European Regional Science Association

ERSA 2003 Congress
University of Jyväskylä – Finland

Fuzzy Spatial Analysis Techniques in a Business GIS Environment

Daniela Wanek

Department of Economic Geography & Geoinformatics
Vienna University of Economics and Business Administration
Rossauer Laende 23/1
A-1090 Vienna, Austria
e-mail: daniela.wanek@wu-wien.ac.at

Abstract

The purpose of the paper is to explore the use of fuzzy logic techniques in spatial analysis. The focus is laid on illustrating whether there is a value added within the context of Business GIS. The issue of geomarketing for illustrative purposes is considered and a case study which works with real world data is discussed.

The objective of the case study is to identify spatial customer potentials for a specific product, a prefabricated family house, using customer data of an Austrian firm. Fuzzy logic is used to generate customer profiles and to model spatial customer potential of the product in question. The use of fuzzy logic in comparison to crisp classification techniques and modelling with crisp operators for solving the problem is illustrated, and more generally how the use of fuzzy logic may be to the advantage of businesses.

1. Introduction

Spatial analysis techniques analyse events which take place in geographical space. Some of the characteristics of these events can be captured with crisp methods, but others are fuzzy in nature. In general, fuzzy information is neglected or simply transformed into crisp information at the price of loosing information.

The present paper explores whether there is a value added of fuzzy logic in spatial analysis techniques in using data from a case study in a Business Geographical Information System (GIS) environment. A Business GIS is a GI System designed to provide useful location-based functions to support management decisions in private and public business enterprises. The application aims to solve a Geomarketing (address-focused marketing) problem. The objective is to find spatial customer potentials for a prefabricated family house (a house which is near completion but not turnkey ready) in north-eastern Austria at the community level for an Austrian producer of family houses.

In order to identify customer profiles, socio-demographic consumer data are classified using crisp and fuzzy cluster analysis. For modelling spatial customer potentials the information about customer profiles and census data are taken. Modelling is done by means of logical and fuzzy logical operators. The results of the two methods are visualised to detect spatial patterns and compared statistically in order to obtain a quantitative comparison.

In section 2, the methodological framework of the case study is briefly described. In section 3, the customer profiles are identified. The methods utilized are described and the results are presented. In section 4, these customer profiles are used for modelling spatial customer potentials. The results are visualised and compared statistically. A conclusion and an outlook sum up the most important findings in the final section.

2. Methodological Approach

"Spatial analysis is a research paradigm that provides a unique set of techniques and methods for analysing events ... that are located in geographical space" (Fischer 2001, p.

14752). Locational attributes are an important source of information for empirical studies in spatial sciences. Spatial data consist of one or few cross-sections of observations for which the absolute location and/or relative positioning is explicitly taken into account on the basis of different measuring units (see Fischer 2001).

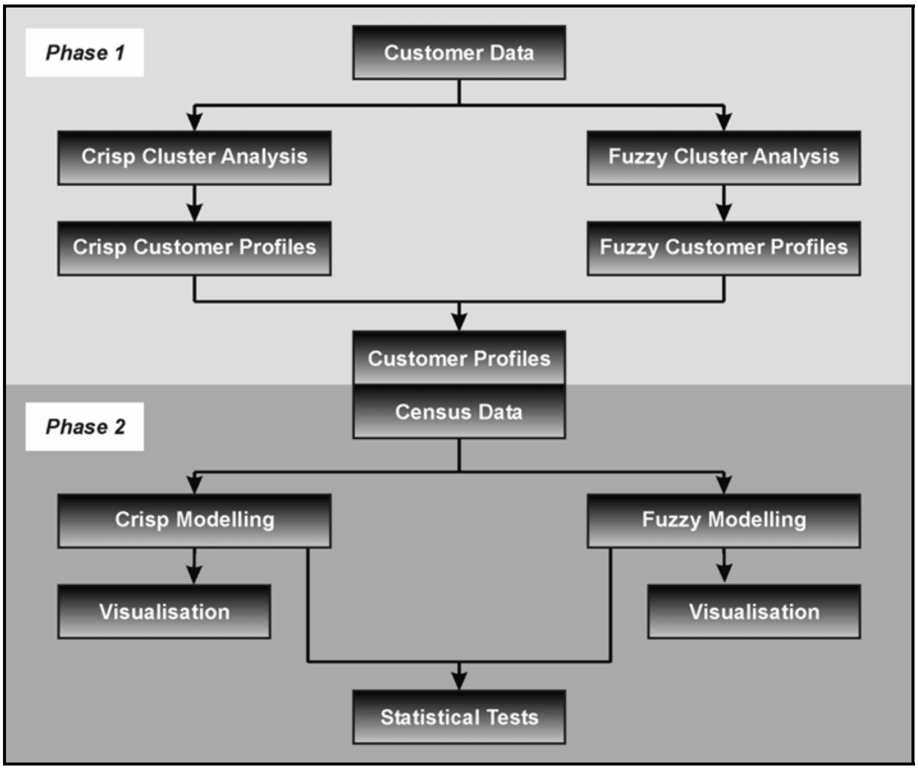
Most data available in a company have a spatial dimension. Decisions in marketing and management often have a spatial dimension, too. Dealing with such data often means dealing with fuzzy information which is difficult to handle with conventional methods. Spatial data analysis techniques are in general crisp and rely on crisp information. Fuzzy logic – more-valued logic – enables to utilise fuzzy information to represent real world situation. Fuzzy sets are defined by functions. Fuzzy logic allows avoiding 'hard' decisions and enables elements to be assigned to more than one set. Elements can be assigned to sets over memberships. That means, for example, one element can be member of set *A* with a membership of 0.2 and member of set *B* with a membership of 0.8. The sum of all memberships to all sets has to be 1.

