

## DELIVERABLE 7.2

# "Applying fuzzy theory concepts to the analysis of employment diversification of farm households: methodological considerations"

Jana Fritzscht<sup>1</sup>

| Dissemination level (see DoW p. 27-30) |   |                                     |
|--|---|-------------------------------------|
| <b>PU</b>                              | Public  | <input checked="" type="checkbox"/> |
| <b>PP</b>                              | Restricted to other programme participants (including the Commission Services)        | <input type="checkbox"/>            |
| <b>RE</b>                              | Restricted to a group specified by the consortium (including the Commission Services) | <input type="checkbox"/>            |
| <b>CO</b>                              | Confidential, only for members of the consortium (including the Commission Services)  | <input type="checkbox"/>            |

<sup>1</sup> The author gratefully acknowledges financial participation from the European Community under the Sixth Framework Programme for Research, Technological Development and Demonstration Activities, for the Specific Targeted Research Project "SCARLED" SSPE-CT-2006-044201.

The views expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.

This deliverable was internally reviewed by Judith Möllers and Matthew Gorton.

### **Abstract**

The Deliverable 7.2 (D7.2) of the SCARLED project provides methodological considerations for applying fuzzy set theory to the analysis of employment diversification of farm households. It presents a Mamdani's type fuzzy inference model and describes its application within the project's framework.

The model consists of ten variables that are grouped into the four factors: (i) necessity to diversify, (ii) internal preconditions, (iii) external preconditions, and (iv) attitudes. The coherence of these four factors with the integrated framework for the analysis of non-farm rural employment is discussed.

The model will be realised in the Fuzzy Logic Toolbox from MATLAB®. Forty four membership functions and 138 rules are going to be implemented, tested, and adapted with survey data from the five countries: Bulgaria, Hungary, Poland, Romania, and Slovenia. The final model will be used to assess the diversification potential of 15 regions in these countries.

## SCARLED Consortium

This document is part of a research project funded by the 6<sup>th</sup> Framework Programme of the European Commission.

**Leibniz Institute of Agricultural Development  
in Central and Eastern Europe (IAMO) -  
Coordinator**

Theodor-Lieser Str. 2  
06120 Halle (Saale)  
Germany

Contact person: Judith Möllers  
E-mail: scarled@iamo.de

**University of National and World Economy  
(UNWE)**

St. Town "Chr. Botev"  
1700 Sofia  
Bulgaria

Contact person : Plamen Mishev  
E-mail: mishevp@intech.bg

**Research Institute for Agricultural Economics  
(AKI)**

Zsil u. 3/5  
1093 Budapest  
Hungary

Contact person: József Popp  
E-mail: poppj@akii.hu

**Banat's University of Agricultural Sciences and  
Veterinary Medicine Timisoara (USAMVB)**

Calea Aradului 119  
300645 Timisoara  
Romania

Contact person: Cosmin Salasan  
E-mail: cosminsalasan@xnet.ro

**The University of Kent, Kent Business School  
(UNIKENT)**

Canterbury  
Kent CT2 7NZ  
United Kingdom

Contact person: Sophia Davidova  
E-mail: s.davidova@imperial.ac.uk

**Catholic University Leuven (KU Leuven)**

LICOS Centre for Institutions and Economic  
Performance & Department of Economics  
Deberiotstraat 34  
3000 Leuven.  
Belgium

Contact person: Johan Swinnen  
E-mail: jo.swinnen@econ.kuleuven.be

**Corvinus University Budapest (CUB)**

Department of Agricultural Economics and Rural  
Development  
Fövám tér 8  
1093 Budapest  
Hungary

Contact person: Csaba Csáki  
E-mail: csaba.csaki@uni-corvinus.hu

**Warsaw University, Department of Economic  
Sciences (WUDES)**

Długa 44/50  
00-241 Warsaw  
Poland

Contact person: Anna Dominika Milczarek  
E-mail: milczarek@wne.uw.edu.pl

**University of Ljubljana (UL)**

Groblje 3  
1230 Domzale  
Slovenia

Contact person: Luka Juvančič  
E-mail: luka.juvancic@bfro.uni-lj.si

**University of Newcastle upon Tyne, Centre for  
Rural Economy (UNEW)**

Newcastle upon Tyne NE1 7RU  
United Kingdom

Contact person: Neil Ward  
E-mail: neil.ward@newcastle.ac.uk

## CONTENT

|   |           |
|---|-----------|
| ABSTRACT .....  | i         |
| SCARLED CONSORTIUM.....   | ii        |
| LIST OF FIGURES .....   | iv        |
| LIST OF BOXES .....   | iv        |
| LIST OF EQUATIONS .....   | iv        |
| LIST OF ABBREVIATIONS .....   | v         |
| <br>  |           |
| <b>1 INTRODUCTION.....</b>  | <b>1</b>  |
| <b>2 FUZZY LOGIC - THE CONCEPT OF WORKING WITH IMPEEFECT INFORMATION.....</b> | <b>3</b>  |
| <b>3 THE MODEL .....</b>  | <b>5</b>  |
| <b>4 IMPLEMENTATION OF THE MODEL.....</b>                                     | <b>9</b>  |
| <b>5 TESTING, ADAPTATION, AND APPLICATION OF THE MODEL .....</b>              | <b>12</b> |
| <b>6 SUMMARY.....</b>   | <b>14</b> |
| <br>  |           |
| LIST OF REFERENCES .....  | 15        |

