Modeling farmers' decisions: a comparison between HDM and CART for oats-vetch adoption in the Ethiopian Highlands

I. Darnhofer, R. Gretzmacher and W. Schneeberger

Modellierung von bäuerlichen Entscheidungen: ein Vergleich zwischen HDM und CART anhand von Hafer-Wicke-Anbau im Äthiopischen Hochland

1. Introduction

The rate of adoption of new or improved agricultural technologies in Third World countries has not been satisfying, as neither the size nor the distribution of the benefits have matched the expectations of the implementing agencies and national governments. This result is generally attributed to the neglect of the "human element" in farming systems in the traditional research approaches (NORMAN and BAKER, 1986; WALKER et al., 1995). Researchers increasingly realized that the effect of a new technology rarely depends solely on its technical importance, and that "human", "social" and "nontechnical" factors need to be taken into consideration, if the full potential of a technology is to be exploited.

As direct and creative farmer participation has been elusive the necessity for a two-way linkage between various participants in the research process was recognized. New approaches seek to thoroughly understand the farmers' situation and perspective and to integrate these factors into agricultural research (CHAMBERS, 1994). Their roots are found in the recognition that a farmer's decision depends on, and is influenced by, his/her own knowledge and perception of a technology, rather than the researcher's knowledge of the technology (GLADWIN et al., 1984).

The underlying assumption is that the decision makers themselves are the experts on how they make the choices they make. With the focus on the farmer whose adoption or rejection of the new technology can make or break a project, the researcher needs to know: (1) what decision the farm household is making, (2) what alternative he/she is considering in each decision context, and (3) why he/she chooses a particular outcome (GLADWIN, 1982).

The present study aims to elucidate the criteria influencing the adoption decision of an improved feed for crossbred

Zusammenfassung

Die Entscheidung von Bauern ein verbessertes Futtermittel (Hafer-Wicke) anzubauen wird analysiert, und mit der Hilfe von zwei Methoden modelliert: hierarchische Entscheidungsbäume (HDM) und Klassifikations- und Regressionsbäume (CART). Die Modelle werden dazu verwendet, zu erklären, warum Bauern im Äthiopischen Hochland die verbesserte Futtermittelbasis nicht im geplanten Ausmaß angenommen haben. Die Stärken und Schwächen der zwei Methoden werden verglichen. Beide Methoden werden als hilfreich befunden, um Informationen über die Meinung, Haltung und Taten der Bauern in sinnvollen Zusammenhang zu bringen.

Schlagwörter: Entscheidungsmodellierung, HDM, CART, Äthiopien, Hafer-Wicke.

Summary

Farmers' decision to adopt an improved feed (oats-vetch) is modeled with the help of two methods: Hierarchical decision trees (HDM) and classification and regression trees (CART). The models are used to help explain why oats-vetch was not adopted by farmers in the Ethiopian highlands to the anticipated extend, and can be used to provide a guide for future extension efforts. The strengths and weaknesses of the two methods are compared. Both are found useful for compiling information about farmers' opinions, attitudes and actions in a meaningful manner.

Key words: adoption decision, HDM, CART, Ethiopia, oats-vetch.

cows in the Selale area of the Ethiopian highlands: intercropped oats-vetch. The study was made in the context of a development project by the Ethiopian Ministry of Agriculture (MoA) and the Finnish International Development Aid (FINNIDA) (MoA/FINNIDA, 1986). The project, which was implemented between 1987 and 1991, distributed crossbred cows against credit to farmers, to enable them to raise milk production both for household consumption and as a source of cash income. As crossbred cows require both quantitatively and qualitatively better feed than local zebu cows, improved feed production was promoted. The recommendations included intercropping oats and vetch. Oats is already well known by the farmers in the area, and intercropping the two species raises both the crude protein content and the dry matter yield, without using supplementary crop land which is scarce.

The models presented hereafter examine reasons why farmers will or will not plant vetch, and if they do, whether or not they will intercrop it with oats, as was recommended by the MoA and FINNIDA. The decision criteria are assembled into decision models using two methods: Hierarchical Decision Models (HDM) as well as Classification and Regression Trees (CART). The juxtaposition of the 'manual' model and the computerized analysis should allow the comparison of the strengths and weaknesses of the two analytical tools. Further details of the research reported herein are presented in DARNHOFER (1997).