In the present paper a case study compares the performance of fuzzy methods to crisp methods. The objective of the case study is to solve a Geomarketing problem for an Austrian company which produces and sells prefabricated family houses. The main focus is laid on locating spatial customer potentials for a specific product, the prefabricated family house, in north-eastern Austria at the community level. A prefabricated family house is a house which is near completion but not turnkey ready. The flooring, wall papers or painting, sanitary facilities, etc. are done by the building owner or professional craftsmen, who are not subcontractors of the prefabricated house producer.

A two phase approach is used to solve this problem (see Figure 1). In the first phase, socio-demographic data from existing customers are classified utilising cluster analyses to obtain customer profiles. Cluster analyses aim to classify a data set into groups that are internally cohesive and externally isolated (see Math Soft, Inc. 1998). Crisp and fuzzy clustering methods will be used (see Figure 1).

In the second phase, spatial customer potentials are modelled. The objective is to find communities in which people matched with the customer profiles are over-represented under specified conditions. Modelling takes place by means of using crisp and fuzzy operators. The results for the spatial customer potentials are compared visually and statistically (see Figure 1). The results for spatial customer potentials will be visualised in ArcView and compared with statistical methods. Statistical methods which demand defuzzification of the results and one method which can work with fuzzy results are used.

Figure 1 The Two Phase Approach



3. Identifying Customer Profiles: Theory and Results

Empirical sciences aim to put objects or phenomena of particular scientific interest in order by classification to simplify the field of research. Ordering tries to point out which aspects – measured on the particular cognitive interest – are important and which are unimportant. To focus on the important aspects, all objects of phenomena which are differentiated in a dispensable amount from one another are subsumed in one class with one common name (see Fischer 1982).

Customer profiles are identified by utilizing iterative partitioning cluster analytical techniques. The techniques used refer to the principle of external separation which means that different classes should be "as much apart in the attribute space as possible with respect to the specified attributes" (Fischer 1979, p. 506). The objective is to divide a set of customers characterised by *p* attributes into an a priori given number (*k*) of classes.

An iterative relocation algorithm modifies an initial classification by moving objects from one group to another if this will reduce the optimisation criterion. Iterative partitioning cluster analyses are characterised by two phase techniques. In the first phase (build), an initial partition must be selected or constructed and the number (*k*) of classes must be specified. In the second phase (swap), the quality of the classification is advanced through swap procedures. Restrictions for the iterative relocation method are that it requires specification of the number of clusters in advance and that it also requires an initial classification (see Math Soft, Inc. 1998). A partitioning algorithm which indicates a method that divides the data set into *k* clusters is used. It must be decided how an adequate objective function for the evaluation of the partitions should look like. There has to be a choice of a technique for iteratively re-allocation of customers and the choice of a stopping rule when further re-allocation does not advance the result (see Fischer 1979). Usually you run the algorithm for a range of *k*-values. "For each *k*, the algorithm carries out the clustering and also yields a 'quality index', which allows the user to select the 'best' value of *k* afterwards" (Math Soft, Inc. 1998, p. 510).

In order to obtain customer profiles, non-hierarchical iterative partitioning cluster analyses from the statistical programme S-Plus 2000 are carried out: the crisp cluster analysis Partitioning Around Medoids (PAM) and the fuzzy cluster analysis Fuzzy Analysis (FANNY). The aim is to get the best partition possible of the *n* objects to *k* clusters with regard to a chosen optimisation criterion. This is a classical optimisation problem.

There is used a dissimilarity matrix which means that the value is low if customers are similar and high if customers are dissimilar. The dissimilarity matrix is computed with the programme DAISY, which is included in the S-Plus 2000 package.

The *n* consumers can be characterised by *p* measurements or attributes. That means working with an *n*-by-*p* matrix. The rows of this matrix correspond to the consumers and the columns correspond to the variables (see Kaufman and Rousseeuw 1990). If the individuals (n) are characterised by p variables of different measurement levels, the dissimilarity is calculated as:

$$d(i,j) = \frac{\sum_{f=1}^{p} \delta_{ij}^{(f)} d_{ij}^{(f)}}{\sum_{f=1}^{p} \delta_{ij}^{(f)}} \qquad (1)$$

where the number of variables $f = 1,...,p$. The indicator $\delta_{ij}^{(f)}$ is 1 if the value for attribute f of customer i and the value for attribute f of customer j are no missing values. Otherwise the indicator will be 0. In this case, $d(i,j)$ must either be assigned a conventional value or one of the two customers has to be removed.