### LIST OF FIGURES

|   |    |
|---|----|
| Figure 1: An integrated framework for the analysis of non-farm rural employment.....                        | 6  |
| Figure 2: Structure of the fuzzy logic diversification model .....  | 7  |
| Figure 3: Overview of the Mamdani's type fuzzy inference system .....                                       | 9  |
| Figure 4: Graphical representation of the membership functions for the variable<br>"dependency ratio" ..... | 10 |
| Figure 5: Output of fuzzy inference.....  | 11 |
| Figure 6: Structure of the fuzzy logic model for assessing the regional diversification<br>potential .....  | 13 |

### LIST OF BOXES

|                                       |   |
|---------------------------------------|---|
| Box 1: Hypothetical interview 1 ..... | 1 |
| Box 2: Hypothetical interview 2 ..... | 1 |
| Box 3: Hypothetical interview 3 ..... | 2 |

### LIST OF EQUATIONS

|   |    |
|---|----|
| Equation 1: Crisp set.....  | 3  |
| Equation 2: Fuzzy sets .....  | 4  |
| Equation 3: Mathematical representations of the membership functions for the variable<br>"dependency ratio" ..... | 10 |

## LIST OF ABBREVIATIONS

|           |  |
|-----------|--|
| AKI       | Research Institute for Agricultural Economics  |
| COA       | Centre of area method  |
| COG       | Centre of gravity method   |
| CUB       | Corvinus University Budapest, Department of Agricultural Economics and Rural Development |
| IAMO      | Leibniz Institute of Agricultural Development in Central and Eastern Europe              |
| KU Leuven | Catholic University Leuven   |
| MOM       | Mean of maxima method  |
| SCARLED   | Structural change in agriculture and rural livelihoods                                   |
| SLF       | Sustainable Livelihood Framework   |
| UL        | University of Ljubljana  |
| UNEW      | University of Newcastle upon Tyne, Centre for Rural Economy                              |
| UNIKENT   | University of Kent, Kent Business School   |
| UNWE      | University of National and World Economy   |
| USAMVB    | Banat's University of Agricultural Sciences and Veterinary Medicine Timisoara            |
| WUDES     | Warsaw University, Department of Economic Sciences                                       |

## 1 INTRODUCTION

Everyone who has ever done empirical work knows that the information one gets from respondents is often imperfect. Different cases of imperfection can be observed. The first and surely most obvious one is that the respondents do not know the correct answer (Box 1).

### Box 1: Hypothetical interview 1

Interviewer: I would like to know what the highest degree of formal schooling of your household members is.

Respondent: I finished primary school and my wife too. My son was at the university and my daughter finished an apprenticeship as secretary last year.

Interviewer: Do you know which degree your son got at the university?

Respondent: He was five years there and he got a certificate.

Interviewer: Was it a bachelor or a master degree?

Respondent: No, no, it was a university degree.

In the hypothetical interview 1 (Box 1) the respondent does not know the exact university degree of his son because he is not familiar with the concept of Bachelor and Master level. For the father it is simply a university degree. The interviewer has now different options. He could leave an empty space and risk that he will be rebuked for not being able to get answers to simple questions. Without doubt, most interviewers will try to avoid this. Hence, there are two options left. He could force the father forward to get an answer by explaining the differences between Bachelor and Master degree. The answer will most probably be an estimation coloured from the wish of a proud father to have provided his son with the best education he could get. The second option that the interviewer has is to use the information that the son had studied for five years to decide that it is a Master degree. In both cases the information is imperfect.

The second and most delicate case of imperfection is that the respondents are reluctant to give correct figures although they know them (Box 2).

### Box 2: Hypothetical interview 2

Interviewer: I would like to know something about your household income. You said that you have a car and that you offer from time to time transport service to other persons in your village. Could you tell me please how much money you earn with this service in an average year?

Respondent: I do not always ask for money. Sometimes, I take what the passenger freely offers. I think it is not really a business and therefore there is no income.

Interviewer: Yes, I understand. But how much do you get approximately per month from the service?

Respondent: Oh, last month was not a good one.

Interviewer: What is a bad month?

Respondent: Well, it was ten units.

In the hypothetical interview 2 the respondent knows how much he earns with his taxi service but he does not like to provide detailed figures. The answer "ten units" is obviously not exact but the interviewer has no other option but to accept it. This kind of problematic answers are typical and will occur in a similar way whenever the issue of income data is touched. Resulting, the information in the data base is imperfect in the way that the income data are biased.

The last case of imperfection are qualitative answers ranging from e.g. "very good" to "very bad" or "low" to "high".

**Box 3: Hypothetical interview 3**

Interviewer: How good are the streets in your region?

Respondent: Five years ago, I had to pay much money because my wife hit a pothole and spoiled the axle. But last year the street was reconstructed and now I think it is good.

Contrary to the first two hypothetical interviews the information given in the hypothetical interview 3 (Box 3) is by its very nature imperfect because everyone may understand different things under "good streets". The same street could be rated as "poor" from someone who is used to Italian motorways or "very good" from people who are used to farm tracks. For analytical purposes such information is usually coded into a rating scale from one to five. The resulting value pretends a precise figure that in fact simply does not exist. Nevertheless, such information is often used because people feel comfortable answering them and thus the survey faces fewer problems to motivate the respondents to keep interested and concentrated.

Notwithstanding the vagueness of the information, the collected data are used in econometric and simulation models as precise data and the results are generalised. No doubt that this methodology gained good results in the last decades but it would be appealing to apply a methodology that considers the vagueness of information in the estimation routine. Such a methodology is fuzzy logic.

