ORIGINAL PAPER

Using fuzzy logic modelling to simulate farmers' decision-making on diversification and integration in the Mekong Delta, Vietnam

Roel Bosma · Uzay Kaymak · Jan van den Berg ·
Henk Udo · Johan Verreth

Abstract To reveal farmers' motives for on-farm diversification and integration of farming components in the Mekong Delta, Vietnam, we developed a fuzzy logic model (FLM) using a 10-step approach. Farmers' decision-making was mimicked in a three-layer hierarchical architecture of fuzzy inference systems, using data of 72 farms. The model includes three variables for family motives of diversification, six variables related to component integration, next to variables for the production factors and for farmers' appreciation of market prices and know-how on 10 components. To obtain a good classification rate of the less frequent activities, additional individual fine-tuning was necessary after general model calibration. To obtain the desired degree of sensitivity to each variable, it was necessary to use up to five linguistic values for some of the input and output variables in the intermediate hierarchical layers. Model's sensitivity to motivational variables determining diversification and integration was of the same magnitude as its sensitivity to market prices and farmers' know-how of the activities, but less than its sensitivity to labour, capital and land endowment. Modelling to support strategic decision-making seems too elaborate for individual farms, but FLM will be useful to integrate farmers' opinions in strategic decision-making at higher hierarchical levels.

R. Bosma (✉) · J. Verreth
Aquaculture and Fisheries, Wageningen University,
P.O. Box 338, 6700 AH Wageningen, The Netherlands
e-mail: roel.bosma@wur.nl

U. Kaymak
Econometric Institute, Erasmus University,
Rotterdam, The Netherlands

J. van den Berg
Fac. Technology, Policy and Management,
Delft University of Technology, Delft, The Netherlands

H. Udo
Animal Production Systems, Wageningen University,
Wageningen, The Netherlands

1 Introduction

A recent overview of crop–livestock simulation models recognised that the household's stage of development and its effect on strategic decision-making (DM) have not been sufficiently studied or considered in model development (Thornton and Herrero 2001). Though it is more and more recognised that motivations other than 'utility maximisation' might guide human decisions, most bio-economic models of farming systems still do not include basic human motivations. Moreover, farmers' motives are mostly expressed in subjective linguistic terms instead of continuous variables. In this context, fuzzy logic models (FLMs) are considered a good alternative to the linear models based on utility theory (Fodor et al. 1998; Guillaume and Magdalena 2006). Fuzzy set theory (Zadeh 1975) can provide a more powerful tool for modelling complex human reasoning than classical models (Türksen 2004). In a recent study in Thailand, the crop choices of about 300 farmers was successfully simulated using fuzzy logic data mining (Ekasingh et al. 2005), but the model did not integrate the social dimensions (Ekasingh and Letcher 2005). In our study, we therefore set out to include in a FLM farmers' social drives and motives for on-farm diversification and integration of farming components,

🖄 Springer

which might lead to integrated agriculture–aquaculture farming systems (IAASs).

Although specialisation is the global trend in agriculture, in the last 30 years IAASs have successfully emerged in Vietnam (Luu 1992; Prein 2002). In the Mekong Delta in particular, the adoption of improved technologies that have allowed two—and more recently three—rice crops per year has improved household's food security and gradually made land available for farmers to invest in other activities, especially when the market price of rice is low (Bosma et al. 2006a). Some of these activities have been closely integrated for reasons of synergy or lack of space, and mixed systems have emerged: fruit–fish, rice–fish, pigs–fish (Sanh et al. 1998; Prein 2002). In previous papers, we analysed empirically the contextual drives and social motives for the on-farm diversification in the Mekong Delta (Bosma et al. 2006a, b).

In the present paper, our main objectives are to explore the effect of farmers' family motivations on farm diversification and integration by including those in a FLM simulating their DM. After presenting our methodology, we will describe and discuss the results and mainly the constraints and opportunities of using a FLM to simulate the composition of mixed farms in the Mekong Delta.

2 Methodology

In Sect. 2.1, we motivate our choice to apply manual procedures in the modelling process, for which we used a 10-step approach (Fig. 1). The corresponding activities to be performed are described in the other sections of this chapter. We summarise steps 4–6 and 8 in Sect. 6, and we give details of steps 1, 3, 7, 9 and 10 in other sections.

2.1 Fuzzy logic modelling

Based on the 3-staged procedure proposed by Emami et al. (2000) and the seven steps described by Jang et al. (1997), we developed an iterative 10-step procedure to develop the FLM-simulating farmers' DM process (Fig. 1). Later, we argue that our choices are determined by the character of the problem and by the related size of the database.

After conceptualising the problem to be dealt with, we selected relevant variables. In agreement with Emami et al. (2000), we first defined the output categories because input variables are relevant in relation to the outputs. The input variables related to family farm household DM can be identified through correlations using statistical analysis of databases and through elicitation of causalities, through knowledge elicitation or through data mining, if a large dataset is available. However, feature extraction is needed to make some variables, such as those related to motivations, operational.

Step 3 aims to identify a FLM structure that mimics the DM process, and is logic-based and knowledge-transparent (Ruspini et al. 1998). When the variables become very numerous, as for human DM making, the number of clauses (or rules) increases exponentially, i.e. the model is exposed to the curse of dimensionality (Setnes 2001). This proliferation of rules can be managed by reducing complexity through: decomposing the FLM, simplifying the rule base and reducing dimensionality (ibid.). The structure of expert systems is identified by eliciting the experts' reasoning, but it may be worthwhile using a data-driven approach to identify the most straightforward and simple structure (Hastie et al. 2001, p. 270; Cai et al. 2005). Straightforward data-driven approaches bring the risk of neglecting DM pathways and reducing transparency,
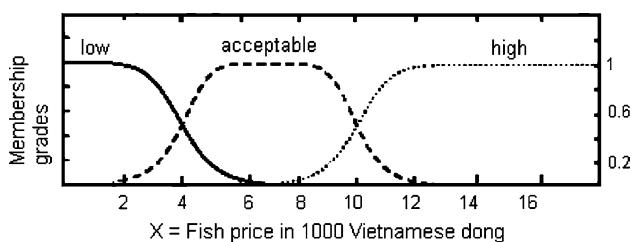

Fig. 1 The 10 steps proposed for the development of a fuzzy-logic model of social decision-making (*DM*) with (*left*) the goal of each step (*centre*) the activities to be performed and (*right*) the goal or the resulting state in the modelling process
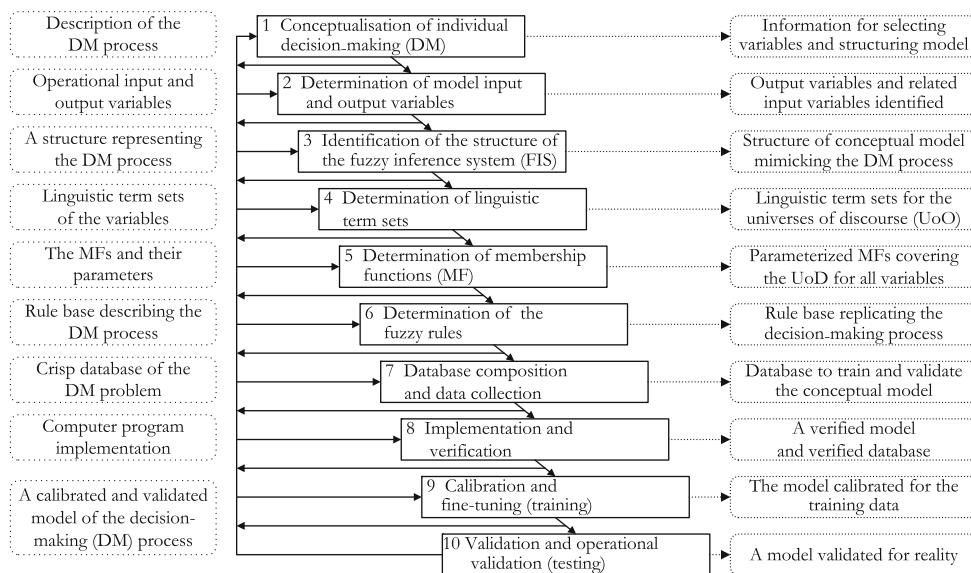
especially when the databases are small (Guillaume and Magdalena 2006). On the basis of the experience with neural networks, it is advisable to deal with complex systems by decomposing them by choosing an overarching structure of various fuzzy inference systems (FISs) before data mining (Geman et al. 1992). For modelling real-world problems, hierarchical FLMs have three advantages: interpretability, accuracy, and dimensionality reduction (Lee et al. 2003; Liu and Li 2005; Zeng and Keane 2005).

In general, for each variable the aggregated surface area of the MFs has to cover the space of discourse of the graph, and the span of the MFs should cover the data dispersion (McCloskey et al. 2006). The type of function chosen depends on the procedures of parameter identification and the character of the variable. Smooth functions, e.g. Gaussian, are required for automated determination of parameters by data-based gradient descent learning and for automated fine-tuning (Gürocak and de San Lazaro 1994). Moreover, smooth functions improve the model's sensitivity, i.e. increase the overlap of the MFs and thus the model's fuzzy character (Zimmermann 1987). We used, in accordance with Jang et al. (1997), either two sigmoidal functions or the smooth asymmetrical polynomial spline-based curves open to the left (z curve) or to the right (s curve), in combination with a Pi curve that is zero at both extremes with a rise in the middle (Fig. 2). For expert systems, the fuzzy partition should be based on automated methods and/or on the distribution and characteristics of the data itself (Medasania et al. 1998).

