen, there are fewer actions available to farmers (the graphs have fewer distinct cases within the scenario, though there are more scenarios defined.) Once again the link between stocking density and shrimp crop failure can be seen.

Finally, an example of post-planting output is shown in Figure 4. The scenario is now defined by 8 conditions, and there is only one choice available to farmers: water colour management option. The optimal actions also mention that fertiliser applied should be above the recommended level. However, this is not a “free” choice – it is determined by the farmer’s assessment of the rice colour, and the factors contributing to the “Platform soil quality”, determined by the choice of scenario. It must be extracted from the BBN data by searching for the fertiliser level with the highest conditional probability corresponding to the optimal actions within the scenario. This highlights a consideration that should ideally be kept in mind when designing BBNs – nodes defining actions to be chosen should not have parent nodes that constrict their conditional probabilities. The only choice left at this stage is water colour management, which has negligible impact on outcomes. While there are only two states within each scenario depending on this choice, there are over 2500 distinct scenarios of detailed specification.

5. Conclusion

The work described in this paper has harvested knowledge from data generated by a Bayesian Belief Network (BBN) derived from extensive prior consultation and development work. It provides an overview of how the knowledge was extracted using simple data mining and visualisation techniques. It is hoped that novel presentation of the outcomes can provide guidance and assistance to farmers on the land, and consultation is planned to try to determine what questions are of prime interest, and how information can best be presented.

A second phase of the broader research program is looking to incorporate scientific data from experiments in the field into a new BBN. The structure of this new BBN is being informed by the attempts at knowledge extraction reported in this paper; in particular, ensuring that nodes that embody decisions should not be subject to constraint of the choices from “parent” nodes. The conceptual partitioning of the network into distinct phases corresponding to different periods of the growing season is also being given further attention, to allow easy entry of field data to “cut-out” effects from “higher” nodes at the boundary between phases.

In terms of practical application of this approach, we would like to transfer the knowledge described in this paper into a mobile app that will be readily available to farmers as a decision support tool. The idea would be for the app to access a data file that is the output of all possible scenarios with decision variables, the decision tree, from the BBN. Users of the app, such as farmers and regional planners, would simply select the scenario (with fields such as wet season onset, and rainfall volume) to return the recommended actions (e.g. soil lime treatment and tilling, rice salt tolerance, shrimp stocking density, etc.) to maximise their chances of simultaneously harvesting successful rice and shrimp crops. The recommendations would be tailored to the needs of one of the many, specific scenarios. Further refinements might include entering production costs and yield income to incorporate economic conditions into the decision making under the model.

References

[1] K. Korb, and A. Nicholson. *Bayesian artificial intelligence*, Chapman & Hall/CRC, Florida, 2003.

[2] J. Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann, 2014.

[3] C. Leigh, L.H. Hiep, B. Stewart-Koster, D.M. Vien, J. Condon, N. Van Sang, J. Sammut, and M.A. Burford. "Concurrent rice-shrimp-crab farming systems in the Mekong Delta: Are conditions (sub) optimal for crop production and survival?." Aquaculture Research 48(10), 2017, pp. 5251-5262.

[4] C. Smith, I. Russell, and C. King. "Rats and rice: belief network models of rodent control in the rice fields of Cambodia." In Proceedings of MODSIM 2005 International Congress on Modelling and Simulation. A. Zerger, R.M. Argent (Eds.), Modelling and Simulation Society

of Australia and New Zealand, vol. 449, 2005, p. 55.

[5] E. Baran, T. Jantunen, P. Chheng, and C.T. Hoanh. "Integrated management of aquatic resources: a Bayesian approach to water control and trade-offs in Southern Vietnam", Tropical deltas and coastal zones: food production, communities and environment at the land-water interface, 2010, pp. 133-142.

[6] C.A. Pollino, O. Woodberry, A. Nicholson, K. Korb, and B.T. Hart, “Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment”, Environmental Modelling & Software, 22(8), 2007, pp.1140-1152.

[7] B. Stewart-Koster, N. Anh, M. Burford, J. Condon, N. Van Qui, D. Van Bay and J. Sammut, “Expert based model building to quantify risk factors in a combined aquaculture-agriculture system”, Agricultural Systems, 2017, pp. 157:230-240.

[8] U.M. Fayyad, A. Wierse, and G.G. Grinstein, (Eds.), Information visualization in data mining and knowledge discovery, Morgan Kaufmann, 2002.

[9] L. Rokach, and O. Maimon, Data mining with decision trees: theory and applications World Scientific, 2014.

[10] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2017, Available at https://www.R-project.org/.