This report is structured as follows: (i) Chapter 2 gives a short introduction into the rational of fuzzy logic, (ii) Chapter 3 presents the model that will be implemented to assess the non-farm income diversification in the survey countries, (iii) Chapter 4 explains how the model will be implemented, (iv) Chapter 5 describes the testing, adaptation, and application of the model within the SCARLED project, and (v) Chapter 6 summarises the main outcomes of the previous chapters.



## 2 FUZZY LOGIC - THE CONCEPT OF WORKING WITH IMPEEFECT INFORMATION

Fuzzy logic gained increasing prominence in the last decade. One reason may be the rapid development of hardware and software that makes it comfortable to develop and run fuzzy logic systems. However, the basic article for the concept of fuzzy logic dates back more than forty years. In 1965 Lofti A. Zadeh published his article "Fuzzy sets" and became the father of the fuzzy set theory. This theory opened the opportunity to include imperfect information into precise data processing routines. It should be mentioned that it is not the methodology that is fuzzy but the data that is processed. The methodology itself is rooted in non-fuzzy mathematics.

Three kinds of imperfection are distinguished: (i) vagueness, (ii) imprecision, and (iii) uncertainty (Kruse et al. 1995). Information is vague when it could be interpreted from different people or in varying contexts in different ways. Linguistic statements like the one in the hypothetical interview 3 (Box 3) are vague data. Information that cannot be observed with optional accuracy is called imprecise. The income data are the most prominent example for this kind of imperfection in empirical research in economics (hypothetical interview 2 in Box 2). Uncertain information is subject to random events like lottery results or caused by subjective estimations. The answer given in the hypothetical interview 1 (Box 1) could be rated as uncertain information.

Whatever the kind of imperfection is, all imperfect information share the characteristic that they cannot be rated as true or false but as partially true and partially false. Classical set theory allows only for true or false statements and operates with so-called **crisp sets** - each datum belongs only to one set, i.e. the statement that the street's condition is "good" implies that the datum belongs to 100% to the set "good" (Equation 1).

### Equation 1: Crisp set

$$\chi_{good}(street\_condition) = \begin{cases} 1 & \text{if } street\_condition = "good" \\ 0 & \text{if } street\_condition \neq "good" \end{cases}$$

Source: Own equation.

But what is needed for processing imperfect information are sets to which a datum belongs only to a certain degree, i.e. the street's condition may be to 80% "good", to 10% "average", and to 10% "very good". This results in what is called **fuzzy sets**, i.e. the datum belongs to 80% to the set "good", to 10% to the set "average", and the 10% to the set "very good" (Equation 2).

Equation 2: Fuzzy sets

$$\chi_{average}(street\_condition) = \begin{cases} 0.0 & \text{if } street\_condition = \text{"very poor"} \\ 0.1 & \text{if } street\_condition = \text{"poor"} \\ 0.8 & \text{if } street\_condition = \text{"average"} \\ 0.1 & \text{if } street\_condition = \text{"good"} \\ 0.0 & \text{if } street\_condition = \text{"very good"} \end{cases}$$

$$\chi_{good}(street\_condition) = \begin{cases} 0.0 & \text{if } street\_condition = \text{"very poor"} \\ 0.0 & \text{if } street\_condition = \text{"poor"} \\ 0.1 & \text{if } street\_condition = \text{"average"} \\ 0.8 & \text{if } street\_condition = \text{"good"} \\ 0.1 & \text{if } street\_condition = \text{"very good"} \end{cases}$$

$$\chi_{very\_good}(street\_condition) = \begin{cases} 0.0 & \text{if } street\_condition = \text{"very poor"} \\ 0.0 & \text{if } street\_condition = \text{"poor"} \\ 0.0 & \text{if } street\_condition = \text{"average"} \\ 0.1 & \text{if } street\_condition = \text{"good"} \\ 0.9 & \text{if } street\_condition = \text{"very good"} \end{cases}$$

Source: Own equation.

The degree to which a datum belongs to the various sets is defined by the so-called membership functions. The membership functions are at the core of fuzzy set theory and their definition could be seen as the most delicate task in developing a fuzzy logic system. The fuzzy sets are subject to mathematical operations that result in a crisp output.

Sivanandam et al. (2007) quotes many applications of fuzzy logic. Most prominent are the industrial and control applications but fuzzy logic also encroached upon expert systems. Smithson and Verkuilen (2006) give an overview of fuzzy logic applications in social sciences. So far, no fuzzy model that simulates the household decision to diversify into non-farm activities is known. Work package 7 of the SCARLED project aims to develop such a model. Compared to common testing statistical methods the advantages of fuzzy logic are as follows (MathWorks 2001: 1-4 and 1-5):

- Fuzzy logic is conceptually easy to understand.
- Fuzzy logic is flexible.
- Fuzzy logic is tolerant of imprecise data.
- Fuzzy logic can model nonlinear functions of arbitrary complexity.
- Fuzzy logic can be built on the top the experiences of experts.
- Fuzzy logic can be blended with conventional control technique.
- Fuzzy logic is based on natural language.

Especially the third, fifth, and seventh point motivate the SCARLED team to apply fuzzy logic.