A rule base, i.e. a collection of the fuzzy 'if–then' rules, specifies the prototypical behaviour of the system under study. A typical fuzzy 'if–then' rule is composed as follows: 'If x is A and y is B and … then z is C', but rules may also use 'or' propositions. Composing a complete rule base of complex problems might be beyond the experts' capacity, especially as our understanding of the real world is incomplete (Gaines 1976). However, experts tend to be rational and might not reveal inconsistency even though it exists in reality (Weisbrod 1998). On the other hand, one might also have to deal with inconsistency between multiple experts: one way of solving this is through fuzzy evaluation (Cornelissen et al. 2001). Data-driven approaches tend to reveal restricted rule bases, but if the original database has a limited scale, the rule base might be incomplete, i.e. neglect logical remainders. Therefore, Guillaume and Magdalena (2006) proposed an integrative method to design compact and non-redundant, but consistent, rule bases. A fuzzy rule base for expert systems can be composed directly by domain experts, derived from experts' opinions or panels, or derived by data mining. To simplify the rule base, it is preferable to limit the number of alternative rules, e.g. by using constraints. All methods of rule-base composition can lead to redundant rules, but these can be pruned by means of automated procedures using algorithms. In relation to agricultural development, we are especially interested in the few individuals who are innovators because they keep abreast of developments in socio-economic and environmental context (Deutschman and Borda 1995), and therefore all prototypical cases need to be included.

Data for expert systems can be collected from long-term records, expert panels, observations, or surveys (step 7). Collecting data by surveys can be a laborious process, which is also an imposition if it does not benefit the interviewed experts directly. Therefore, most studies on social change and adoption of innovations in rural agriculture use small samples. Successful data mining requires a large database, as if smaller databases are used, exceptional cases may be missed.

Implementation involves choosing the FIS, the type of t-norm to calculate the degree of membership, the type of t-conorm to determine the combined degree of fulfilment for each rule, and the software. Our goal is to identify an inference system that maintains both the transparency and the uncertainty of the reasoning in the intermediate layers of the hierarchical FLM. Two main types of fuzzy systems, named after their developers, are available for the reasoning mechanism of the inference: Mamdani and Takagi–Sugeno–Kang (TSK) models (Jang et al. 1997). TSK models delivering a crisp (yes/no) output are very appropriate for use in data-driven procedures (ibid.). Mamdani-based FLMs deliver a fuzzy (graphical) output and are popular in low-level direct control but also appropriate for high-level hierarchical control systems and expert systems (Karray and deSilva 2004, p. 470). The fuzzy output allows to maintain the advantage of dealing with uncertainty at the intermediate layers of hierarchical FLMs and, at the final stage, to check the model's sensitivity. For the inference of rules and MFs related to fuzzy 'and' rules, the min-operator is a natural choice above algebraic, bounded or drastic product (Zimmermann 1991).

Model calibration is done through training and fine-tuning—either manually or automatically. A training

Fig. 2 An example of the asymmetrical polynomial spline-based curves open to the left (z curve) or to the right (s curve), in combination with a Pi curve that is zero at both extremes with a rise in the middle

dataset may contain 50–90% of the original dataset (Hastie et al. 2001, p. 533). Automated training does not change the rule base during calibration (Jang et al. 1997). During manual calibration models can be adjusted vis-à-vis the rules, the number of linguistic values, and the parameters of the MFs at all layers of the hierarchical FLM. Manual calibration entails running the model on a training dataset for a range of values for all variables, checking face validity and then adjusting and checking the model. Face validity is checked by comparing the model output against the real-world outcome (Sorensen 1990). The procedures are repeated until a satisfactory fit is obtained, i.e. until the face validity of the model output and calibration data is optimal. Using multiple predicates and modifiers for the variables increases the complexity of FLMs, but this can be solved by superposing proximate membership functions (MFs) to reduce the number of linguistic values, which is especially necessary after data mining (Setnes 2001). If after adjusting membership parameters during training some prototypical cases, i.e. logical remainders, are not revealed among the consequences, the rule bases should be adjusted. The last two argue in favour of using manual procedures for determining linguistic term sets, MFs and rules, and for calibrating the model.

2.2 Problem analysis and data collection

To make the DM processes explicit, i.e. to identify the stakeholders' personal context, their options, choices and general ways of reasoning, we used four steps: literature study, conversations with domain experts in order to become familiar with IAASs in the Mekong Delta, a field study to assess farmers' motivations for diversification and integration, and a data analysis. To assess farmers' motives and drives to practise a particular activity and integrate it into their family farm, in 2004, we conducted semi-open interviews in three hamlets in the freshwater floodplains of the delta, using methods of socio-technical analysis of rural livelihoods (ODG 2001). In each hamlet, 24 farmers were selected from available lists of farming households, using stratified random sampling based on wealth rankings of poor, intermediate and well-off households (Chambers 1994).

To establish trust, we started each semi-open interview by accompanying the farmer on a walk through the homestead and its neighbouring fields. After this, we mapped the farm together with the farmer, recorded its physical resources (e.g. location of fields, distances, areas, products, number of harvests per year, duration and depth of flooding) and collected data on the family composition, the present farming components and the components' internal and external relations in a resource flow diagram; e.g. see Nhan et al. (2007). The open part of the interview

followed and dealt with past changes in farm composition, the motives, or conditions under which farmers implement a change or innovation, and—if applicable—the farmer's motives for not applying other components. Subsequently, data were collected on the distance to the input and output markets, and the net income generated from each component over the past year. After a test of the interview procedures, we decided to collect financial data for ten farm household activities: irrigated field, orchard, upland, aquaculture, pigs, chickens, ducks, goats, large ruminants (buffalo, cattle), and off-farm labour.

All data were recorded on maps and in MS-Excel® spreadsheets in the form of quantitative data and qualitative information (brief farm history and decision rules for the changes). We used the capital assets framework of rural livelihoods (Carney 1998) and performed correlation analysis on the data (Bosma et al. 2006a), to assess the input variables relevant for the decision-making.

2.3 The hierarchical decision-tree

The open-ended interviews on the changes in farm composition revealed that farmers practise one or more of the 10 activities if they need to for food security, if they have the required land, water source, capital, and labour at their disposal, if they have the know-how, and/or if they consider the marketability of the product promising. Most of these features are determined by several variables, e.g. the availability of capital depends on the area of land, the other assets and the risk behaviour. If both the factors and the opportunity for a product are favourable the farmer may decide to practise only one, or several components depending on his personal context, his vision on the relationship between the components, and on his motives (Bosma et al. 2006a). As a result, the farmer's DM is represented by a three-layer hierarchical tree with five subsets: the primary production factors, the product opportunities, the product options, the farmer reference frames (FRFs) and the final output layer (Fig. 3). For each of these decision factors, several explanatory variables were identified (Table 1).

The farmers frequently mentioned two motives for change or innovation: improving income and diversifying the diet, both mainly for the well-being of their children. Therefore, we used the number of young children as an operational variable. Older farmers with no successors change the farming system to reduce the labour requirement. In the model, these driving forces were inferred in the FRF for diversification that comprised three operational variables: the number of young children in the household, the age of the household head, and the phase in the household life-course (Bosma et al. 2006a). The FRF for the integration of farm components related to six variables:

Fig. 3 Simplified structure of the hierarchical fuzzy model simulating farmers' decision-making on their farm composition: *left-hand column* shows the input variables for 18 first-layer fuzzy inference systems; *extreme right* shows the output (the third layer). *FRF* farmer reference frame, *ha* hectare, *LQI* land quality index, *dotted line* input variables for the farmers' references frames, *dashed line* ditto for the production factors, *solid line* ditto for opportunity to make a profit
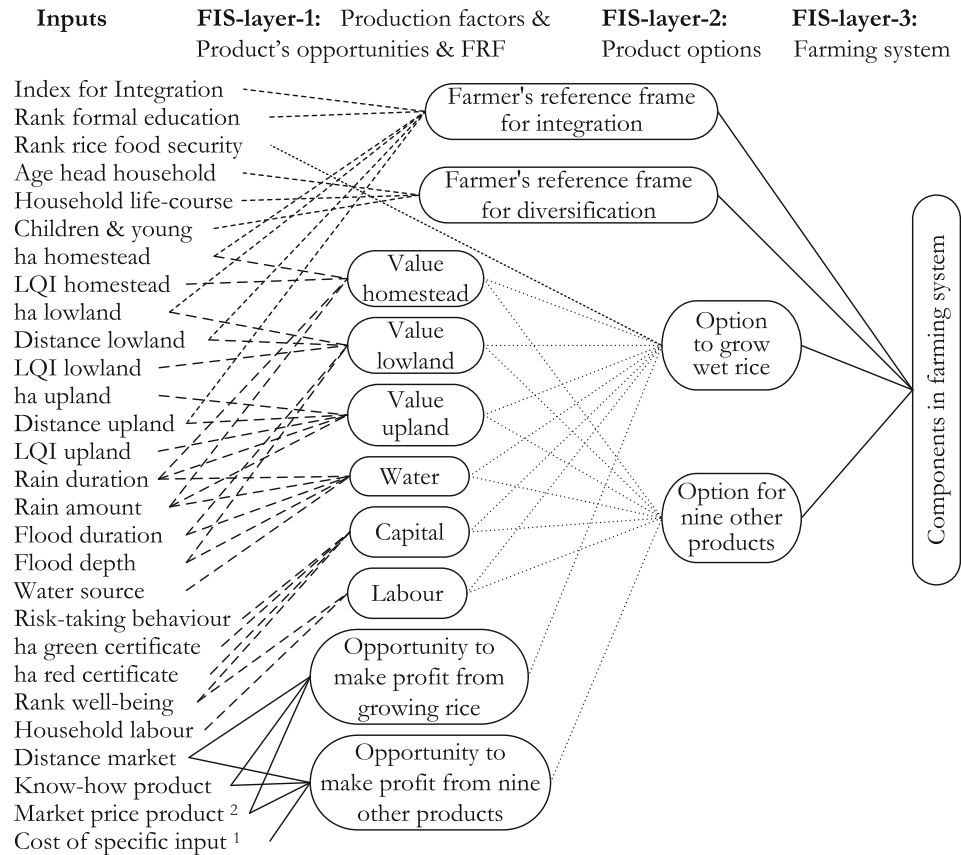
1: Cost for piglets only; 2: for chickens and ducks the prices of eggs and broilers were considered.