2. The methods

2.1 Hierarchical Decision Models

Christina GLADWIN (1976) developed a method called hierarchic decision modeling (HDM) to model the way people make real-life decisions. Its distinctive features are: (1) a reliance on ethnographic fieldwork techniques to elicit decision criteria, and (2) combining these criteria in the form an inverted tree which is read like a flow-chart.

The form of a decision tree is simple: the possible outcome of the decision to be modeled is formulated at the top of the tree for easy reference. Each decision criteria forms a node of the tree. The decision variables or criteria have discrete values (GLADWIN, 1975). These discrete criteria can be either rejected or accepted. After a number of criteria have been passed, the decision outcome is reached at the end of a path of the tree. These outcomes are of the general form: 'do A', 'do not do A'. A decision tree is thus a sequence of discrete decision criteria, all of which have to be passed along a path to a particular outcome or choice (GLADWIN, 1989).

2.2 Classification and Regression Trees

The authors of CART and developers of its computational algorithm are the statisticians Leo BREIMAN, Jerome FRIED-MAN, Richard OLSHEN and Charles STONE (1984). Their aim was to develop an easy to use statistical package for treestructured nonparametric data analysis to tackle classification problems.

The classification trees are also drawn in the form of an inverted tree and are read like a flow chart. To decide how to split a node, the CART software examines all possible splits for all variables included in the analysis, and ranks them on a goodness-of-split criterion. The most commonly used criterion is how well the splitting rule separates the classes contained in the parent node, i.e. decreases class impurity.

The CART output is thus composed of several elements: the optimal tree, with detailed information for each node, such as the variable on which the node was split, the number of cases going right and left, and the improvement in class purity. Additional information on variables are also provided, such as possible competitor variables on which the node could also have been split but with a less optimal result and surrogate variables whose split would have divided the data similarly and which can be used as alternative in case of missing data.

2.3 Data collection

Following the method set forth by FRANZEL (1984), the data for this study were collected in two stages: first all the decision criteria were elicited from key informants and, in a second stage, the criteria were compiled into a formal questionnaire and data gathered from 50 randomly selected farmers, providing the data used in the models.

The key informants for the interviews held in the first stage were knowledgeable enumerators, farmers and MoA officials. These were instrumental in acquiring a better understanding of influencing factors and the problems that farmers may face with vetch. These first stage interviews took the form of informal conversations, without pre-formulated questions. During the interviews care was taken to follow ethnographic guidelines (SPRADLEY, 1979; ATTES- LANDER, 1993). This first stage is an iteration between eliciting criteria, building draft decision trees to organize the criteria, eliciting further criteria and modifying the draft tree. Once the draft decision tree seemed reasonably complete, the questions underlying the decision criteria were written up and a formal questionnaire designed.

In the second stage of data collection, 50 randomly sampled farmers who had participated in the MoA/FINNIDA project were interviewed. The data collected during this formal survey were then used to build the final decision model. For the CART analysis all questions were entered for the software to be able to select the most appropriate splitting variables.

3. Decision models of oats-vetch

3.1 HDM of oats-vetch

Figure 1 starts with reasons why some farmers have never planted vetch, and will therefore be excluded from the further model, since they have no direct experience with the crop. It is better to illicit criteria from a farmer who has actually made a decision, rather than one who has no firsthand knowledge, as his/her answers will likely be hypothetical. Of the 50 interviewed farmers, eight have never planted vetch, as they never got seeds, or do not think vetch will grow in their area, or do not have enough land to plant vetch. This last group of two farmers, one of whom has never planted, and the other having planted vetch once, shows a reluctance of the farmers to intercrop vetch at the first trial. Also, "lack of land" is a surprising argument as if they would intercrop oats and vetch, as recommended, it would not require any additional land, making it a practice particularly relevant to those farmers who do not have enough crop land. The reluctance to intercrop might partly be due to the fact that intercropping is not a traditional practice in the area.

The remaining farmers have all planted vetch at least once, and the criteria influencing their choice whether or not to plant it this year can be analyzed. A first group of seven farmers exits the decision tree, as they are 'not satisfied' with the performance of vetch, when they first grew it. This poor performance might be due to cool temperatures at the higher altitudes, or a lack of water at the lower altitudes, depending on the rainfall in the year when the farmer first planted the crop.