For crisp cluster analysis the program PAM is used. Starting from a number of characteristics (p) which characterise the customers, a number of k clusters should be built. Classification takes place with the dissimilarity measure. Each cluster has to contain at least one customer. Each customer has to be part of exactly one cluster. This means that every customer is a member of a cluster with membership 1 or no member with membership 0.

The algorithm utilized in PAM computes k representative customers (medoids). Each customer is "assigned to the cluster corresponding to the nearest medoid" (Math Soft, Inc. 1998, p. 515). This means that customer i "is put into cluster v_i when medoid m_{v_i} is nearer than any other medoid m_w" (Math Soft, Inc. 1998, p. 515):

$$d(i, m_{v_i}) \leq d(i, m_w) \qquad \text{for all } w = 1,\ldots,k \qquad (2)$$

PAM aims to compute k representative customers (medoids) by minimising the sum of the dissimilarities of all customers to their nearest medoid. The objective function C looks as follows:

$$C = \sum_{i=1}^{n} d(i, m_{v_i}) \qquad (3)$$

The algorithm used in PAM consists of two phases. In the first phase (called BUILD), an initial clustering is computed by the successive selection of representative objects until k objects have been found. "The first object is the one for which the sum of the dissimilarities to all other objects is as small as possible. This object is the most centrally located in the set of objects. Subsequently, at each step another object is selected. This object is the one which decreases the objective function as much as possible. ... This process is continued until k objects have been found. In the second phase of the algorithm (called SWAP), it is attempted to improve the set of representative objects and therefore also to improve the clustering yielded by this set" (Kaufman and Rousseeuw 1990, p. 102f). Through iterative display of individual customers from one cluster to another it is attempted to optimise the classification. A singular allocation to one cluster is not definite. A revision is possible if a better classification can be achieved. The swap is carried out until the objective function no longer decreases (see Math Soft, Inc. 1998).

Fuzzy cluster analyses allow the assignment of a customer to more than one cluster. Customers are assigned over memberships to the clusters. These memberships cannot take a negative membership value. The sum of all membership values must be 1. That means, for example, that a customer may be member of cluster one with a membership of 0.3, member of cluster two with a membership of 0.5 and member of cluster three with a membership of 0.2. The outcome of this is that hard clustering solutions are limiting cases of fuzzy ones. The fuzzy cluster analysis called FANNY minimises the following objective function C:

$$C = \sum_{v=1}^{k} \frac{\sum_{i,j=1}^{n} u_{iv}^2 u_{jv}^2 d(i,j)}{2\sum_{j=1}^{n} u_{jv}^2} \tag{4}$$

in which $d(i,j)$ represent the given dissimilarities between objects i and j, and u_{iv} is the unknown membership of object i to cluster v. Using Dunn’s partition coefficient, it can be evaluated how far off a fuzzy solution is from a hard clustering. Dunn’s partition coefficient is defined as the sum of squares of all membership coefficients, divided by the number of objects

$$F_k(U) = \sum_{i=1}^{n} \sum_{v=1}^{k} u_{iv}^2 / n \tag{5}$$

where U is the matrix of all memberships (see Kaufman and Rousseeuw 1990). "It can be seen that for a partition (u_{iv} restricted to 0s and 1s) $F_k(U)$ gets the maximum value of 1, whereas it takes on the minimum value of $1/k$ when all $u_{iv}=1/k$" (Kaufman and

Rousseeuw 1990, p. 186). To get a representative customer for a fuzzy cluster the results have to be defuzzified. The crisp classification which converges best to the fuzzy result is called closest hard classification. This nearest crisp clustering method assigns each object to the cluster in which it has the highest membership.

The partition of the customers can be visualised with a silhouette plot. The silhouette value $s(i)$ which represents the membership of an element to a cluster (see Kaufman and Rousseeuw 1990) is computed for each object i and is represented in the plot as a bar of length $s(i)$. The value for $s(i)$ always lies between -1 and 1. The value $s(i)$ may be interpreted as follows:

- $s(i) \approx 1 \Rightarrow$ object i is well classified.
- $s(i) \approx 0 \Rightarrow$ object i lies between two clusters.
- $s(i) \approx -1 \Rightarrow$ object i is badly classified.

"The silhouette of a cluster is a plot of the $s(i)$, ranked in decreasing order, of all its objects i. The entire silhouette plot shows the silhouettes of all clusters next to each other, so the quality of the clusters can be compared. The overall average silhouette width of the silhouette plot is the average of the $s(i)$ over all objects i in the data set. ... It is possible to run PAM several times, each time for a different k, and to compare the resulting silhouette plots. The user can then select that value of k yielding the highest average silhouette width. If even the highest width is below (say) 0.25, one may conclude that no substantial structure has been found" (Math Soft, Inc. 1998, p. 516).