### 3 THE MODEL

Buchenrieder et al. (2007) discussed diversification in the theoretical context of the sustainable livelihood framework (SLF) and the so called demand-pull and distress-push concept (Efstratoglou 1990, Barrett et al. 2001, Buchenrieder 2005, Möllers and Buchenrieder 2005, Möllers 2006). Since both concepts do not address the diversification decision itself, they complemented their analytical framework with the theory of planned behaviour (Ajzen 1985). In the result of the discussion an integrated framework (Figure 1) for the analysis was proposed. The framework is too comprehensive to implement in a fuzzy logic model that is developed from scratch in this project. Nevertheless the variables for the model are selected in a way that the final model touches all aspects of the discussed theories.

Ten variables will be used to determine the potential of a household to diversify into non-farm activities. To keep the model's structure comprehensive these variables will be grouped into the four factors: (i) necessity to diversify, (ii) internal preconditions, (iii) external preconditions, and (iv) attitudes (Figure 2). In the following all factors and their variables will be described. Each factor and its variables are shortly discussed in the integrated framework of Figure 1.

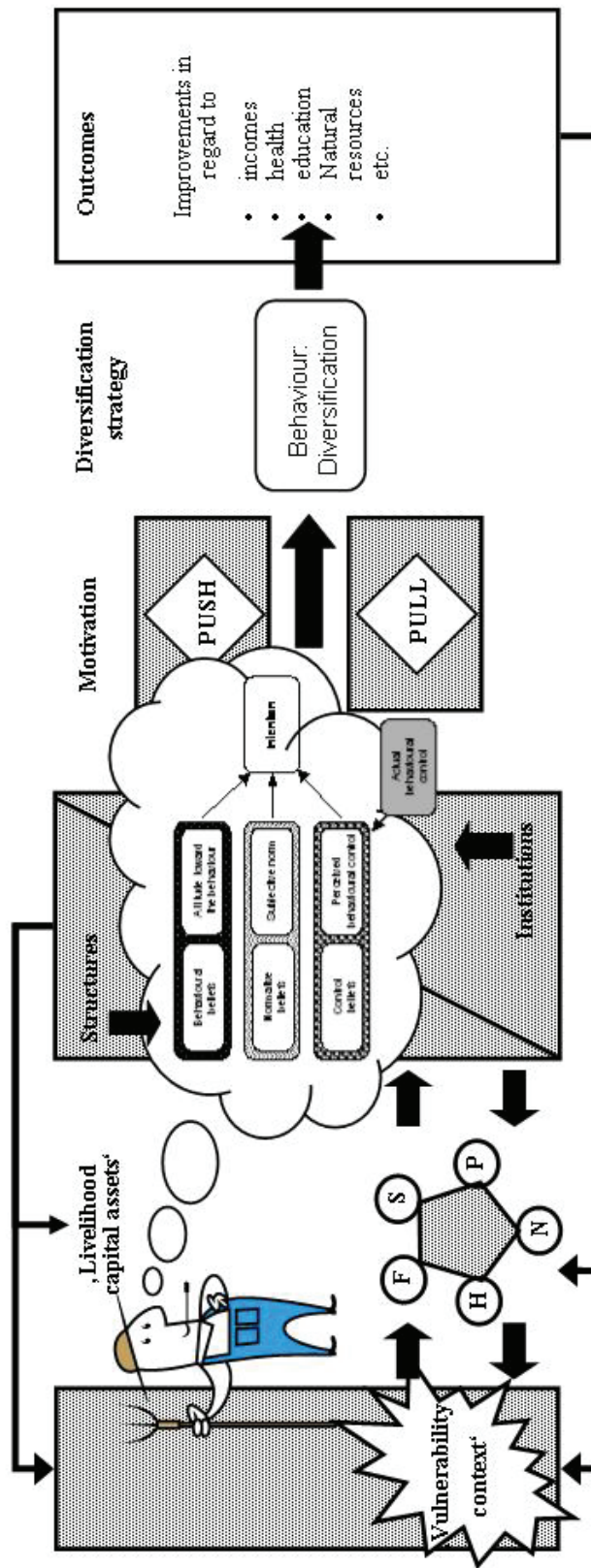
The necessity to diversify is seen as a distress-push factor. It depends on the income that a household can achieve from agricultural activities and the number of household members that have to be supported from this income. Households with a high agricultural income and few dependent household members feel less pressure to diversify because the income is sufficient to cover all expenditures. As an indicator for the agricultural income that the household could earn, the farm size could be used in the model.<sup>1</sup> The farm size measured in available land stands for natural assets in the SLF. The second variable that determines the necessity to diversify in the model is the dependency ratio. It is supposed that it is not primarily the number of household members that pushes a household into non-farm diversification but the relation of dependent household members to economically active ones. Economically active persons could migrate and sustain themselves but especially children and sometimes pensioners<sup>2</sup> do not have this opportunity and must be supported from the economically active household members.

---

<sup>1</sup> Whether or not this indicator is the best will be discussed with the SCARLED partners in the survey countries before a final decision is taken.

<sup>2</sup> In some cases, the so called unearned income of pensioners, i.e. the pensions exceeds the factor income from labour, land, and capital. Whether or not it is sufficient to sustain a decent livelihood by itself is not always clear.