The decision tree then subdivides the farmers in two altitude sub-locations: those living at an altitude ranging from 2500 m to 2700 m and those living at an altitude range between 2700 m and 3000 m. Above the 2700 m range, frosts are likely to occur between November and December, the period during which vetch flowers. These occasional night time frosts seem to hinder the seed production of vetch.

In the view of the forage experts the impaired seed production is not a problem, as the oats-vetch mixture is meant to be harvested and fed green. But to farmers this is a major factor. On the one hand, a number of farmers remarked that planting a crop which does not produce seeds is useless. On the other hand, farmers are well aware that if they cannot produce their own seeds, they entirely depend on the MoA for the supply of vetch seeds, as these are not available on the market. Since the MoA does not have sufficient vetch seed for distribution to all interested farmers every year, it is a constraining factor. In other words, a farmer who is not able to produce his own seed, will most likely not be able to plant it again the following year. This is reflected in the answer of 13 farmers, who, despite the fact that they are interested in growing the crop, "will not plant vetch this year, due to a lack of seeds".

In figure 1 most farmers are satisfied with the results of intercropping and will plant vetch this year if they receive seeds from the MoA or if they have left-over seeds from a previous year. As seed production is not an option due to the risk of frost, farmers in this area are more likely to intercrop. The two farmers who are not satisfied with intercropping oats and vetch, could still plant vetch as a sole crop, but the farmer who would like to do so is prevented by his lack of seeds.

Figure 2 presents the decision process of farmers living at lower altitude, and who can secure their own supply of seeds if they plant at least part of the vetch as a sole crop. Two problems appear at that level: first a palatability problem, as farmers say that their cows do not like vetch. The other problem the uncertainty concerning the 'right' planting time for intercropped oats and vetch.

The MoA recommendation is to intercrop and plant towards the end of the rainy season, i.e. in September. To farmers vetch (*Vicia datycarpa*) is a novel crop with which they have no experience, but due to the likeness of vetch seeds with the ones of rough pea (*Lathyrus sativa*), a crop most of the farmers plant, they conclude that these two crops will have similar characteristics. This link in the farmers' belief is shown in the fact that they use the same oromo word ('guaya') for both plants, differentiating them only through specifying whether it is 'guaya' (rough pea) or



Figure 1: Oats-vetch decision tree – Part 1 Abbildung 1: Hafer-Wicke Entscheidungsbaum – Teil 1

Die Bodenkultur

48 (4) 1997



Figure 2: Oats-vetch decision tree – Part 2 Abbildung 2: Hafer-Wicke Entscheidungsbaum – Teil 2

Die Bodenkultur

'horii guaya' ('animal *guaya'*, i.e. vetch). Rough pea is a crop common to the area, which many farmers plant, albeit in very small quantities. Its main advantage is that it grows on poor soils and requires very little water, so that a harvest can be assured, even when there is a poor rainy season and most other crops fail.

Rough pea is normally planted towards the end of the rainy season, i.e. in September, which is therefore, to the farmers, also the optimum time to plant vetch. But because oats are traditionally planted around May-June, intercropping these two crops results in some confusion and doubts. On the one hand farmers know that rough pea (and therefore expect vetch to be the same) does not grow when planted around June, as it does not tolerate waterlogging, which occurs usually in July-September. On the other hand, planting oats in September does not seem to be appropriate either, as it will not mature on the residual moisture at the end of the rainy season, as does rough pea. For the farmers intercropping oats and vetch therefore causes confusion concerning the planting time, and they are likely to proceed with caution. Also, as mentioned above, few farmers will intercrop all of their vetch, as they then will not harvest any seeds, making them entirely dependent on the MoA for seeds in the next year.

Of the 21 farmers included in Figure 2, four will not plant vetch due to a lack of interest, caused by the lack of palatability of vetch, six would be interested but cannot plant as they do not have own seeds and did not receive any from the MoA. The remaining 11 will plant vetch this year.