the distance between the fields and the homestead, the area of the homestead, the index for integration, and the farmers' level of education, i.e. the number of classes the farmer passed at school. The index of integration was calculated from the number of flows between the farm components (ibid.). The FRFs were inferred in the first layer but implemented in the third layer of the FLM. In the first layer, the variables related to the production factors and the products' opportunities were inferred, and in an intermediate layer each product's opportunity was related to all of the production factors to establish whether or not the farmer has the option to practise the component in question.

The economic drivers for innovation were assessed through the individual product opportunities. Farmers' opinions confirmed the results reported in the literature (Phong et al. 2008), that four variables influence the economic opportunity to practise a component: distance between the farm and the market, cost of inputs, market price of the produce, and the farmer's know-how on the component. We applied the prices per kilogram of product and are aware that the latter prices do not reflect the net margin of the component. Our justification for using this approach is that for crops grown and livestock raised, the

farmers are aware of the price level that resulted in break-even or a profit, or caused financial losses.

The availability of labour related to two variables: the household labour and the capacity to hire labour, which was determined by the level of income. The availability of capital did indeed depend on the collateral value of the land owned, the rank of risk behaviour, and the level of income. In the database, the level of income was represented by the rank of well-being.

The farmers' preference for having their own rice-field for food security affected their decisions about land-uses other than rice. This importance of a rice field for food security was rated from 1 to 5 by the farmers themselves. The plots of most Vietnamese farms are scattered, and each has its own characteristics relating to e.g. soil quality, water availability, and thus supports different types of activities; we took this variation into account by using three categories of land: homestead, upland, and irrigated land. The FIS of the homestead contained four input variables: its area, its soil quality, the duration of the rainy season(s) and the amount of rainfall. In addition to the variables applied for the homestead, the FIS for the upland contained the distance from the plot to the homestead. The FIS of the irrigated land also had five variables, as was the case for

Table 1 An overview of the 28 fuzzy inference systems (FIS), with the number of input and output variables, the number of linguistic terms (LT) used, and the initial number of fuzzy rules (777) and the final number (668) after trimming the non-firing rules

Level in system	Title of FIS	Input variables	Linguistic terms	Fuzzy rules		Output vars.	LT
				Initial	Final		
1	Value of irrigated land	5	4 × 2, 1 × 3	32	25	1	3
1	Value of upland	5	4 × 2, 1 × 3	26	8	1	3
1	Value of homestead	4	3 × 2, 1 × 3	17	6	1	3
1	Labour availability	2	2 × 3	9	9	1	3
1	Capital availability	4	3 × 2, 1 × 3	24	14	1	4
1	Water availability	5	4 × 2, 1 × 4	28	6	1	3
1	Opportunity for rice	3	1 × 2, 2 × 3	16	16	1	5
1	Opportunity for fruits	3	1 × 2, 2 × 3	18	15	1	4
1	Opportunity for cattle/goats	3	1 × 2, 2 × 3	18	18	1	4
1	Opportunity for fish/veg./crops	3	1 × 2, 2 × 3	18	18	1	3
1	Opportunity for pigs	5	2 × 2, 3 × 3	44	38	1	4
1	Opportunity for ducks	5	2 × 2, 3 × 3	62	54	1	3
1	Opportunity for chickens	5	2 × 2, 3 × 3	34	34	1	3
1–2	FRF for diversification	3	2 × 2, 1 × 3	10	10	1	3
1–2	FRF for integration	7	4 × 2, 3 × 3	28	19	1	4
2	Option to crop a rice field	8	6 × 2, 2 × 3	76	76	1	3
2	Option to grow upland crops	8	7 × 2, 1 × 3	21	21	1	2
2	Option to produce vegetables	7	7 × 3	43	36	1	3
2	Option to produce fruits	7	6 × 2, 1 × 3	21	21	1	3
2	Option to produce fish	7	4 × 2, 3 × 3	21	21	1	3
2	Option to produce ducks (eggs)	7	3 × 2, 4 × 3	21	21	1	3
2	Option to raise cattle	7	5 × 2, 2 × 3	21	20	1	3
2	Option to raise goats	6	4 × 2, 2 × 3	8	8	1	3
2	Option to raise chickens/pigs	5	2 × 2, 3 × 3	13	12	1	3
3	Components in the farming system	12	8 × 2, 4 × 3	81	72	1	11

the upland FIS, but the two factors related to rainfall were replaced by the duration of the flooding and the flood depth, both of which restrict the period the land can be used. The water availability related to five variables: the duration of both rainy season and flooding, the amount of rain and depth of flooding, and the source of the water.

2.4 The fuzzy inference system

Most used FLMs are composed of several FIS; FIS consist of a number of 'if then' rules relating the input and output variables where so-called membership functions are used to define the linguistic values of each variable; for an example see Table 2. We applied the so-called Mamdani inference, and the 'minimum–maximum' operators for computing the degree of membership of the rule antecedents, the degree of fulfilment of the rules, and the combined rule output. The fuzzy outputs of the FISs in the first and second layers were fed directly into the FISs of the second

and third layers of the hierarchical tree, respectively. To take account of the continuous character of most input variables and to mimic the normal distribution of most human behaviour made operational by non-continuous ratings, we represented the linguistic values by smooth curves. For the MFs of the input variables and the intermediate output variables we used either: (1) a combination of a z curve and a s curve, or (2) a combination of these z- and s curves with a Pi curve, or (3) a combination of two sigmoidal functions, one open to the left, and the second open to the right when a high input value corresponded to a low linguistic appreciation. The initial values of parameters were set using medians and quartiles of the data.

To mimic the multiple outcomes, i.e. one farmer practising several components, the output was represented by a discrete set of possible alternatives and by repeated rules having the same antecedents but different consequences. The fuzzy output of the third layer could have a value between 0 and 1; a farmer was assumed to have a particular

Table 2 Example of a FIS: the membership functions of the inputs and output, and the rule base for *Capital* needing three input values for *Wellbeing* in order to distinguish consequences and four output values to prevent the domination of the effect of *Risktaking*

Inputs: *Wellbeing* [0 4]: poor, zmf[1 1.5], medium, pimf[1 1.5 2.5 3], rich, smf[2.5 3]; *Redbook*[0 10]: small, zmf[0.2 0.5], large, smf[0.2 0.5]; *Greenbook*[0 5]; small, zmf[0.5 1]; large, smf[0.5 1]; *Risktaking*[0 5]; low, zmf[1 3]; high, smf[1 3]);

Output *Capital* [0 1]; bad, zmf[0.2 0.3]; fine, pimf[0.2 0.3 0.5 0.6], good, pimf[0.5 0.6 0.8 0.9], excellent, smf[0.8 0.9];

1='if wellbeing is poor and redbook is small and risktaking is low, then capital is bad';
2='if wellbeing is poor and redbook is small and greenbook is small and risktaking is high, then capital is bad';
3='if wellbeing is poor and redbook is small and greenbook is large and risktaking is low, then capital is bad';
4='if wellbeing is poor and redbook is large and risktaking is low, then capital is bad';
5='if wellbeing is poor and redbook is large and risktaking is high, then capital is fine';
6='if wellbeing is medium and redbook is small and greenbook is small and risktaking is low, then capital is bad';
7='if wellbeing is medium and redbook is small and risktaking is low, then capital is fine';
8='if wellbeing is medium and redbook is small and greenbook is large and risktaking is high, then capital is fine';
9='if wellbeing is medium and redbook is large and risktaking is low, then capital is fine';
10='if wellbeing is medium and redbook is large and risktaking is high, then capital is excellent';
11='if wellbeing is rich and redbook is small and greenbook is small and risktaking is low, then capital is fine';
12='if wellbeing is rich and redbook is small and greenbook is small and risktaking is high, then capital is good';
13='if wellbeing is rich and redbook is large and risktaking is low, then capital is good';
14='if wellbeing is rich and redbook is large and risktaking is high, then capital is excellent';

farm component if the membership for that output was larger than 0.5 (Bosma et al. 2005). We also calculated the centre of gravity of the graphical output of every FIS and used these as indicators during calibration. We implemented the model in Matlab®7 using the Fuzzy Logic toolbox (Mathworks 2004).