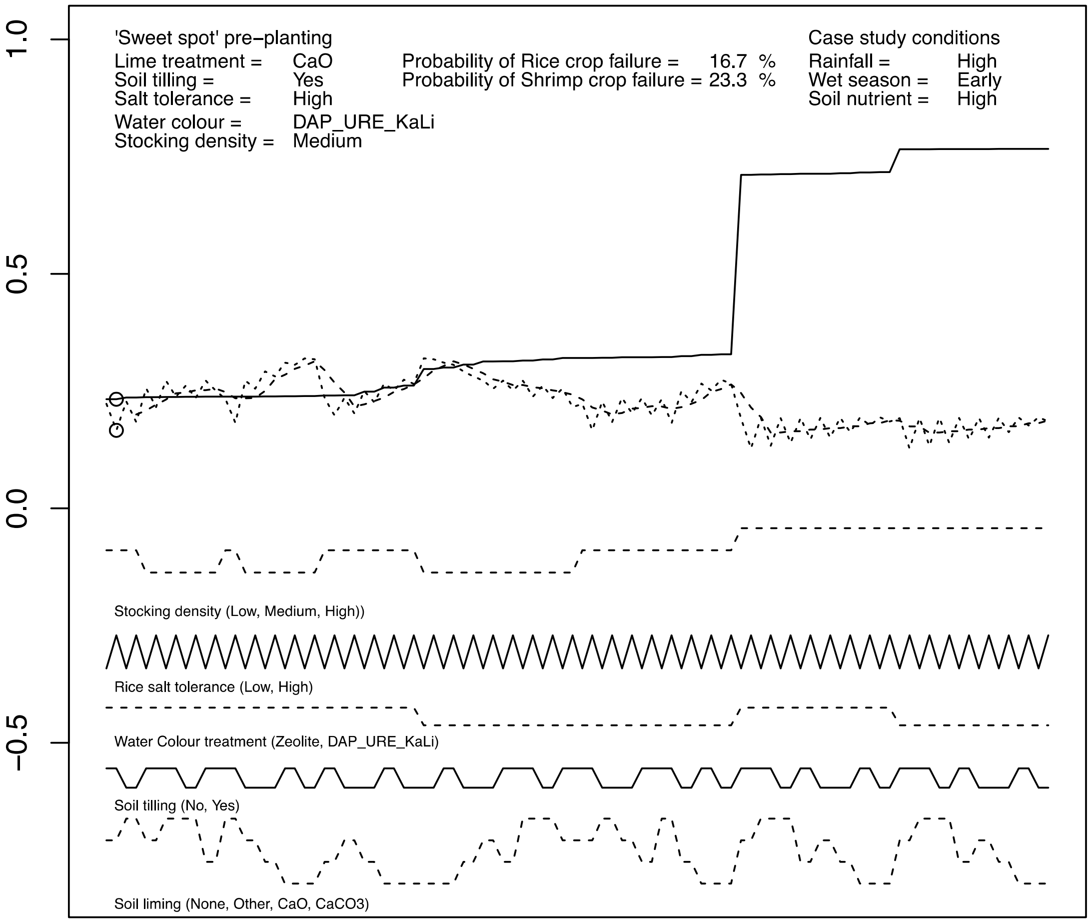

Figure 2. Pre-planting experimental output.