To summarize the outcome of the hierarchical decision model for oat-vetch: of the 50 interviewed farmers, 41 have planted vetch at least once and can therefore base their decision whether or not to plant vetch on their own experience with this novel crop. Twelve have no further interest in the crop, as they were not satisfied with the growth performance, or their cow did not find it palatable. Of the 29 farmers who are interested in growing vetch, only 55 % will do so, because the remaining 45 % of farmers do not have the necessary seeds. The large majority (75 %) of the farmers who will plant vetch this year, will follow the recommendation of the MoA and intercrop at least part of the vetch. Still, where the climate allows, they will also plant some vetch for seed multiplication, as they cannot rely on the MoA to provide them with seeds in the following year.

3.2 CART of oats-vetch

The tree in Figure 3 shows the CART tree, classifying farmers in five classes: (0) farmers who have never grown vetch, (1) whether the farmer will plant vetch as a sole crop, or (2) intercrop it, or (3) do both, and (4) those farmers who will not plant vetch this year. These categories are similar those of the HDM tree in Figure 1 and Figure 2.

The first question selects those farmers who have never grown vetch, as these were not asked the remaining questions due to their lack of personal experience with the crop. Those who have grown vetch at least once are then divided between the area above 2700 m and those below. The next question for farmers in frost-prone areas concerns the issue whether or not they can get vetch seeds from the MoA. Of the seven farmers who can get seeds from the MoA and who therefore are classified as farmers who will intercrop, only five will plant and intercrop, and two are misclassified as they will not plant vetch.

The farmers located in the lower area, where vetch produces seeds are asked whether or not they think that oats and vetch grow well together, a question addressing the problem of optimal planting time. Those who do not think intercropping advisable are classified as farmers who will not plant. Of the 12 farmers in this terminal node, nine will not plant vetch at all, and three are misclassified, as they will plant it as a sole crop. The farmers who think that intercropping oats and vetch is not a problem, are asked if their crossbred cow likes the mixture of green oats and vetch. Two farmers say that their cow does not like it, of which one will not plant vetch at all, and the other will plant some alone, and some intercropped. Of the 10 farmers whose crossbred cows like the oats-vetch mixture, six farmers will plant vetch both as a sole crop and intercropped with oats, one will intercrop all of its vetch and three will not plant any vetch this year.

The consiceness of the CART tree allows the decisive factors to become more apparent than in a larger tree with more details. On the other hand, the price for such a succinct tree is misclassification: of the 50 interviewed farmers, 10 are misclassified. Only two terminal nodes are 'pure': those farmers who have never grown vetch, and those who will not plant due to lack of seeds from the MoA. The four remaining terminal nodes all have some 'impurity', i.e. misclassified cases. For CART splitting these nodes further does not sufficiently reduce node impurity to be worth the higher complexity, i.e. the penalty imposed per additional terminal node. The 'right' size is a trade-off between misclassified



Figure 3: Oats-vetch classification tree

Abbildung 3: Hafer-Wicke Klassifikationsbaum

cases and tree size, as in virtually all applied statistics parsimony is considered a desirable feature in a model (STEIN-BERG and COLLA, 1995).

3.3 Summary of influencing factors

Overall the main problem seems to be seed availability, as the Ministry of Agriculture does not have sufficient seed to supply all interested farmers. Therefore farmers have to secure their own seed which is only possible if the farmer lives in the lower altitude area and if he does not intercrop all of his vetch since when intercroped it will be harvested at the flowering stage. The seed supply is therefore one reason some farmer are reluctant to intercrop oats and vetch, but the reluctance might also be due to the fact that intercropping is not a traditional practice in the area.

Another factor that became clear in the course of the interviews is that the similarity between the seeds of rough pea and of the vetch misleads the farmers to assume the two crops have similar characteristics concerning planting time. Linking the two crops is also problematic as rough pea is toxic when consumed in larger quantities, so that farmers might erronerously think that vetch could be toxic for their cows. This fear of toxicity might be one reason some farmers do not put too much energy in getting their cow used to this novel feed and complain of palatability problems.

Most of the factors inhibiting the wider adoption of oatsvetch in the Selale area can be adressed through extension and demonstrations at field-days by the MoA. To reinforce the extention message, seeds need to be made available to farmers who are already interested and who can therefore serve as demonstration farmers to their neighbors.