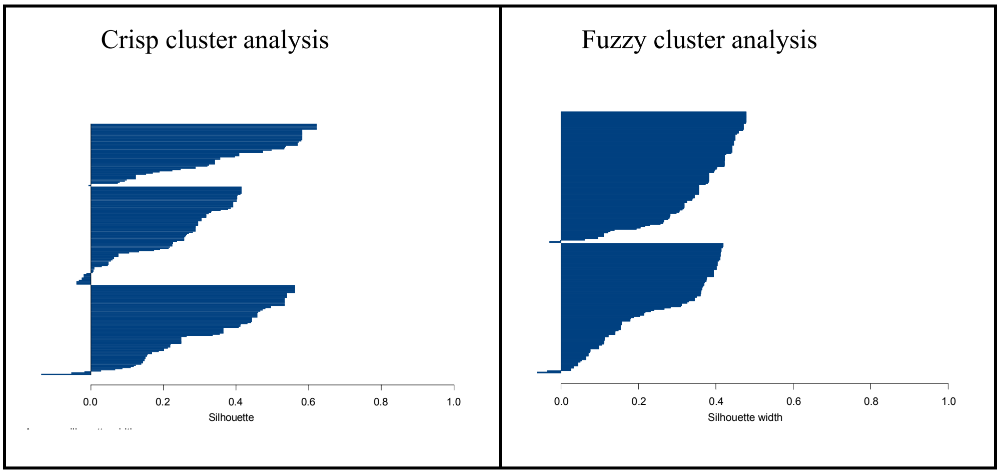
The data base of the case study are 217 customers who bought a prefabricated family house in the period from 1996 to 1998. The data were collected in a mail survey (see Stadelhofer 2000, Staufer-Steinnocher and Stadelhofer 2000). Customer data are classified using iterative partitioning cluster analysis – both fuzzy and crisp. The following socio-demographic characteristics of the customers are used for classification:

- marital status,
- age,
- size of household, and
- highest finished education level.

Using the available data PAM allows 2 to 4 partitions. The results are compared and the best partition is searched considering important statistical indices. Build is one of these indices which gives information about the average dissimilarity of the results in phase one. Swap gives information about the average dissimilarity for the result in the second phase. The value for separation shows the minimal dissimilarity of one customer who is a member of one cluster to a customer who is a member of another cluster. The value for average width tells how well a customer fits into a cluster.

There are indices for each cluster which show the suitability of a cluster. The medoid stands for the customer who lies in the centre of a class and therefore represents the class best. The number of customers who lie in a cluster is shown. The maximum dissimilarity and the average dissimilarity are specified for each cluster. The diameter stands for the maximum dissimilarity between two customers of two clusters. The smaller the value of the diameter the better the clustering.

Figure 2 Silhouette Plot of Crisp and Fuzzy cluster analysis for *k*=3



PAM is computed for initial partitions from *k*=2 to *k*=4. The result of the initial partition with *k*=3 brought best results in consideration of statistical indices. In the following the result for *k*=3 is shown as an example for all computed cluster analysis. There are no isolated clusters. The largest number of customers is member of cluster 2 followed by cluster 3 and cluster 1. Maximum dissimilarity, average dissimilarity and diameter lie in all three clusters in similar intervals. The average width of the silhouette plot is small for cluster 2 followed by cluster 3 and cluster 1 (see Figure 2). Summing up for the whole crisp classification of *k*=3, the average width is 0.3 and the separation is 0.25.

Crisp cluster analysis for *k*=3 is the classification which gives best results and yields in the following three customer profiles (see Table 1 and Figure 2). The first cluster comprises young unmarried individuals with apprenticeship certificate who live in small households; the second cluster comprises married individuals between 35 and 39 with higher education entrance qualification who live in households with 3 persons; and the third cluster comprises married individuals between 35 and 39 with apprenticeship certificate who live in households with 4 persons.

Table 1 Crisp Customer profiles for *k*=3

Customer Profile		Socio-demographical Characteristics			
No.	*Characteristics*	*Marital Status*	*Age*	*Size of House-hold*	*Highest Finished Education*
1	Young unmarried individuals with apprenticeship certificate who live in small households	Un-married	25-29	2 Persons	Apprenticeship
2	Married individuals between 35 and 39 with higher education entrance qualification who live in households with 3 persons	Married	35-39	3 Persons	School-leaving examination certificate (Reifeprüfungs-zeugnis)
3	Married individuals between 35 and 39 with apprenticeship certificate who live in households with 4 persons	Married	35-39	4 Persons	Apprenticeship

Using the available data FANNY allows 2 to 12 partitions. The value for iteration indicates how many iteration steps are necessary to achieve this result. The value for objective tells the result which is achieved for the objective function. The Dunn's partition coefficient gives information about how 'hard' or 'sharp' a result of a fuzzy cluster analysis is. The result for the closest hard clustering shows the hard classification which approximates the fuzzy classification best.