2.5 Database composition

We pre-processed data for some variables in the original spreadsheet and subsequently transferred all data for the operational variables to a matrix. The availability of household labour was derived from the weighted number of family members living on-farm in the following age categories: adult $-0.25 \times$ non-working $+ 0.5 \times$ youngster $+ 0.75 \times$ elder; because the effort that people can deliver varies according to age and a non-working person (e.g. a baby) reduces the availability of the adults. Children contributing to farm activities were classified as youngsters; grandparents still working on the farm were classified as elders. Grandparents and children not participating in work were classified as non-working.

We used the three categories of wealth as indicators for capital endowment and also for income, because they correlated significantly with the farm income (Bosma et al. 2006a). If a plot of upland or a ditch–dike-based orchard bordered on the homestead, both were considered part of the homestead. Land that flooded seasonally was classified as irrigated; flood level and duration were collected

individually. Land with a "red certificate" (which attributes rights of ownership) had a collateral value that was double that of land with a green certificate (which attributes user rights and confers obligations) (ibid.).

The selection of variables and the model's structuring revealed that after the first round of interviews, data for seven variables were lacking. During a second round of interviews in 2005, we collected data on two of these variables by asking the farmers to rate their preference for having their own rice-field for food security and their know-how on the various farming activities, on a Likert scale 1–5 (Matell and Jacoby 1971). The other five variables (soil quality, water availability, index for integration, stage in household life-course, and risk behaviour) were derived from the dataset collected during the first interviews. The soils were classified into 10 quality categories (Bosma et al. 2006b). Nine sources of water were ranked in order of diminishing availability: river, primary and secondary canal, natural source, seasonal river, rainwater reservoir, permanent well, deep well or bore-hole, and shallow well. To represent the farmers' tendency to integrate several farm components, we extracted an index of integration by counting the flows between the farm components on the bio-resource flow diagram. From the available data on the household's marital status and its age composition we determined the stage of each household's life-course (Bosma et al. 2006a). Using the data recorded on the source of credit and the activity it was used for, we classified each household's risk-taking behaviour, using six

categories: none, relatives' loan, bank loan, input providers, private money lenders or high risk credit.

In the FIS of products' opportunity (see Sect. 3.2, paragraph 2), we implemented the same distance between the farm and the input or output market for all products, though in reality this distance differed for some products. The opportunity to raise pigs, ducks and chickens was related to two types of product and the know-how and prices were represented by both specialisations: fattening and reproduction (offspring or egg). A high price for eggs was always a positive incentive for raising ducks or chickens. A high price for piglets was positive if the farmer's know-how on breeding was good, but negative if the farmer had little know-how and piglets were an input he had to buy. For pigs, we therefore used the market price for piglets to represent the cost of input, which was an exception; for none of the other activities we applied the cost of an input because we considered farmers' awareness of bad, acceptable and good prices at a particular moment for particular set of prices. The market prices applied were equal for all farmers: the average of the farm gate prices for the various product categories (Table 3). The past prices collected during the open interviews were adjusted to real values, using the inflation rates acquired from the Faculty of Economics of Can Tho University. In the available dataset, the financial outputs for goats and large ruminants were pooled due to their low frequency; the model included separate estimates of both for future use.

2.6 Calibration and fine-tuning

Calibration aims to achieve optimal fit between model result and average real world situation, and fine-tuning aims at maximising fit by calibrating individual cases. For calibration and fine-tuning, we used a training dataset of 48 cases randomly sampled from the dataset of 72 farmers; sampling was weighted for the frequency distributions for the rank of wealth.

To guide manual calibration we used face validation: i.e. we compared the model's output with the number of farmers practising the component in reality (Sorensen 1990). To take account of the farming systems' traditional economic feature, we used two thresholds for face validation: the lower threshold was the number of farmers earning cash income from a component and the upper threshold was the total number of farmers practising that component. The difference between the thresholds is the households that consume all the produce of the component themselves or that did not sell a larger asset such as ruminant or a pig, during the period in question. When a result fell between the two thresholds without appreciably affecting the fit of other outputs, we deemed the output to be a realistic fit.

For the calibration and the subsequent fine-tuning we used product prices from 2003 (Table 3). To guide the optimisation of fit, we observed the model's sensitivity by consecutively running the model for a range of values for the prices of each product and for the other variables (Ascough et al. 2005). For each of the output variables, a graph was composed for the averages of the centres of gravity and of the number of practising farmers for each activity. We optimised face validity by shifting the membership functions' parameters, adjusting the rules if shifting the parameters did not lead to a desired result, and if needed by adjusting the number of linguistic values, to obtain sensitivity and to make the model's implementation perform according to rational expectations. Output variables for which the simulated number of practising farmers did not fall between the two thresholds after calibration were individually fine-tuned using the data training set.

2.7 Validation and testing

To validate the model, we ran it on the 24 cases of the delta dataset not used for training. For performance assessment (testing) we ran the model on the dataset of 72 cases with prices for previous years. For years other than 2003, the

Table 3 Product' prices applied [×1,000 VND, per kg or head for livestock (except pigs)]

Model run	Rice	Crops	Fruit	Fish	Veg	Duck	Hen	Egg	Pig	Piglet	Lrum	Goat
1995	1.05	0.26	1.3	7.8	2.6	9.2	13.1	0.65	10,450	260	650	70
1997	1.34	0.45	2.2	8.9	4.5	11.2	16.8	0.78	16,810	450	1,120	110
1999	1.46	0.73	3.1	8.4	6.3	13.6	18.8	0.94	9,400	840	1,570	160
2003[a]	2.10	1.0	4.0	8.0	6.0	15.0	18.0	0.80	10,000	800	2,000	200

The prices for 1995, 1997 and 1999 were transformed into real values for 2003 by correcting for inflation. Annual inflation in Vietnam was close to 3% in 2003, 0.8% in 2002, −1.7% in 2000, 4% in 1999 and 1998, and estimated at 4% between 1995 and 1997

Crop crop other than rice, *Veg* vegetables, *Lrum* large ruminants (cattle, buffalo)

[a] Year of calibration and validation

farm history only revealed the total number of farmers practising a component, and only one threshold (all-practising) was used for testing.

To verify our hypotheses, the model was run for a range of values of the variables most related to the paper's objective, i.e. inclusion of the farmers' social motivations. To test the influence of their inclusion, the calibrated model was also run without implementing FRFs for innovation and integration in the third layer of the FLM.

The performance of the model was checked for face validity and quantified by calculating individual classification rates (ICRs) and overall error. The ICR of the positives is the quotient of the correctly classified number of farmers practising a specific activity on the number of farmers' actually practising this activity: $ICR^+ = \{n_{yes}\text{-type I error}\}/n_{yes}$. The ICR of the negatives is this quotient for non-practising: $ICR^- = \{n_{no}\text{-type II error}\}/n_{no}$. The model's performance is evaluated by the overall performance rate $ICR^{overall}$, which is calculated as follows: $ICR^{overall} = \sqrt{[(\{n_{yes}\text{-type I errors}\}/n_{yes}) \times (\{n_{no}\text{-type II errors}\}/n_{no})]}$.

After eliminating the non-firing rules, i.e. rules that were not activated during the various model' runs, we quantified the model's sensitivity, using the sum of the first derivatives. By lack of another comparable dataset, we ran the model on the aggregated dataset for all decimal values of the various variables, calculated the centres of gravity, the components practised by each farmer, and the average number of components each individual farmer was estimated to practise. The series of results for each variable was transposed to an MS-Excel® spreadsheet to calculate the first derivatives (∂) for the average number of components a farmer practises (∂_{NC}) (Crouch 1998). The first derivatives were averaged ($\sum \partial_{NC}/n$) and presented as a percentage indicating the relative sensitivity of the output to a specific input variable:

$$\sum \partial_{NC}/n = \left[\sum_n \{(Y_{i+1} - Y_i)/(X_{i+1} - X_i)\} \right] \Big/ n,$$

in which $i = 1, 2, 3, \ldots, n$.

3 Results

3.1 Calibration and validation

Focussing on the positive ICR, i.e. reducing the type I errors, resulted in large type II errors, which is a commonly observed trade-off during calibration. Therefore, calibration and fine-tuning addressed the reduction of both error types. The model's performance indicated by the overall ICR was too optimistic, as it approached either the best of

the ICR of positives or negatives (Table 4) and therefore we calculated an overall performance rate as given above.

Using two linguistic values in the intermediate layers of the HFS was the modeller's starting point. However, to maintain sensitivity and to obtain optimal fit during calibration, more terms were needed, for four reasons:

1. To define a constraint, i.e. a rule with one input and one output, while maintaining variation in the remaining section of the space of discourse of the rule base.
2. To define a rule base in which the effects of all the inputs were distinguishable and non-confounded (Table 2). The number of values of the input variable affects the range the output' values can take (Fig. 4).
3. To obtain acceptable effects of the variation in the original variables.
4. To be able to simulate the synergetic effect between two activities.