4. Evaluation of HDM and CART

Although simple in appearance, the above described decision trees of a feed crop in the Ethiopian highlands, have shown that decision models are useful for assembling information on farmers' opinions and perceptions in a systematic way so as to show the logic behind farmers' decisions. This is a strong advantage compared to the more classical statistical analysis which will focus mainly on the frequency that a factor is mentioned by interviewed farmers without revealing the chain of though of the farmers and the interconectedness of the factors. The two methods have their respective strengths and weaknesses, which will be displayed in different research settings. One difference lies in the influence of the sample size. In many agricultural data collection settings, the number of farmers interviewed rarely exceed 100 or 200, due to the high demands in personnel and time necessary to interview each farmer, as well as the limited funds and time available for most data collection. This relatively small sample size, particularly its lower range, is not a problem for analysis through HDM, if the number of variables on which the data are collected is relatively limited. CART, on the other hand, will analyze several thousand cases and a large number of variables effortlessly. The larger the number of cases, the more appropriate CART is likely to be, because patterns are easier to distinguish with larger sample sizes.

HDM has limitations concerning the type of information which can be included in the model: the decision to be modeled needs to be narrowly framed, i.e. have only a few possible outcomes, and the alternatives and decision criteria must be discrete. Where a wide array of choices is available, and most farmers select several different alternatives, the analysis becomes muddled with multiple nodes and branches. This makes the analysis excessively complex, and the interpretation difficult or impossible. Should the decision variable be continuous (e.g. area planted with oats-vetch), it can only be modeled if, in the farmers' view, there is a logical reason allowing interval formation, thereby making the variable discrete.

In most decision models this should not be a problem. But in certain circumstances it might be desirable to widen the data collected to variables such as household size, cattle number or hectare land cultivated, and analyze it together with the ethnographic data. CART will be able to find the most appropriate grouping and therefore split between the groups, even with continuous variables, finding the cut-off value through the same method of reduction in node impurity.

Another difference between HDM and CART is the size of the resulting tree. Trees built by CART will in almost all cases be more compact than trees built through HDM. This also means that the amount of information contained in a CART tree will be lower compared to a HDM tree on the same topic. For example, the HDM tree in Figure 2 specifies why the farmer will not grow vetch this year (e.g. 'lack of interest'), whereas in the CART tree Figure 3, the outcome is simply labeled as 'will not plant'. Which one is preferable will depend on the context and the main aim of the study. If the study aims at an overview of the important

Die Bodenkultur

factors, so as to address them in detail later, a concise tree will be preferable, as the decisive factors are more apparent. On the other hand the nuances and individual differences and therefore the complexity of a real-life situation tend to get lost, so that a larger tree might be preferable.

A limitation to decision trees in general, is that they will be difficult to apply successfully if a farmer does not have a definitive opinion. In this case, his or her criteria will be vague and it will be hard to pin down what the perceived problem is. For example, after some unsuccessful initial trials to feed the oats-vetch mixture to their cow, it is likely that farmers loose interest since they do not have enough information as to the advantages of persevering and getting the cow used to the novel feed. This inaction is then rather due to a lack of knowledge of potential advantages, rather than a strong feeling concerning perceived drawbacks.

HDM and CART also differ in their link to the data collecting process. A distinguishing feature of HDM is that the tree building process cannot be separated from the data collection process. The ethnographic interviews in the first stage will reveal and focus on those variables considered important by the farmers. During the formal interviews in the second stage, only the pertinent questions will be asked, and each question will correspond to a node in the final tree. With CART on the other hand, the burden is still with the analyst in constructing a set of questions which can effectively extract information from the data. However, CART permits a considerable 'overkill': many questions that may or may not turn out to be informative can be included in the analysis. This is particularly useful in problems where it is not clear which features and what threshold values are instructive. CART can therefore also be used to analyze data that were not specifically collected for decision models.

Due to the largely manual data analysis underlying the HDM trees, they have a large amount of flexibility, relying mainly on the understanding the analyst has of his/her subject. This allows for a good analyst to build a tree in which the logical links and the thought process of the farmers are clearly reflected and easily understood, even for outsiders. For CART, the data structure needs to be standardized and it is analyzed with algorithms which focus on the best split to lower node impurity. A number of the trees produced by CART therefore were not a 'logical' succession of factors affecting the farmers' choice, as in the HDM trees, but reflected the well known discrepancy between correlation and causation. It is then necessary to guide and modify the analysis so as to make CART build a tree with a good explanatory power.