The result for a fuzzy cluster analysis with *k*=3 gives the best classification with respect to the statistical reference numbers. After defuzzification we get two clusters for the closest hard clustering. For all customers, the membership to two of the three clusters is greater than to the third (see Figure 2). Dunn's partition coefficient is 0.33, that means the classification is totally fuzzy. The average width is 0.3.

Fuzzy cluster analysis for *k*=3 is the classification which gives best results. Defuzzifying the fuzzy results yields in two customer profiles (see Table 2). The first cluster comprises married individuals between 35 and 39 with higher education entrance qualification who live in households with 3 persons; and the second cluster comprises married individuals between 30 and 34 with apprenticeship certificate who live in households with 3 persons.

Table 2 Fuzzy Customer profiles for *k*=3

Customer Profile		**Socio-demographical Characteristics**			
No.	*Characteristics*	*Marital Status*	*Age*	*Size of House-hold*	*Highest Finished Education*
2	Married individuals between 35 and 39 with higher education entrance qualification who live in households with 3 persons	Married	35-39	3 Persons	School-leaving examination certificate (Reifeprüfungs-zeugnis)
4	Married individuals between 30 and 34 with apprenticeship certificate who live in households with 3 persons	Married	30-34	3 Persons	Apprenticeship

One of the customer profiles is the same for crisp and for fuzzy version (see profile number 2 from crisp and from fuzzy). Altogether, the crisp and fuzzy analysis result in four customer profiles (see numbers in Table 1 and Table 2). These customer profiles are the source for modelling spatial customer potentials.

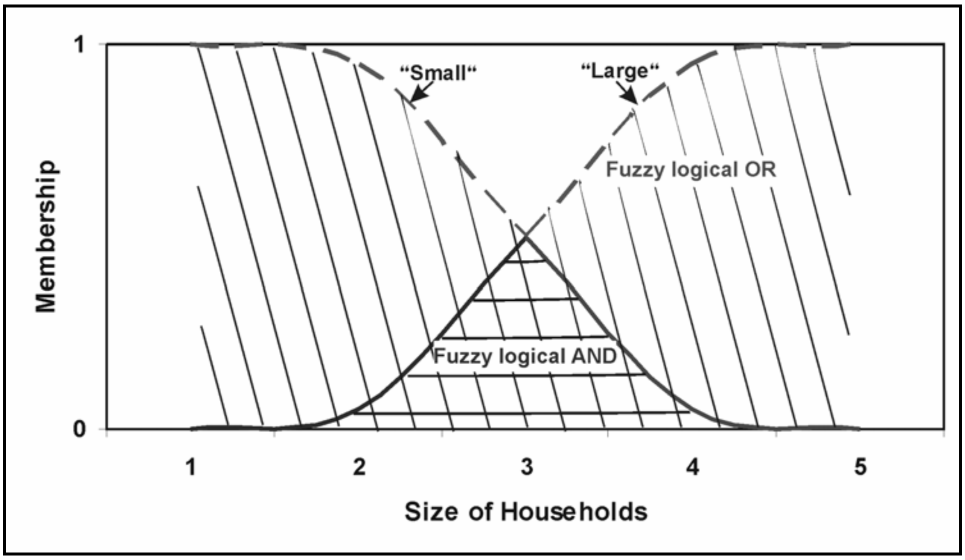
In this example defuzzifying the fuzzy result for *k*=3 yields into two clusters. That means that one cluster is getting lost by defuzzification. In this context classification with crisp cluster analysis achieves a more adequate solution because the information which should be kept by using fuzzy logic is lost again by the necessity of defuzzification.

4. Identifying Spatial Customer Potentials: Theory and Results

The objective of modelling spatial customer potentials is to find spatial units – here communities – where people who match with the customer profiles are over-represented. Companies which have knowledge about the spatial allocation of potential customers have a very important instrument in their hands. One advantage is, for example, to get better possibilities for shaping marketing strategies. Knowledge about the location of potential customers enables for instance to minimise loss through dispersion (see Fischer and Staufer-Steinnocher 2001).

For the crisp variant, the critical value for over-representation is defined as the median of a characteristic over all communities. For the fuzzy variant you need two critical values, one for a membership of 1 which means a customer is a full member of the set and one for a membership of 0 which means a customer is no member of the set. The critical value for a membership of 1 is given if the value of a community is greater than or equal to the value which amounts if you add half of the standard deviation to the median. The critical value for a membership 0 is given if the value of a community is lower than the median minus half of the standard deviation. Between these two values, the membership varies between 0 and 1 according to an s-function.