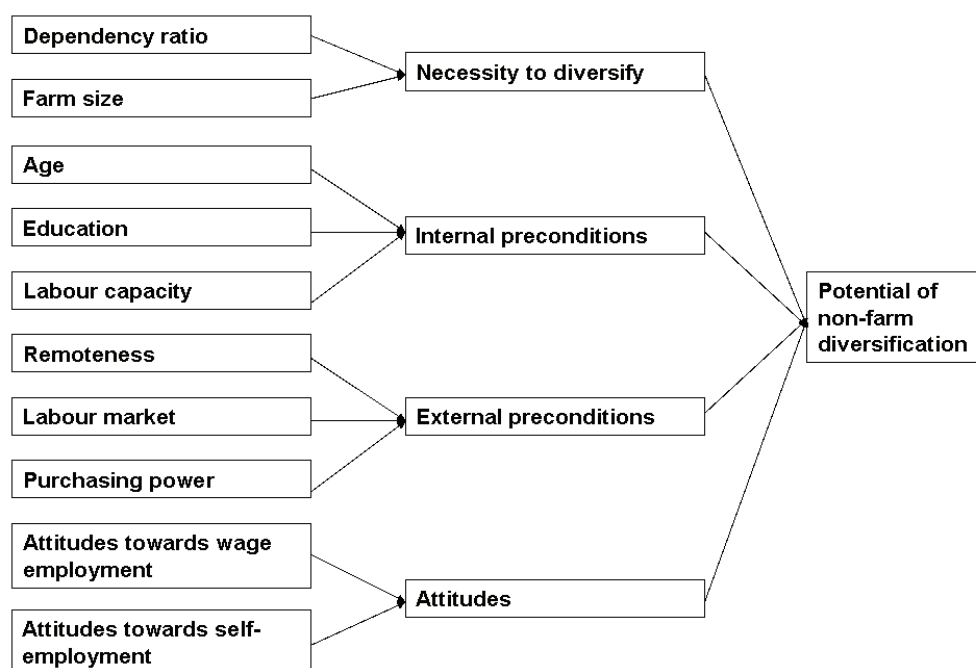
Figure 1: An integrated framework for the analysis of non-farm rural employment



Source: Möllers (2006)

Internal preconditions describe the ability of a household to diversify. They stand for human capital in the SLF and work as a switcher in the demand-pull and distress-push concept because they determine whether a household could grab favourable opportunities to earn a higher income or whether it will persist in low income activities. It is unquestionable that elderly people do not tend to alter their living situation. But even if they should have the wish to find a job, they will usually find themselves confronted with labour market constraints. But age is not the only limiting variable; also people with a deficient education may find it difficult to get a wage job or to start up an own business due to insufficient skills. Last but not least, the labour capacity of a household determines its ability to earn additional income. Wage-employment in rural regions usually implies commuting long distances and also self-employment normally means more than eight hour daily work. Whether it is a wage job or a self-employed activity, long absence on business is usually the result. Households with small children or elderly people in need of care must have at least two economic active persons to save the labour capacity for non-farm diversification. In the model the variables age, education, and labour capacity determine the internal preconditions of a household to diversify.

**Figure 2: Structure of the fuzzy logic diversification model**



Source: Own figure.

The factor external preconditions describes the possibility to diversify. It summarises variables that define a demand-pull environment. The labour market is part of the institutional framework under which a household operates and remoteness is caused by an unfavourable basic infrastructure that belongs to the physical assets in the SLF. The key question for external preconditions is whether there is a demand for paid labour or additional products in the respective region. Citizens of remote areas face difficulties to

get a wage job despite their willingness to commute and migrating may not be an option. Even for self-employed activities there may be, on the one side, only limited market capacities to earn a decent income and, on the other side, it may be difficult to attract skilled employees. Besides the remoteness of the village, the situation on the labour market is used as an indicator for wage job opportunities in the model. The local demand for new products or services from profit-oriented business will be measured in the regional purchasing power.

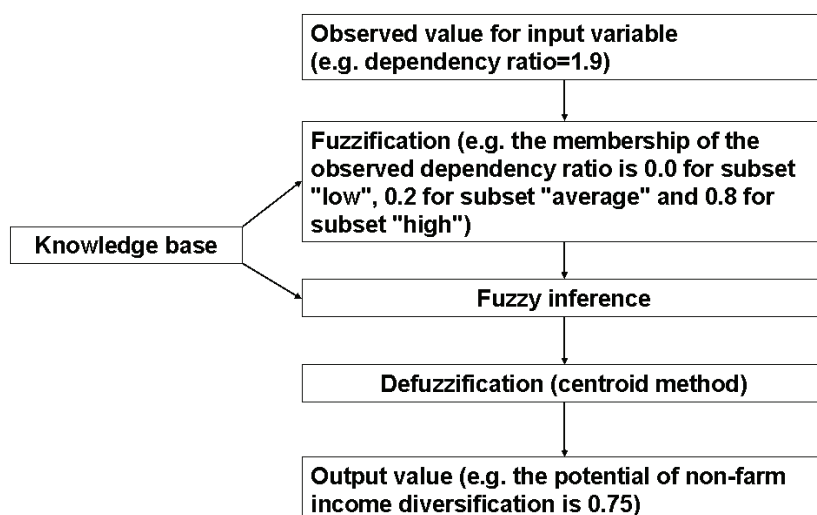
Even if the household sees the necessity to earn an additional income, the internal preconditions are positive, and the external preconditions make diversification possible the decision what is actually done depends to a high degree on what is called socio-psychological factors. Is there a culture of entrepreneurship or is it rather the civil servant with a pension who is admired? What says the old patriarch when his granddaughter migrates into the big city? Will the schoolmates stay friends when one becomes an entrepreneur? What will the neighbours do when one earns more money than the rest of the villagers? This is a large field and it is not the focus of this model to be exhaustive in the used variables. It is assumed that all these factors result in attitudes towards a certain activity. By the factor attitudes the psychological aspects of diversification decisions is represented in the model. Since the attitudes towards self-employment may be diametric to the ones towards wage employment both attitudes are used into the model.