The number of linguistic values for the inputs of the second and third layers was not related to those of the previous layer but determined by the need to prevent domination by one of the other inputs. After calibration, the parameters of the MS of inputs for layer 2 or 3 could be different also from the outputs of layer 1 or 2, respectively (Table 5). Individual fine-tuning was needed for the activities with a small number of practising farmers.

After model testing, the rule base contained 767 rules. We eliminated the non-firing rules except those for all price levels and their subsequent product opportunities (Bosma 2007). The remaining number of rules was 658 distributed over 28 FISs.

3.2 Performance assessment

Except for ruminants, the model's estimate of the number of farmers practising the other components was intermediate between the number of all practising farmers and those farmers engaged in the activity for cash also (Fig. 5). With the exception of the estimates generated for farmers

Table 4 The ICRs of positives and negatives, the average ICR, and the overall performance rate, for the aggregated dataset (all practising farmers)

|	Rice	Fruit	Fish	Pigs	Ducks	Chickens	Ruminants
ICR of positives	0.92	0.83	0.79	0.64	0.53	0.68	0.33
ICR of negatives	0.67	0.50	0.60	0.47	0.89	0.23	0.92
Overall ICR	0.88	0.81	0.78	0.60	0.67	0.60	0.88
Overall performance	0.78	0.65	0.69	0.55	0.69	0.40	0.56
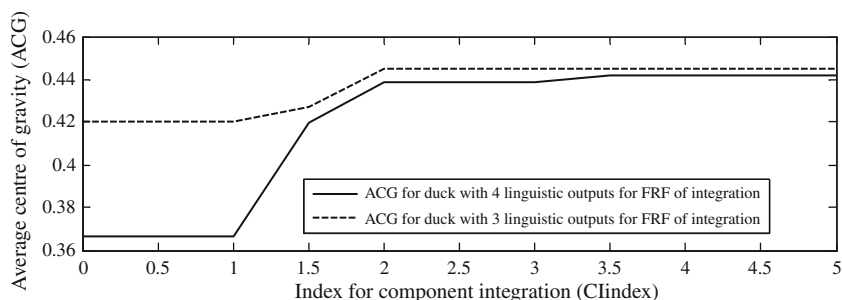
Fig. 4 An example of how increasing the number of linguistic values from 3 to 4 for an intermediate output affects the range of the final output: the Average of the Centres of Gravity (*ACG*) for the fuzzy output of the number of farms raising ducks for the Index for component integration, a factor in the farmers' reference frame (*FRF*) of integration

Table 5 Four examples of the linguistic values and the MFs' parameters at the intermediate layers (compare numbers in left and right columns)	Output of first layer	Input for second layer
	Labour = var (output, *labour*,[0 1]);	Ricefield = var (input, *labour*,[0 1]);
	Labour = mf (bad, zmf,[0.2 0.4]);	Ricefield = mf (bad, zmf,[0.2 0.4]);
	Labour = mf (fine, pimf,[0.2 0.4 0.6 0.8]);	Ricefield = mf (fine, pimf,[0.2 0.4 0.6 0.8]);
	Labour = mf (good, smf,[0.6 0.8]);	Ricefield = mf (good, smf,[0.6 0.8]);
	Capital = var (output, *capital*,[0 1]);	Ricefield = var (input, *capital*,[0 1]);
	Capital = mf (bad, zmf,[0.2 0.8]);	Ricefield = mf (bad, zmf,[0.1 0.2]);
		Ricefield = mf (fine, pimf,[0.1 0.2 0.5 1]);
	Capital = mf (good, smf,[0.2 0.8]);	Ricefield = mf (good, smf,[0.5 1]);

One example of equivalent values (*Labour*) and others examples of shifting the values in order to calibrate the model. Two example of a different numbers of inputs and outputs at the intermediate layer: an increase for *Capital* to maintain sensitivity and a decreased for *Rice field* to reduce the number of rules needed

	Output of second layer	Input for third layer
	Ricefield = var (output, *ricefield*,[0 1]);	Iaas = var (input, *rice*,[0 1]);
	Ricefield = mf (bad, zmf,[0.1 0.2]);	Iaas = mf (bad, zmf,[0.3 0.7]);
	Ricefield = mf (fine, pimf,[0.1 0.2 0.5 0.6]);	
	Ricefield = mf(good, smf,[0.5 0.6]);	Iaas = mf (good, smf,[0.3 0.7]);
	Fish = var (output, *fish*,[0 1]);	Iaas = var (input, *fish*,[0 1]);
	Fish = mf (bad, zmf,[0.2 0.4]);	Iaas = mf (bad, zmf,[0.4 0.45]);
	Fish = mf (fine, pimf,[0.2 0.4 0.6 0.8]);	Iaas = mf (fine, pimf,[0.4 0.45 0.7 0.8]);
	Fish = mf (good, smf,[0.6 0.8]);	Iaas = mf (good, smf,[0.7 0.8]);

Fig. 5 The simulated number of farmers engaging in a given activity (Simulated) versus the actual numbers of all farmers engaging in the activity (Actual-all) and those generating cash income (Actual-cash), for the 72 farmers in the Mekong delta
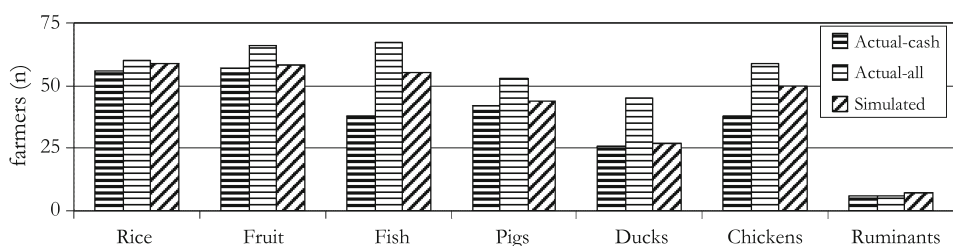

raising fish and chickens, the simulated numbers were close to the number of farmers practising the component to generate cash income. The ICRs of the positives of the training dataset for rice were similar in the training and validation datasets (90%), but in the validation set they were lower for pigs, ducks and chickens, while they were higher for fruit, fish, and ruminants. For the aggregated dataset, the ICRs of the positives were higher for the land-based activities than for livestock activities, especially for ruminants (Table 4). Except for ducks and ruminants, the ICRs of the negatives were lower than the ICRs of the positives: i.e. error type II was larger than error type I. The overall performance rate, i.e. the identification of the individual farmers engaged in (or not engaged in) a specific activity, was on average close to 75% for the land-based activities (rice, fruits and fish) and 55% for the livestock activities.

The validation for various price levels showed an overestimation of the positive trend for the number of farmers raising ruminants and chickens, and slight

Fig. 6 Comparison of trends in the % of farmers practising the component in reality (Bosma et al. 2007), and the simulated % for the price levels in four specific years (*Chick* chickens, *Rumi* cattle and goats)
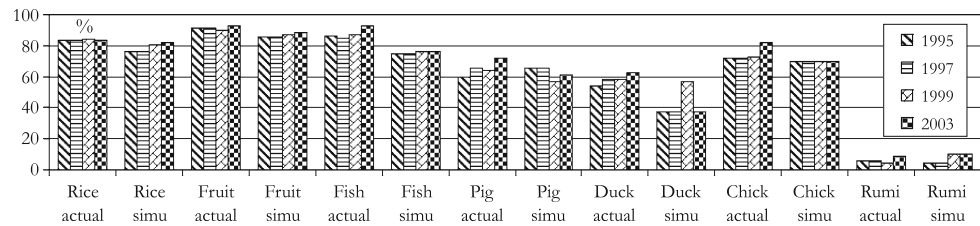
underestimations of the number of farmers having a fruit orchard, and raising fish and ducks (Fig. 6). The stagnating number of farmers cropping rice and fruits was not well simulated: the model showed 5–10% less farmers cropping rice and fruit in the past, while in reality these stayed stable.

The simulated number of farmers fattening fish and raising chickens was about 5–10% too low and the recent rising trend was not represented. According to the simulation, the number of farmers keeping ducks fluctuated, while in reality a steady increase was observed. In reality, the number of farmers raising pigs also increased, yet the simulation showed a decreasing trend. The fall in the number of farmers raising pigs due to the low price around 1999 was overestimated, while the increase of 2003 was underestimated.

3.3 Motives and drivers for diversification and integration

The inclusion of operational variables of family related motivations through the FRFs for diversification and integration improved the simulation accuracy of the number of farmers' engaged in growing or raising fruits, pigs, and ducks for generating cash income (Table 6). The implementation of both FRFs reduced the simulated numbers of practising farmers by around 10%, except for rice. Including the FRFs improved the overall performance indicator for fruit, fish, and for chickens slightly; the performance for rice, pigs, chickens, and ruminants was hardly affected.