Figure 3 Fuzzy Logical OR and Fuzzy Logical AND



For modelling spatial customer potentials a lot of different variants are possible. The characteristics of the customer profiles can be combined with different operators, for example with logical AND or with logical OR. The combination of the customer profiles can be varied too, for example with fuzzy logical AND or with fuzzy logical OR (see Figure 3). In the following, one crisp and one fuzzy variant are shown as an example for all computed models.

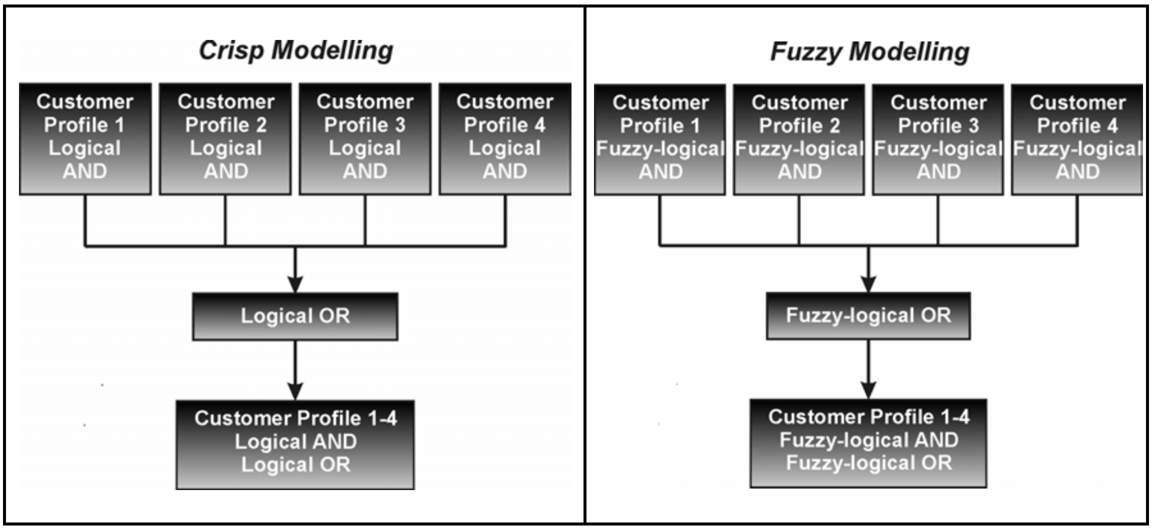
The computation is carried out in ArcView using a non-commercial add-on module, called MapModels, which has been developed at the Vienna University of Technology.

This extension allows fuzzy modelling and contains fuzzy operators (see Riedl and Kalasek 1998, Riedl 1999).

Modelling spatial customer potentials uses the customer profiles and data from the Austrian census of 1991 which includes the information about the socio-demographical characteristics at the community level. To avoid an influence of the size of the communities, percentage data and not absolute data are used.

The following crisp variant is computed: The characteristics of the customer profiles are combined with logical AND and the profiles are combined with logical OR (see Figure 3). That means for example for customer profile 1 that communities with individuals who are between 25 and 29 years old AND communities with individuals who are unmarried AND communities with individuals with apprenticeship certificate AND communities with individuals who live in small households are searched. The same procedure is executed for the other customer profiles. Than the results of the customer profiles are combined with a logical OR.

Figure 4 The Two Approaches for Modelling Spatial Customer Potentials



For the fuzzy variant, the characteristics of the profiles are combined with fuzzy logical AND (minimum operator) and the profiles are combined with fuzzy logical OR (maximum operator) (see Figure 4).

The models which are calculated in ArcView-GIS are compared visually and statistically. The results show similar patterns for the fuzzy and the crisp modelling (see Figure 5 and 6). Both maps show a high spatial customer potential in the catchment areas of bigger cities like Vienna, St. Pölten and Linz, and there is a high spatial customer potential in the north of the Burgenland. The difference is that the variant using fuzzy logic contains more differentiated information. A statement about the degree of the interesting characteristics can be given. Higher colour value in grey shows higher membership to the set of interesting communities. Defuzzification of the fuzzy result via a hard critical value (0.5 and more get membership 1) gives an identical pattern as the crisp result.