All four factors in their various combinations determine the potential that a household has for diversification of its income activities.

## 4 IMPLEMENTATION OF THE MODEL

The model introduced in Chapter 3 will be implemented as a Mamdani's type fuzzy inference systems (Mamdani and Assilian 1975) using MATLAB® with the module Fuzzy Logic Toolbox. A graphical overview of the system is given in Figure 3. The core of the system is the knowledge base. It includes for each variable the codomain, the number of fuzzy subsets, a linguistic term and the membership function for each fuzzy subset, and the rules for fuzzy inference. The membership functions and the rules will be set according to the project team's expert knowledge. Forty four membership functions and 138 rules are planned at the current state of work.

Figure 3: Overview of the Mamdani's type fuzzy inference system



Source: Own figure in adaptation of Kruse et al. (1995: 164).

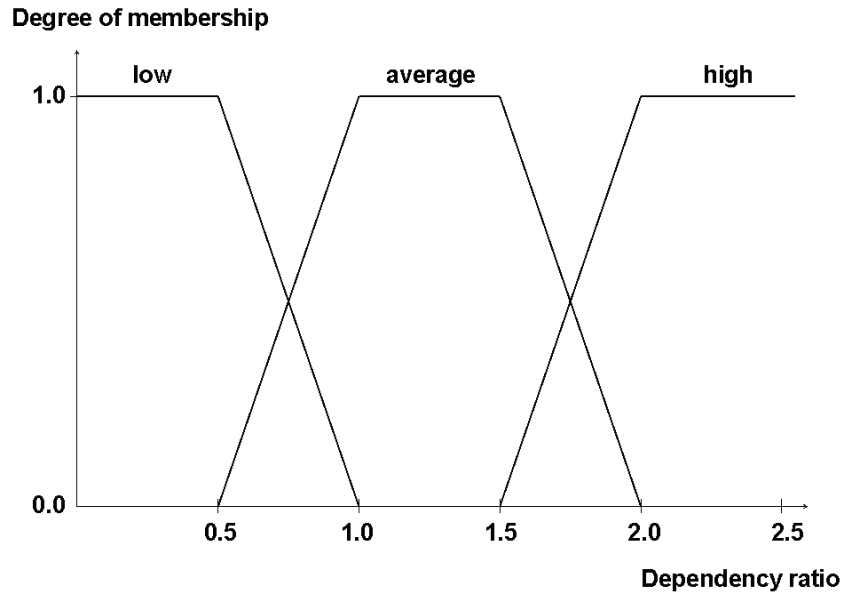
The system works in the three steps:

1. Fuzzyfication,
2. Fuzzy inference, and
3. Defuzzification.

The input values for the system are crisp, e.g. the observed dependency ratio in the hypothetical household here is 1.9.

In the **fuzzification** step the degree of membership for an input value in the defined fuzzy subsets of the respective variable is determined according to the membership functions (Figure 4 and Equation 3). An observed dependency ratio of 1.9 has a membership degree of 0.0 in the fuzzy subset "low", of 0.2 in "average", and of 0.8 in "high".

**Figure 4:** Graphical representation of the membership functions for the variable "dependency ratio"



Source: Own figure.

**Equation 3:** Mathematical representations of the membership functions for the variable "dependency ratio"

$$\mu_{low}(dep\_ratio) = \begin{cases} 1 & \text{if } dep\_ratio < 0.5 \\ -2.0 * (dep\_ratio - 1.0) & \text{if } 0.5 \leq dep\_ratio \leq 1.0 \\ 0 & \text{if } dep\_ratio > 1.0 \end{cases}$$

$$\mu_{average}(dep\_ratio) = \begin{cases} 2.0 * (dep\_ratio - 0.5) & \text{if } 0.5 \leq dep\_ratio \leq 1.0 \\ 1 & \text{if } 1.0 < dep\_ratio < 1.5 \\ -2.0 * (dep\_ratio - 2.0) & \text{if } 1.5 \leq dep\_ratio \leq 2.0 \\ 0 & \text{if } dep\_ratio < 0.5 \text{ or } dep\_ratio > 2.0 \end{cases}$$

$$\mu_{high}(dep\_ratio) = \begin{cases} 0 & \text{if } dep\_ratio < 1.5 \\ 2.0 * (dep\_ratio - 1.5) & \text{if } 1.5 \leq dep\_ratio \leq 2.0 \\ 1 & \text{if } dep\_ratio > 2.0 \end{cases}$$

Notes: The membership functions refer to the three fuzzy subsets "low", "average", and "high" for the variable dependency ratio in Figure 4.  
dep\_ratio: dependency ratio

Source: Own equations.