5. Conclusion

Decision trees allow to structure and present information gathered from farmers through ethnographic interviews in a logical and easily understood form. Some of the weaknesses and strengths of the two methods used in this research have been illustrated. Still, with both HDM and CART, the analyst needs a 'feel' for the data in order to reach a comprehensive assessment of the variables under consideration. Therefore, with both methods, as with all data analysis, yielding good and honest results is "a mixture of art and science, involving considerable subjective judgment" (BREIMAN et al., 1984).

Acknowledgments

I would like to thank the farmers for their patience during the interviews, the many people at ILRI who helped me during the field research in Ethiopia and my professors, Dr. Schneeberger and Dr. Gretzmacher for their support during the write-up of the study.

References

- ATTESLANDER, P. (1993): Methoden der empirischen Sozialforschung. 7. Auflage. Walter de Gruyter, Berlin, Sammlung Göschen.
- BREIMAN, L., J. FRIEDMAN, R. OLSHEN and C. STONE (1984): Classification and regression trees. Wadsworth and Brooks, Cole advanced books and software. Monterey, California (USA).
- CHAMBERS, R. (1994): The origins and practice of participatory rural appraisal. World Development, 22(7), 953– 969.
- DARNHOFER, I. (1997): Tagasaste and oats-vetch production for crossbred cows in the Ethiopian highlands. Animal feeding, farmer decision making and economic considerations. Doctoral dissertation. University of Agriculture Vienna.
- FRANZEL, S. (1984): Modeling farmers' decisions in a farming systems research exercise: The adoption of an improved maize variety in Kirinyaga District, Kenya. Human Organization, 43(3), 199–207.
- GLADWIN, C. (1975): A model of the supply of smoked fish from Cape coast to Kumasi. In: S. PLATTNER (ed): Formal methods in economic anthropology. Special publication

of the American Anthropological Association, No. 4, 77-127.

GLADWIN, C. (1976): A view of the Plan Puebla: An application of hierarchical decision models. American Journal of Agricultural Economics, 58(5), 881–887.

GLADWIN, C. (1982): The role of a cognitive anthropologist in a farming systems program that has everything. In: Report of an exploratory workshop on the role of anthropologists and other social scientists in interdisciplinary teams developing improved food production technology. Los Banos, The Philippines: IRRI, 73–92.

GLADWIN, C. (1989): Indigenous knowledge systems, the cognitive revolution, and agricultural decision making. Agriculture and Human Values, VI(3), 32-41.

GLADWIN, C., R. ZABAWA and D. ZIMET (1984): Using ethnoscientific tools to understand farmer's plans, goals, decisions. In: P. MATLON, R. CANTRELL, D. KING and M. BENOIT-CATTIN (eds): Coming full circle: farmers' participation in the development of technology. Ottawa, Canada: IDRC, 27–40.

MoA/FINNIDA (1986): Selale peasant dairy development pilot project. Project document. Draft. Addis Abeba: Ministry of Agriculture, Ethiopia and FINNIDA, Ministry of Foreign Affairs, Finland.

NORMAN, D. W. and D. C. BAKER (1986): Components in farming systems research, FSR credibility and experiences in Botswana. In: J. L. MOOCK (ed.): Understanding Africa's rural households and farming systems. Westview Press, Boulder and London, 36–57.

SPRADLEY, J. P. (1979): The ethnographic interview. Holt, Rinehart and Winston, New York.

- STEINBERG, D. and P. COLLA (1995): CART. A Salford Systems implementation of the original CART program. Interface and documentation. San Diego, CA: Salford Systems.
- WALKER, D. H., F. L. SINCLAIR and B. THAPA (1995): Incorporation of indigenous knowledge and perspectives in agroforestry development. Part 1: Review of methods and their application. Agroforestry Systems 30(1–2), 235–248.

Corresponding Authors

Dr. Ika Darnhofer, Marschallplatz 5/12, A-1120 Wien.

Ass.-Prof. Dr. Ralph Gretzmacher, Institut für Pflanzenbau und Pflanzenzüchtung, Universität für Bodenkultur Wien, Peter Jordan-Straße 82, A-1190 Wien.

O. Univ.-Prof. Dr. Walter Schneeberger, Institut für Agrarökonomik, Universität für Bodenkultur Wien, Peter Jordan-Straße 82, A-1190 Wien.

Eingelangt am 29. April 1997 Angenommen am 6. Juni 1997