Figure 5 Case Study: Crisp Modelling Results

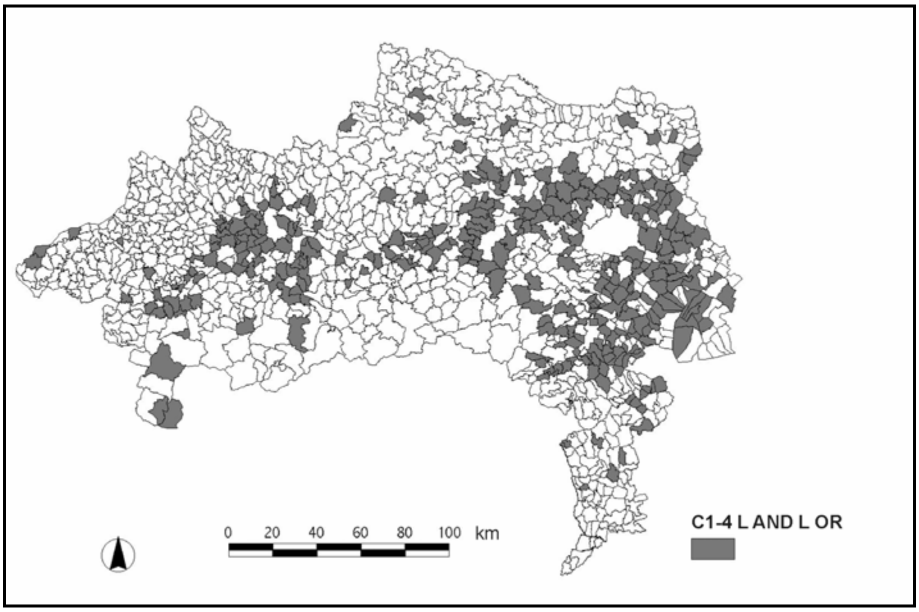
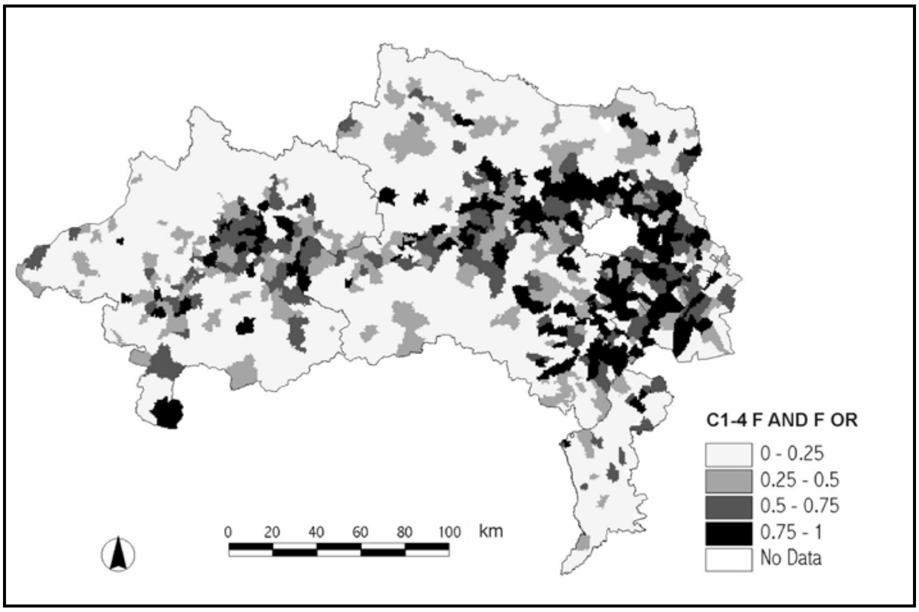


Figure 6 Case Study: Fuzzy Modelling Results



The results are compared with three statistical tests: the Wilcoxon's test, the Kappa Coefficient and the Contingency Coefficient. For comparison with the Wilcoxon's test you do not need to defuzzify. For comparison with the Kappa-Coefficient and with the Contingency Coefficient it is necessary to defuzzify.

The Wilcoxon's signed-rank test compares matched pairs. This test is used to examine differences in the central tendencies of distributions. With the Wilcoxon's test it is possible to test if differences of matched pairs are symmetrically distributed with the median equal to 0. In a first step differences of the measurements for pairs are computed. In a second step absolute differences are ordered by rank digit rows. Pairs with equal measurements are eliminated. The algebraic signs of the differences are assigned to the rank digits. Under the null hypothesis there is no difference between the two data sets. Summing up each with positive and negative rank digits under condition of the null hypothesis it is expected that the sums are added to 0. The alternative hypothesis expects that the sums are different. The z-value is the border value for which the null hypothesis is indifferent. The error probability for which the test decision is indifferent is given over the significance level (see Janssen and Laatz 1999). In this example the z-value is -3.704 and the asymptotic significance is 0.00. That means that the two distributions are different.

The Kappa Coefficient tests concordance between two or more measurements of the same issue. The results of two different modellings should be compared. Kappa is a correspondence measurement which takes into account that by chance a special share of accordance could be expected. A high value for Kappa stands for high concordance whereas a low value stands for little concordance between the two measurements. The null hypothesis says that there is no concordance between the two measurements. That means that Kappa is 0. The value for the approximate T is used as a border value for the t-distribution when the null hypothesis is 0. The error probability from which the test decision is indifferent is given by the significance level (see Janssen and Laatz 1999). In this example the value for Kappa is 1, the approximate T is 34.351 and the approximate significance is 0.00. That means that the comparison of crisp and fuzzy modelling results shows complete concordance.