Note that the sensitivity to the rank of well-being on the number of components practised by a farmer (Fig. 7, Table 7) was more important than the sensitivity to availability of family labour and to the total sensitivity of the three operational variables of family motivations from the FRF for diversification. The model's sensitivity to the three operational variables of family motivations from the FRF for diversification—number of young children, phase in the life-course, and age of household head—was small compared to the sensitivity to the availability of household labour. The availability of household labour is strongly related to the first two variables mentioned. The sensitivity to the attitude to risk-taking was 7%, which was higher than sensitivity to the age of the household head, but

Table 6 The actual numbers of farmers generating cash and all practising various farm components, compared to the simulated numbers with and without implementation of farmers' reference frames (FRFs), and the overall performance rate

Type of rule base	Rice	Fruits	Fish	Pigs	Ducks	Chickens	Rums
Numbers (*N* = 72)							
Actual generating cash	56	57	38	42	26	38	6
Actual all practising	60	66	67	53	45	59	6
Simulation							
With FRFs	59	58	55	44	27	50	7
Without FRFs	59	64	63	48	30	57	8
Overall performance							
With FRFs	0.75	0.59	0.61	0.44	0.50	0.33	0.32
Without FRFs	0.75	0.52	0.39	0.45	0.53	0.31	0.32

For the 2003 price level

Rums ruminants

Table 7 The sensitivity of the number of components practised to the three variables in the FRF diversification and to the classical production factors household (hh) labour and of well-being, expressed as the relative change (%), and the range of the variables

	Age head of household (hh)	Phase in the hh life course	Number of children in hh	hh labour	Rank of well-being
$\sum \partial_{NC}$	−0.6	8.9	18.4	49.3	66.0
Range	25–75	1–5	0–5	1–7.5	1–3

intermediate to the sensitivity to the household life-course (Table 7) and the level of education (Table 8).

As for the variables determining the FRF for integration, the sensitivity to the index of integration was double that of the level of education, while the sensitivities to area of lowland and distance between lowland and homestead were close to zero or slightly negative, respectively (Table 8). However, the impact of those variables was dominated by the sensitivity to the area of the homestead. This sensitivity could be direct or indirect, because this variable was also implemented in a FIS for land.

The sensitivity to the rating of the importance of having one's own rice field for food security was strong (47%), and comparable to the sensitivity to price and know-how
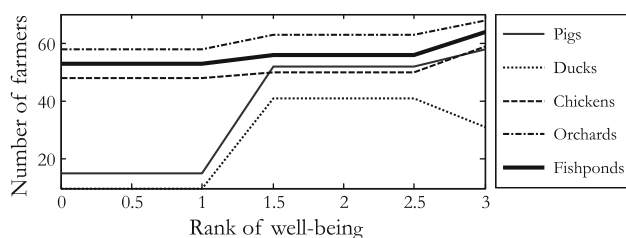
Fig. 7 Example of the result of a sensitivity analysis of changes in the rank of well-being on the number of farmers practising an activity. The effect on raising cattle and goats was also strong but is not included in the graph

Table 8 The sensitivity of the number of components practised to the five variables in the FRF of integration expressed as the relative change (%), and the range of the variables

	Area (ha) homestead	Area (ha) lowland	Distance lowland homestead	Index of integration	Level of education
$\sum \partial_{NC}$	57.4	0.8	−2.0	11.4	5.8
Range	0–3	0–3	0–5	1–5	1–5

Table 9 The sensitivity of the model, expressed as relative change (%) in the number of farm components practised due to the market prices of the components' product and to the ratings of components' know-how

Components	Relative change in number of components due to:	
	Product' price	Farmer' know-how on product
Rice	28	9.7
Fruits	0.4	0
Fish	0.8	32
Cattle	19	53
Goat	23	16
Pigs (fattening)	23	17
Piglets (breeding)	10	8.3
Ducks	30	0
Chickens	1	0
Duck eggs	23	31
Chicken eggs		1.4

for cattle (Table 9), but limited to a specific range of this rating. The average rating of this importance of a rice field for food security was 4.3. For an index below 3, the simulated number of farmers having rice fields was around 25, and this number more than doubled if the index was above 3.5. A high index reduced the number of farmers having fruit orchards, raising pigs, ducks or chickens only slightly.

The sensitivity to the market prices was high for some of the individual activities (ducks, rice, pigs, goats, cattle), but

very low for fruit, fish, chickens (Table 9), and chicken eggs. The sensitivity of the number of components practised to the farmers' rating of know-how on fruit, ducks, broilers, and laying hens was very low (Table 9). The sensitivity of the predicted number of components to farmer's know-how on raising ruminants, keeping ducks for eggs, and fattening pigs and fish was higher compared to the sensitivity of the predicted number of components to the farmer's know-how on rice. Mostly, the effect on the individual activity was reflected in the total number of components that each farmer practised. However, the sensitivity to the rating of know-how of the centre of gravity for fruit was higher (19%) than for e.g. pigs (17%), but this rating of know-how did not affect the number of farmers having a fruit orchard, nor the number of components practised.

4 Discussion

The model's sensitivity to the operational variables for farmers' family motivations, and to farmers' ratings of know-how on the components, was of the same magnitude as the sensitivity to the product's market price. This suggests that models simulating farmers' adoption of technology that do not include farmers' motives and know-how might be less reliable than generally concluded. The identification of individual practising farmers for the land-based activities (between 78 and 88%) were slightly higher than those acquired from a linear simulation of land use in the Philippines and Malaysia, 65–85% (Verburg et al. 2002), but lower than those from a FLM developed by data mining in North Thailand, 86–96% (Ekasingh et al. 2005). The high classification rates show that fuzzy logic allows using only farmers' awareness of too low, breakeven and profitable product prices instead of, e.g. production functions. Simulating the correct number of practising farmers for a particular context was simpler than improving the classification of individual farmers, especially for small numbers of practising farmers. Later we discuss the results, in particular referring to the variables selected and to the sensitivity analysis, but first on the modelling procedure and on the improvements to consider.

4.1 On manual calibration of hierarchical fuzzy systems

We calibrated the FLM manually. Methods of rule definition based on computer learning generally produce meaningless intermediate variables (Lee et al. 2003). We implemented the output of the FRFs in the third layer only, which is identical to the solution presented by Lee et al. (2003) to reduce rules in HFS. This procedure contributed

to the production of intermediate fuzzy outputs and inputs having a logical meaning and that were interpretable for each individual case by calculating the centre of gravity separately. Intermediate inputs and outputs with a logical meaning, maintain the advantage of transparency and allow the participation of stakeholders.

The manual procedures for rule-base definition, calibration and fine-tuning were complex and remain subjective; an alternative might be the integrative method of rule selection (Guillaume and Magdalena 2006). However, for the application of an automated procedure using gradient descent optimisation, more data are needed. Collecting the required information from a large numbers of farmers not only imposes on these farmers, mostly without bringing them any benefit, it is also a costly and time-consuming exercise. At this stage of the development of the approach, we used expert knowledge only, but ideally the rule base should be submitted to the scrutiny of the main stakeholders in the process. Engaging farmers in the development of the tool allows them to learn, which they consider a benefit (Doppler and Floquet 1999; Paassen 2004; Aklilu 2007). Therefore, such decision support tools should be developed in a participatory approach, needing an 11th step to develop the user-friendly interface (Karray and deSilva 2004).

The number of rules still seems massive (658), but a large part is related to the inclusion of various price levels. Fusion of some of the FISs to reduce the number of rules is not an option neither, because it would reduce the transparency for the different farm components. However, during a participatory approach farmers and scientists could opt to focus on some farm activities and leave aside those that are not relevant from a livelihood point of view at that point in time.

4.2 Improvements to consider

According to farmers in the Mekong Delta, the most crucial contextual variable for agriculture after natural disasters and credit availability, is the market price of the products (Phong et al. 2008). It was for this reason that the first step of the sensitivity analysis focussed on the product prices; however, a control of the effect of all factors was needed to make the model perform according to modellers' expectations. The inclusion of the price for piglets demonstrated the feasibility of applying input prices and prevented the model from predicting that all farmers would stop raising pigs at the price level of 1995, or from predicting that most farmers would raise pigs in the future. The underestimation of the increase in 2003 was an effect of expectations that prices for pigs would improve due to the Avian Influenza. After 2003, a decrease in pig keeping was observed, due to an increase in the price of the main

input: rice-bran (an effect of the so-called pig cycle). Together with the overestimation of the fall in the number of farmers raising pigs due to the low price around 1999, these are strong grounds for integrating the cost of crucial inputs in the model.

The overall performance of the model was lower for the activities with few practising or non-practising farmers, than for components practised by most farmers. For example, for the case of ruminants the output after fine-tuning on the training dataset contrasted with the output after fine-tuning on the smaller validation dataset but containing twice as many farmers raising ruminants. The minimum sample size should take account of the frequency of the individual events within the problem area: the fewer the events, the larger the sample size must be.

4.3 Variables and sensitivity

The most decisive factors for the model output were the classical production factors: labour, capital and land endowment, but the model's sensitivity to the product's market price was of the same magnitude as the sensitivity to the operational variables for farmers' family motivations, and to farmers' ratings of know-how on the components. Dutch and US farmers also considered the non-economic (family related) goals at least as important for their decision-making than the economic motives (Bergevoet et al. 2004). Among farmers in New Zealand, family values were also important in distinguishing farm styles (Coughenour and Swanson 1988; Fairweather and Keating 1994). This implies that the reliability of models simulating farmers' DM can be improved by considering not only utility maximisation but also a farmer's individual know-how and operational variables of his family related motivations.