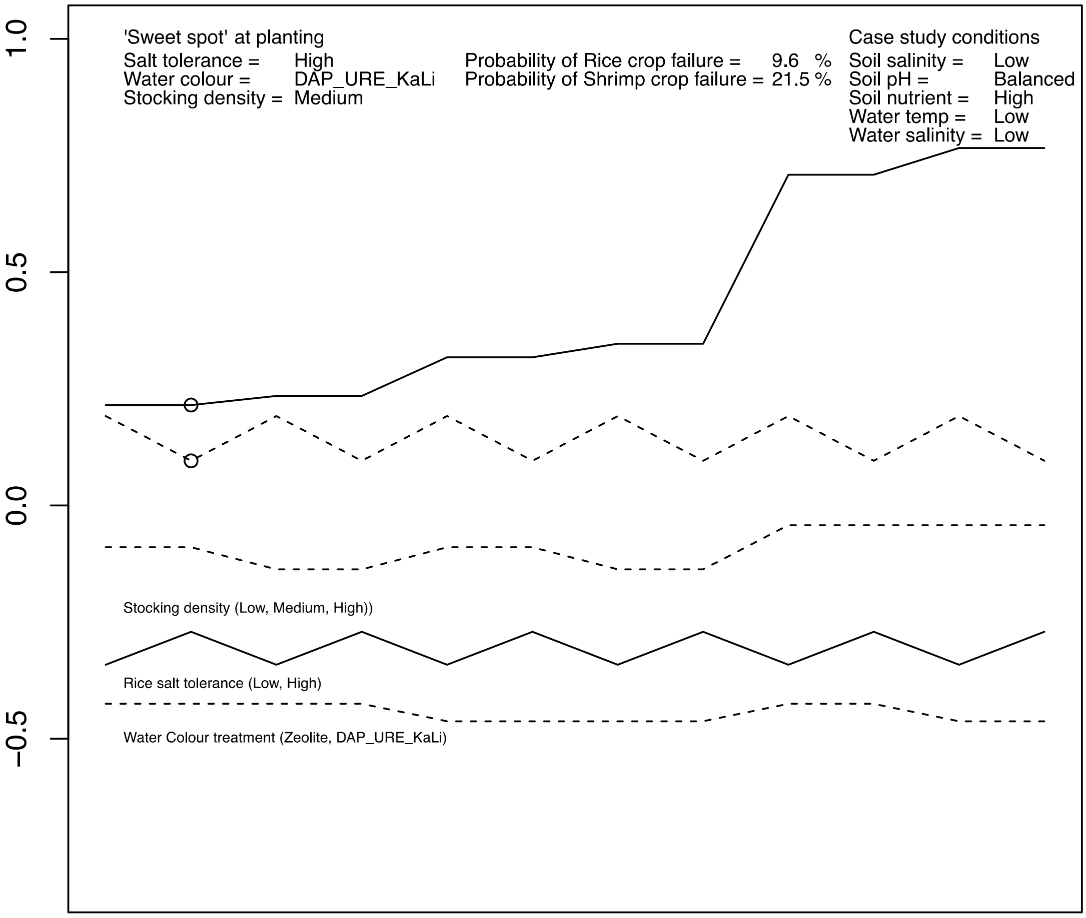

Figure 3. Planting season experimental output.