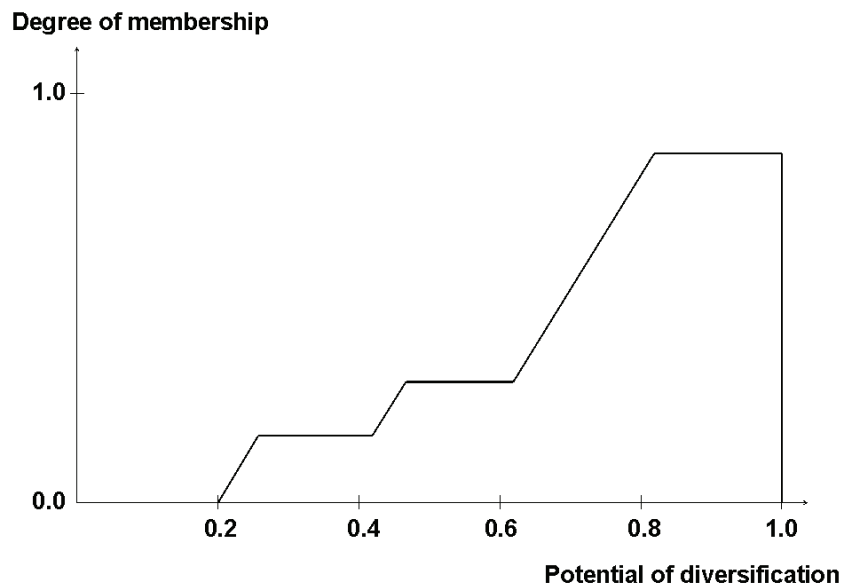


**Fuzzy inference** is the calculation step of the system. It analyses all rules and results in a fuzzy set for the output variable (Figure 5). Rules link the input variables with the output variable and have the following form:

IF dependency ratio IS high AND farm income IS low THEN diversification IS high

The **defuzzification** step transforms the fuzzy output set into a crisp value for the output variable, e.g. the potential that the household diversifies into non-farm activities is 0.75. In this project the centroid method will be used for defuzzification. This method determines the value that divides the area under the curve into two halves as output value (MathWorks 2007). In literature, the centroid method is often synonymously used with the centre of gravity method (COG) and the centre of area method (COA) (Kruse et al. 1995).

**Figure 5: Output of fuzzy inference**



Source: Own figure.

## 5 TESTING, ADAPTATION, AND APPLICATION OF THE MODEL

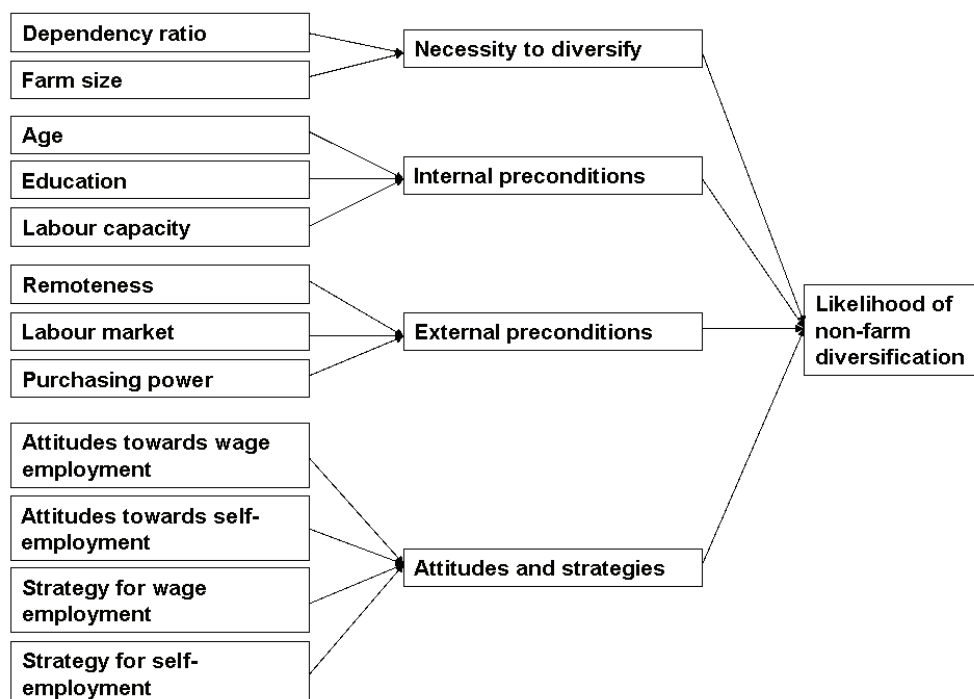
After the model is implemented, it will be tested separately with survey data from Poland, Hungary, Slovenia, Romania, and Bulgaria. The conformance of observed diversification behaviour with the simulated potential of diversification will be used as criterion for model quality.

In a next step, the model will be improved so that the conformance of observed behaviour and simulated results becomes higher. There are three opportunities to adapt the model. Firstly, the defuzzification method will be altered from the centroid method to the mean of maxima method (MOM). This defuzzification method returns the value at which the horizontal line that stands for the highest membership is divided into two halves (Nguyen and Walker 2000). Using the fuzzy output set from Figure 5, the result for the non-farm diversification potential is 0.9. Secondly, the membership functions will be revised. This will without doubt be the main work in adapting the model because it can be assumed that the membership functions are different in each of the five survey countries. Finally, the rules will be checked. Only minor adaptations are expected for the rules because they should be the same in all five survey countries. If they vary substantially across countries, the model does not represent the determinants of non-farm diversification. For the improvement of the model, 50% of the surveyed households will be used. The new conformance of observed behaviour with the simulated one will be calculated using the 50% of the sample that was not used for the adaptation procedure. The calculated conformance indicator shows how well the model pictures the reality.

The model will be applied within the project to estimate the diversification potential of 15 survey regions. This implies further calculations because not every household that has the potential to diversify will actually use it. It is assumed that it is a combination of a household's potential to diversify and its strategy that determines the likelihood of future diversification.