The traditional economic characteristic of farmers practising a component either exclusively for home consumption or for both this and cash income had two consequences. First, the simulation of the effect of market prices was weak; i.e. in reality, farmers may continue practising an activity notwithstanding a low market price. The latter is reflected in several aspects of the validation with historical prices: e.g. the model simulated lower numbers of farmers cropping rice and fruit in the past, while in reality these stayed stable. The number of farmers practising a component for cash was probably not as constant as the number of all practising farmers shown in Fig. 4. Since 1995, overall half of the farmers have intensified or increased their existing activities: in the delta, 26% of the farmers have expanded the area of fruit trees and 18% the area of fish ponds, 8% have increased the number of pigs, and 11% are raising more chickens (Bosma et al. 2007). Such expansions were mainly due to the farmers' intention to earn (more) cash from a

component. Secondly, it remains a challenge to simulate whether a farmer with a small diversified farm will use, e.g. his fruit trees, fish, chickens or ducks to generate cash income or merely for home consumption. Avian influenza also caused the non-simulated increase in fish farming and poultry keeping and the decrease in egg prices. Contrary to expectations, the Avian Influenza epidemic of 2003 was not reflected in a reduction of the numbers of farmers raising chickens. Instead, farmers in the delta reacted by investing in fish, pigs and some even in poultry, hoping for an exceptional increase of prices and benefits (Phong et al. 2007).

The poor performance of the model in predicting farmers raising ruminants, chickens, ducks, or pigs may be due to the rule base not including all the farmers' motivations to keep these animals. For example, chickens and ducks are important for home consumption, especially when receiving guests, for offering to friends, and at ceremonies (Aklilu 2007), while pigs are important for recycling on-farm waste and produce valuable manure. In general, the institutional context, e.g. the availability of training and extension, was not represented in the model we proposed, though it was implicitly included in the individual farmer's know-how.

4.4 On linear and fuzzy decision support

In management of agriculture three horizons of decisions are distinguished: strategic, tactical, and operational. The linear simulation models exploring strategic policy options have attained a high degree of sophistication and level of calibration, but these models tend to skip the farm level, as was also done, e.g. for grasslands (Gimona et al. 2006) and as is the custom in climate models (Bussel and Ewert, personal communication, 2007, Wageningen University). As such these strategic DM support models ignore the complex interaction with human behaviour (Young et al. 2006). Coupling FLMs to multiple goal linear models, as was done to explore land use scenarios in Indonesia (Kok et al. 2006), might be an option to include subjective notions in strategic models. Motives, other than utility maximisation, have been neglected in developing such tools. We demonstrated that hierarchical fuzzy models offer an opportunity to integrate family motivations into models of farmer's DM in a transparent way.

Models to support decisions at farm level mostly focus on tactical (Aklilu 2007; Azadi et al. 2007) and operational (Fisher et al. 1997; Pelta et al. 2003) level, though some address strategic decisions (Doppler and Floquet 1999; Stroosnijder and Rheenen 2001). The question remains whether decision-support tools based on fuzzy logic can be useful to support strategic DM of farmers? Farmers in IAASs have demonstrated to be very flexible, and tools supporting decisions on their natural resource management

need to cope with this continuous learning especially as sustainability is not a fixed state but an emergent property of farming systems (Woodhill and Röling 1998). Maintaining sustainability is a continuous process of learning as every change induces a new condition that we need to manage or study (Andel 2004). This continuous learning and changing makes explorative modelling a hazardous enterprise: one cannot predict what farmers will learn, nor in the context of IAASs, to which activity they will give priority, as was shown above in the reaction to the Avian Influenza. However, they could not anticipate the reactions of consumers fearing their health which lead to a reduction in demand and price. Even sophisticated linear models may fail to include trade-offs for other stakeholders as was demonstrated during a foot-and-mouth disease epidemics in the Netherlands (Cuijpers and Osinga 2002); the predicted economic gains from the non-vaccination strategy were cancelled by stress and death in households due to massive animal culling and by reduced income from tourist industry as a consequence from roadblocks.

Above we argued that more motivations need to be assessed and cost factors integrated in order to simulate farmers' decisions related to their choices. Decisions remain personal and can be guided, among others, by training to increase farmers' know-how, and by credit programmes to provide the financial asset. Models might not be needed to support these decisions at farm level. However, participating in the development of models gives stakeholders, including farmers, an opportunity to learn (Doppler and Floquet 1999; Paassen 2004; Aklilu 2007). Model development mostly includes only a small group of farmers (Aklilu 2007), and scaling up has not yet been successful in developing countries. While easy to handle, models need to be adapted regularly to emerging needs of stakeholders which will make them money, time and expertise consuming learning devices. In a dynamic environment, scientists, extension agents and farmers need simple, flexible and cost-effective learning tools. If computer based models, whether using linear mathematics or fuzzy set theory, offer such tools needs to be assessed. We tend to agree with van Paassen et al. (2007) that the design of models for natural resource management, is a learning tool for scientists, experts and planners especially.

5 Conclusion

The satisfactory classification rates of the land-based activities for the Mekong Delta show that hierarchical FLMs can be a useful method of simulating farmers' DM, using only farmers' awareness of too low, breakeven and profitable product prices instead of production functions. We demonstrated that fuzzy logic allows taking into

account more individual motives of farmers than just utility maximisation. Using manual procedures for fine-tuning rule base and membership functions contains subjectivity, but automated procedures need more data. Collecting more data at farm level is both embarrassing and costly, especially in mixed farming systems where the minimum sample size should consider the frequency of the individual events within the problem area. The inclusion of farmers' opinions in an 11th step for the development of a strategic decision-making support tool for higher hierarchical levels, will compensate for the scientists subjectivity in the first 10 stages and be more valuable because it allows farmers to learn. We doubt whether using such models to support strategic decision-making for individual farms is efficient.

Whether or not a farmer diversifies his farm and integrates these components depends mainly on the availability of household labour, the farmer's rank in terms of well-being, and the area of the homestead, in decreasing order. Thus, the classical production factors still dominated variables such as, in decreasing order, the number of young children, index of integration, level of education, phase in the life-course, attitude to risk-taking, and age of household head, which have much less impact. The model's sensitivity to variables determining the farmers' reference frames and to farmers' ratings of their know-how, was of the same magnitude as its sensitivity to the product's market price.

References

Aklilu HA (2007) Village poultry in Ethiopia. Animal production systems/technology & agrarian development. PhD thesis, Wageningen University, Wageningen, p 178. http://library.wur.nl/WebQuery/wda/lang/1854090

Andel Jv (2004) Tussen theorie en werkelijkheid: oneindige wetenschap. Van Denderen bv, Groningen

Ascough JC II, Green TR, Ma L, Ahjua LR (2005) Key criteria and selection of sensitivity analysis methods applied to natural resource models. In: Zerger A, Argent RM (eds) MODSIM 2005, International congress on modelling and simulation, Melbourne pp 2463–2469

Azadi H, Shahvali M, Berg Jvd, Faghih N (2007) Sustainable rangeland management using a multi-fuzzy model: how to deal with heterogeneous experts' knowledge. J Environ Manag 83:236–249

Bergevoet RHM, Ondersteijn CJM, Saatkamp HW, Woerkum CMJv, Huirne RBM (2004) Entrepreneurial behaviour of Dutch dairy farmers under a milk quota system: goals, objectives and attitudes. Agric Syst 80:1–21

Bosma RH (2007) Using fuzzy logic models to reveal farmers' motives to integrate livestock, fish and crops. Aquaculture and fisheries. PhD thesis, Wageningen University, Wageningen, p 140. http://edepot.wur.nl/2499

Bosma RH, Kaymak U, Berg Jvd, Udo HMJ (2005) Fuzzy modelling of farmer motivations for integrated farming in the Vietnamese Mekong Delta. 14th IEEE international conference on fuzzy systems, Reno (Nevada, USA), pp 827–832

Bosma RH, Nam CQ, Udo HM, Verreth JA, Visser L (2006a) Agriculture diversification in the Mekong Delta: farmers' motives and contribution to livelihoods. Asian J Agric Dev 2:49–66

Bosma RH, Phong LT, Kaymak U, Berg Jvd, Udo HM, Mensvoort MEFv, Tri LQ (2006b) Assessing and modelling farmers' decision-making on integrating aquaculture into agriculture in the Mekong Delta Farmer. Neth J Agric Sci 53:281–300

Bosma RH, Cao NQ, Udo HMJ, Verreth JAJ (2007) Assessing farmers' motives for livelihood diversification in the Mekong Delta: household life cycle, virtual farm size, and index of integration. In: Zijpp AJvd, Verreth JAJ, Tri LQ, Mensvoort MEFv, Bosma RH, Beveridge MCM (eds) Fish ponds in farming systems. Wageningen Academic Publishers, Can Tho, pp 261–269

Cai J, Durkin J, Cai Q (2005) CC4.5: cost-sensitive decision tree pruning. In: Zanasi A, Brebbia CA, Ebecken N (eds) Data mining VI: data mining, text mining and their business applications. Wessex Institute of Technology, Southampton, pp 239–245

Carney D (1998) Implementing the sustainable rural livelihoods approach. In: Carney D (ed) Sustainable rural livelihoods. DFID, London, pp 3–23

Chambers R (1994) Participatory rural appraisal (PRA): analysis of experience. World Dev 22:1253–1268