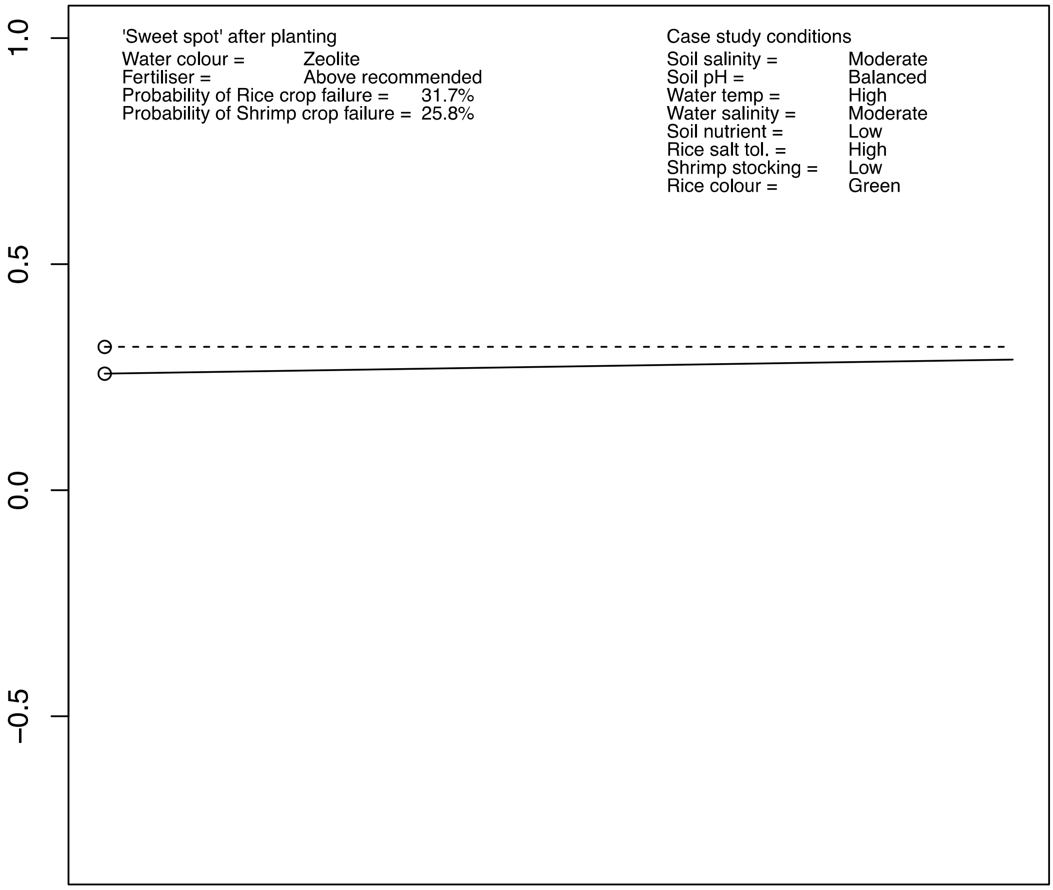

Figure 4. Post-planting experimental output

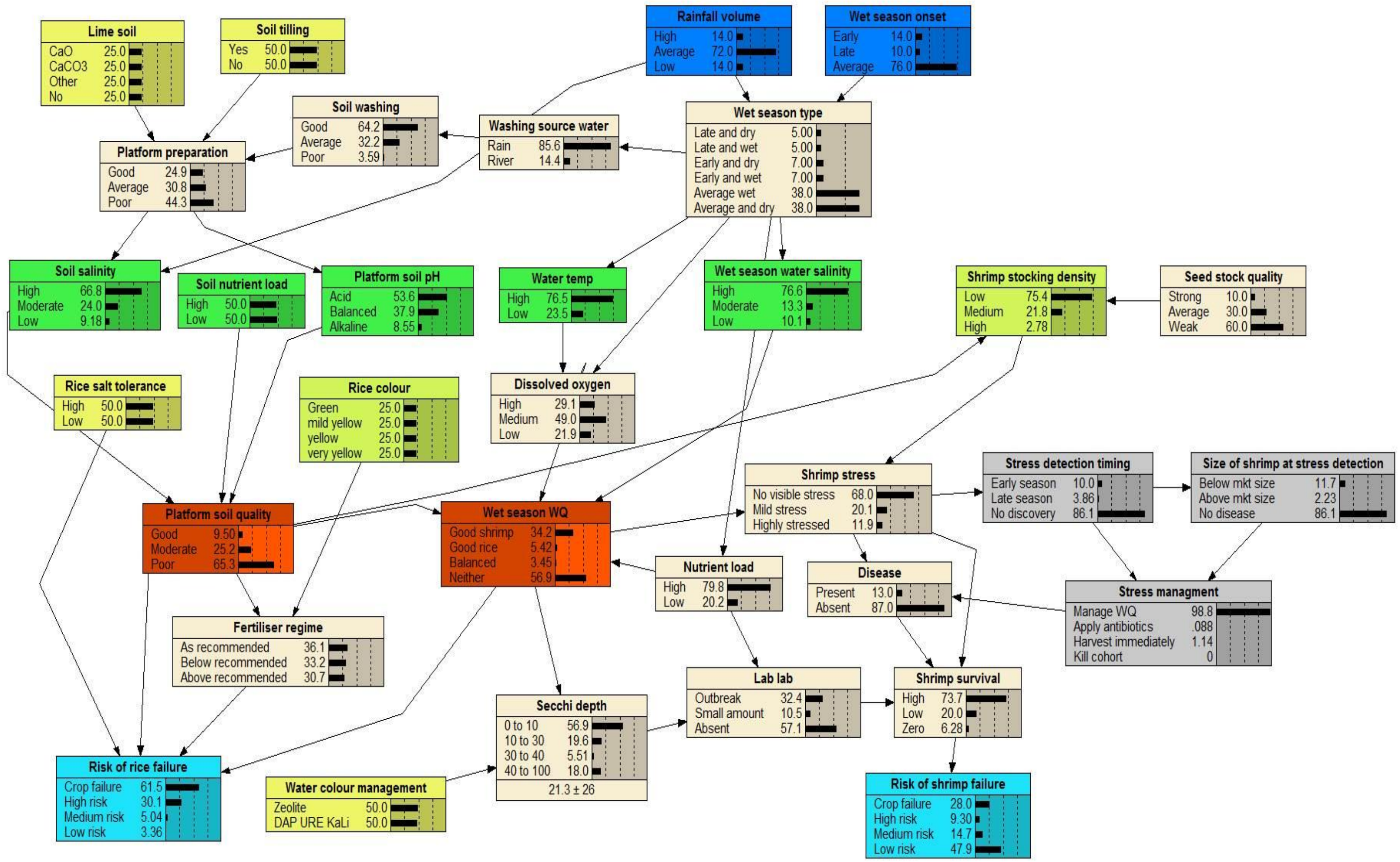

Figure 1. The BBN representing the knowledge and experience of the participating farmers and extension officers involved in the broader research program.