Two methodologies for combining the diversification potential with the household strategy are applicable. The simplest would be to multiply the potential figure with an index for the household strategy ranging from zero to one. A more advanced procedure would be to include the household strategy into the fuzzy logic model (Figure 6). Without doubt, the latter is the more appealing methodology. It will result in six additional membership functions. The rules for the factor attitudes will be adapted to the factor attitudes and strategies but no additional rules are necessary. The drawback of this methodology is that the resulting adapted model could not be verified with survey data. Nevertheless it is the proposed procedure that will be applied to assess the diversification potential of three regions in the five survey countries.

**Figure 6: Structure of the fuzzy logic model for assessing the regional diversification potential**



Source: Own figure.

The procedure will be demonstrated using the survey households that are not diversified. For these households the non-farm diversification potential is known from the fuzzy logic model (Figure 2) and the chosen household strategy was asked in the surveys. The share of households for which diversification is likely in the number of non-diversified households indicates the diversification potential of a region.

It has to be acknowledged that the resulting indicator is only truly representative for the region from which the sample is drawn. Whether or not the results can be generalised to the whole country has to be seen, as generalising recommendations would rely on a relative small subsample of 300 households per country.

## 6 SUMMARY

Information from survey data are often imperfect. This is caused by lacking knowledge of the respondents on the respective case in question, his/her unwillingness to give exact figures, and the qualitative nature of questions. Nevertheless, the information is used as precise data in econometric models. It would be appealing to have a methodology that allows explicitly for imperfect information in the calculation routine. Such a methodology is known as fuzzy logic and roots in the fuzzy set theory from Zadeh (1965). Fuzzy sets allow information to be partially true and false. The resulting fuzzy sets are subject to precise mathematical operations that result in a non-fuzzy output value.

In the project a Mamdani's type fuzzy inference system for assessing household's potential for non-farm diversification will be implemented. The model consists of ten variables grouped into the four factors: (i) necessity to diversify, (ii) internal preconditions, (iii) external preconditions, and (iv) attitudes. These factors cover key concepts from the SLF, the demand-pull and distress-push concept (Efstratoglou 1990, Barrett et al. 2001), and the theory of planned behaviour (Ajzen 1985).

The model will be tested and adapted using survey data from Poland, Hungary, Romania, Bulgaria, and Slovenia. The final model will be used to assess the diversification potential of 15 regions in the survey countries.

## List of references

- Ajzen, I. (1985). From intentions to actions: A theory of planned behaviour. In *Action-control: From cognition to behaviour*, Kuhl J., and J. Beckmann (eds.): 11-39, Heidelberg, D: Springer Verlag.
- Barrett, C., Reardon, T., and T. Webb (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: Concepts, dynamics, and policy implications. *Food Policy* 26 (4): 315-331.
- Buchenrieder, G. (2005). Non-farm rural employment - review of issues, evidence and policies. *Quarterly Journal of International Agriculture* 44 (1): 5-18.
- Buchenrieder, G., Möllers, J., Happe, K., Davidova, S., Fredriksson, L., Bailey, A., Gorton, M., Kancs, d'A., Swinnen, J., Vranken, L., Hubbard, C., Ward, N., Juvančič, L., Milczarek, D., Mishev, P. (2007). Conceptual framework for analysing structural change in agriculture and rural livelihoods. Deliverable 2.1 of SCARLED (Structural Change in Agriculture and Rural Livelihoods), Specific Targeted Research Project (STREP), 6<sup>th</sup> Framework Program for Research Contract SSPE-CT-2006-004420. Halle, D: Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO).
- Efstratoglou-Todoulou, S. (1990). Pluriactivity in different socio-economic contexts: A test of the push-pull hypothesis in Greek farming. *Journal of Rural Studies* 6 (4): 407-413.
- Kruse, R., Gebhardt, J., Klawonn, F. (1995). Fuzzy-Systeme. Stuttgart, D: B.G. Teubner.
- Mamdani, E.H. and Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies* 7 (1): 1-13.
- MathWorks (2001). Fuzzy Logic Toolbox User's Guide Version 2. Natick, MA, USA: The MathWorks, Inc.
- MathWorks (2007). Defuzzification methods. Online support system of MathWorks, <http://www.mathworks.de/products/demos/shipping/fuzzy/defuzzdm.html>, last access 23<sup>rd</sup> November 2007.
- Möllers, J. (2006). Außerlandwirtschaftliche Diversifikation im Transformationsprozess - Diversifikationsentscheidungen und -strategien ländlicher Haushalte in Slowenien und Mazedonien. Studies on the Agricultural and Food Sector in Central and Eastern Europe, Vol. 35, Halle (Saale), Germany.
- Möllers, J. and Buchenrieder, G. (2005). Theoretical Concepts for the Analysis of Non-Farm Rural Employment. *Quarterly Journal of International Agriculture* 44 (1): 19-34.
- Nguyen, H.T. and Walker, E.A. (2007). A first course in fuzzy logic. Boca Raton, Florida, USA: Chapman & Hall, 2<sup>nd</sup> edition.
- Sivanandam, S.N., Sumathi, S., Deepa, S.N. (2007). Introduction to Fuzzy Logic using MATLAB. Heidelberg, D: Springer Verlag.
- Smithson, M. and Verkuilen, J. (2006). Fuzzy Set Theory. Applications in the Social Sciences. Series: Quantitative Applications in the Social Sciences, No. 07-147, Thousand Oaks, Florida, USA: Sage publications.
- Zadeh, L.A. (1965). Fuzzy sets. *Information and Control* 8: 338-353.