Cornelissen AMG, Berg Jvd, Koops WJ, Grossman M, Udo HMJ (2001) Assessment of the contribution of sustainability indicators to sustainable development: a novel approach using fuzzy set theory. Agric Ecosyst Environ 86:173–185

Coughenour CM, Swanson LE (1988) Rewards, values and satisfaction with farm work. Rural Sociol 53:442–459

Crouch S (1998) CEM333: using Excel to find 1st and 2nd derivatives. http://wwwcemmsuedu/~cem333/POTTITRhtml. Accessed 03 Aug 2007

Cuijpers MP, Osinga KJ (2002) The position of the Dutch Farmers' Union on lessons learned and future prevention and control of foot and mouth disease. Rev Sci Tech 21:839–850

Deutschman PJ, Borda OF (1995) Communication and adoption patterns in an Andean village, San José, Costa Rica. Programa Interamericano de Informacion Popular, 1962 In: Rogers (ed) Communication and innovation, pp 253–256

Doppler W, Floquet A (1999) A participatory livelihood planning approach: farmers participation in modelling and optimisation procedures. In: Renard R, Krieg S, Lawrence P, Oppen Mv (eds) Farmers and scientists in a changing environment: assessing research in West Africa. Margraf, Cotonou, Benin, February 22–26, pp 407–418

Ekasingh BS, Letcher RA (2005) Successes and failures of attempts to embed socio-economic dimensions in modelling for integrated natural resource management: lessons from Thailand. In: Zerger A, Argent RM (eds) MODSIM 2005, international congress on modelling and simulation. MODSIM, Melbourne, pp 1–10

Ekasingh BS, Ngamsomsuke K, Letcher RA, Spate JM (2005) A data-mining approach to simulating land use decisions: modelling

farmer's crop choice from farm level data for integrated water resource management. In: Kachitvichyanukul V, Purintrapiban U, Utayopas P (eds) International conference on simulation and modelling, Nakorn Pathom, Thailand, pp 335–344

Emami MR, Goldenberg AA, Türksen IB (2000) Fuzzy-logic control of dynamic systems: from modelling to design. Eng Appl Artif Intell 13:47–69

Fairweather JR, Keating NC (1994) Goals and management styles of New Zealand farmers. Agric Syst 44:181–200

Fisher PR, Heins RD, Ehler N, Lieth JH (1997) A decision-support system for real-time management of Easter Lily (*Lilium Iongiflorum* Thunb.) scheduling and height-I. system description. Agric Syst 54:23–37

Fodor J, Perny P, Roubens M (1998) Decision-making models. In: Ruspine EH, Bonissone PP, Pedrycz W (eds) Handbook of fuzzy computation. Institute of Physics Publishing, Bristol, pp F5.1–F5.14

Gaines BR (1976) Fuzzy reasoning and the logics of uncertainty. Multiple-valued logic. In: Proceedings of 6th international symposium. IEEE Computer Society Press, Logan, pp 179–188

Geman S, Bienenstock E, Doursat R (1992) Neural networks and the bias/variance dilemma. Neural Comput 4:1–58

Gimona A, Birnie RV, Sibbald AR (2006) Scaling up of a mechanistic dynamic model in a GIS environment to model temperate grassland production at the regional scale. Grass Forage Sci 61:315–331

Guillaume S, Magdalena L (2006) Expert guided integration of induced knowledge into a fuzzy knowledge base. Soft Comput 10:773–784

Gürocak H, de San Lazaro A (1994) A fine-tuning method for fuzzy logic rule bases. Fuzzy Sets Syst 67:147–161

Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning; data mining, inference and prediction. Springer, NY

Jang J-SR, Sun C-T, Mizutani E (1997) Neuro-fuzzy and soft computing. Prentice Hall, Upper Saddle River

Karray FO, deSilva C (2004) Soft computing and intelligent systems design, theory, tools and applications. Pearson, England

Kok J-Ld, Engelen G, White R, Wind H (2006) Modeling land-use change in a decision-support system for coastal-zone management. Environ Model Assess 6:123–132

Lee M-L, Chung H-Y, Yu F-M (2003) Modelling of hierarchical fuzzy systems. Fuzzy Sets Syst 138:343–361

Liu P, Li H (2005) Hierarchical TS fuzzy systems and its universal approximation. Inf Sci 169:279–303

Luu LT (1992) The VAC system in northern Vietnam. In: ICLARM, IRRI (eds) Farmer-proven integrated agriculture aquaculture: a technology information kit. IIRR & ICLARM, Los Banos, p 183

Matell MS, Jacoby J (1971) Is there an optimal number of alternatives for Likert scale items? Study I: reliability and viability. Educ Psychol Measur 31:657–674

Mathworks (2004) Fuzzy logic toolbox (version 2) for use with Matlab. The Mathworks Inc, Natick

McCloskey A, McIvor R, Maguire L, Humphreys P, O'Donnell T (2006) A user-centred corporate acquisition system: a dynamic fuzzy membership functions approach. Decis Support Syst 42:162–185

Medasania S, Kimb J, Krishnapuram R (1998) An overview of membership function generation techniques for pattern recognition. Int J Approx Reason 19:391–417

Nhan DK, Phong LT, Verdegem M, Duong LT, Bosma RH, Little D (2007) Integrated freshwater aquaculture, crop and livestock production in the Mekong Delta, Vietnam: determinants and the role of the pond. Agric Syst 94:445–458

ODG (2001) Methods manual for fieldwork. LADDER working paper. DFID, Norwich, p 35

Paassen AV (2004) Bridging the gap: computer model enhanced learning for natural resource management in Burkina Faso. Communication and innovation studies. PhD thesis, Wageningen University, Wageningen, p 219. http://library.wur.nl/WebQuery/wda/lang/1705802

Paassen AV, Roetter RP, Keulen HV, Hoanh CT (2007) Can computer models stimulate learning about sustainable land use? Experience with LUPAS in the humid (sub-) tropics of Asia. Agric Syst 94:874–887

Pelta DA, Verdegay JL, Cadenas JM (2003) Introducing SACRA: a decision support system for the construction of cattle diets. In: Yu X, Kacprzyck J (eds) Applied decision support with soft computing. Springer, Heidelberg, pp 391–401

Phong L, Udo H, Jansen J, Bosma R, Zijpp Avd (2007) Resilience of smallholders: impact of Avian Influenza on mixed farming systems in the Mekong Delta, Vietnam. In: Camus E, Cardinale E, Dalibard C, Martinez D, Renard F, Roger F (eds) 12th AITVM. Does control of animal infectious risks offer a new international perspective? Mont Pellier, France, 20–23 August, pp 383–387

Phong LT, Udo H, Mensvoort Tv, Bosma R, Tri LQ, Nhan DK, Zijpp Avd (2008) Integrated agriculture-aquaculture systems in the Mekong Delta Vietnam: an analysis of recent trends. Asian J Agric Dev 4:51–66

Prein M (2002) Integration of aquaculture into crop-animal systems in Asia. Agric Syst 71:127–146

Ruspini EH, Bonissone PP, Pedrycz W (eds) (1998) Handbook of fuzzy computation (cited from introduction of chapter fuzzy models). Institute of Physics Publishing, Bristol

Sanh NV, Xuan VT, Phong TA (1998) History and future of farming systems in the Mekong Delta. In: Xuan VT, Matsui S (eds) Development of farming systems in the Mekong Delta of Vietnam, Ho Chi Minh City Publishing House, pp 16–80

Setnes M (2001) Complexity reduction in fuzzy systems. PhD thesis, Delft University of Technology, Delft

Sorensen JT (1990) Validation of livestock herd simulation models: a review. Livest Prod Sci 26:79–90

Stroosnijder L, Rheenen TV (2001) Agro-silvo-pastoral land use in Sahelian villages, Reiskirchen, Germany

Thornton PK, Herrero M (2001) Integrated crop-livestock simulation models for scenario analysis and impact assessment. Agric Syst 70:581–602

Türksen IB (2004) A foundation for CWW: meta-linguistic axioms. Fuzzy information 2004 (NAFIPS '04). IEEE, pp 395–400

Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura SSA (2002) Modelling the spatial dynamics of regional land use: the CLUE-S model. Environ Manage 30:391–405

Weisbrod J (1998) A new approach to fuzzy reasoning. Soft Comput 2:89–99

Woodhill J, Röling N (1998) The second wing of the eagle: the human dimension in learning our way to more sustainable futures. In: Röling N, Wagemakers M (eds) Facilitating ustainable agriculture. Cambridge University Press, Cambridge, pp 46–72

Young OR, Lambin EF, Alcock F, Haberl H, Karlsson SI, McConnell WJ, Myint T, Pahl-Wostl C, Polsky C, Ramakrishana PS, Schroeder H, Scouvart M, Verburg PH (2006) A portfolio approach to analysing complex human-environment interactions: institutions and land change. Ecol Soc 11:31

Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning-II. Inf Sci 8:301–357

Zeng X-J, Keane JA (2005) Approximation capabilities of hierarchical fuzzy systems. IEEE Trans Fuzzy Syst 13:656–659

Zimmermann HJ (1987) Fuzzy sets, decision making and expert systems. Kluwer Academic Publishers, Boston

Zimmermann HJ (1991) Fuzzy set theory and its applications. Kluwer, Boston