The Contingency Coefficient gives information about the strength of a coherence of variables. The Contingency Coefficient of Pearson is a measure for tightness of the coherence of two characteristics of four- and more-fold tables. The maximum accessible value for contingency depends on the number of rows and columns of the table. The null hypothesis says that there is no contingency between the two variables. That means that the Contingency Coefficient is 0. The maximum value of the contingency for a fuor-fold table is 0.7071 and appears by completely contingency. The error probability from which the test decision is indifferent is given over the significance level (see Sachs 1999). In this example the value for the contingency is 0.707 and the approximate significance is 0.00. That means that the results of both modelling variants are completely contingent.

Using fuzzy logic operators for modelling spatial customer potentials brings more sophisticated results because the information about membership values gives a more differentiated spatial pattern. The quality of the obtained information is on a higher niveau due to the fact that membership values allow stepwise spatial customer potentials.

5. Conclusions

A lot of the events which are analysed using spatial analysis techniques are not crisp in nature. To analyse such events it is common to commute fuzzy data into crisp data which stands for a loss of information. This paper attempts to explore if one gets a value added by using fuzzy logic to solve such problems. A case study in a Business GIS environment attempts to find a more adequate solution for working with information which is fuzzy in nature.

The case study aims to detect spatial customer potentials for an Austrian producer. To solve this problem customer data are classified using on the one hand crisp cluster analysis on the other hand fuzzy cluster analysis to get customer profiles. The customer profiles and data from the Austrian census are used for modelling spatial customer potentials. Modelling takes place using on the one hand crisp on the other hand fuzzy operators.

In the context of the case study it is demonstrated that classification with crisp cluster analysis achieves a more adequate solution. There is a loss of information using crisp borders but in the following the information could be handled in a better way. A fuzzy classification uses the information in more detail because the restriction of crisp borders can be avoided. The interpretation of the results of a fuzzy cluster analysis is very difficult. To get customer profiles it is necessary to defuzzify the results. By doing this the added information is lost again.

Modelling with fuzzy logic operators yields more sophisticated results. Through the information about memberships we get a more differentiated pattern of the communities. Using this information offers more possibilities for marketing strategies. Fuzzy modelling allows including communities with membership values closely below 1. This means that supplementary information can be obtained that would not be considered with a crisp modelling approach.

A combination of crisp and fuzzy methods seems to be best in this context. By using crisp for classification and a fuzzy modelling approach it is possible to profit from the advantages of both methods.

In this paper modelling spatial customer potentials takes place using basic fuzzy operators. Using more sophisticated fuzzy operators for modelling spatial customer profiles and testing how the results will change would be very interesting. A disadvantage of doing so is that the possibility of a direct benchmark with a crisp analogue will be lost. To exhaust the possibilities fuzzy logic offers completely you have to know a lot about fuzzy logic and you need a high amount of expert knowledge. However the interpretation of the results can be very difficult.

Natural sciences already use the possibilities fuzzy logic offers. This paper attempts to start a discussion to bring fuzzy logic into standard methodologies of social sciences too.

References

Fischer, M. M. (1979): Regional Taxonomy: A Comparison of some Hierarchic and Non-Hierarchic Strategies. *Regional Science and Urban Economics* 10, pp. 503-537.

Fischer, M. M. (1982): Eine Methodologie der Regionaltaxonomie: Probleme und Verfahren der Klassifikation und Regionalisierung in der Geographie und Regionalforschung. *Bremer Beiträge zur Geographie und Raumplanung*, Bd. 3. Bremen: Universität Bremen, Presse- und Informationsamt.

Fischer, M. M. (2001): Spatial Analysis in Geography, in Smelser, N. J. and Baltes, P. B. (eds.): *International Encyclopedia of the Social and Behavioral Sciences*, Vol. 22, pp. 14752-14758. Oxford: Elsevier.

Fischer, M. M. and Staufer-Steinnocher, P. (2001): Business GIS und Geomarketing: GIS für Unternehmen, in Institut für Geographie und Regionalforschung der Universität Wien (ed.): *Geographischer Jahresbericht aus Österreich*, Bd. 58, pp. 9-24. Wien: Eigenverlag.

Höppner, F. et. al. (1999): *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis, and Image Recognition*. Princeton: Wiley.

Jaanineh, J. and Maijohann, M. (1996): *Fuzzy-Logik und Fuzzy-Control*. Würzburg: Vogel.

Janssen, J. and Laatz, W. (1999): *Statistische Datenanalyse mit SPSS für Windows: eine anwendungsorientierte Einführung in das Basissystem Version 8 und das Modul Exakte Tests*. Berlin, Wien and New York: Springer.

Kaufman, L. and Rousseeuw, J. (1990): *Finding Groups in Data: An Introduction to Cluster Analysis*. Chichester: Wiley.

Math Soft, Inc. (1998): *S-Plus 4 Guide to Statistics. Data Analysis Products Division.* Seattle: Math Soft.

Riedl, L. and Kalasek, R. (1998): MapModels - Programmieren mit Datenflußgraphen, in Strobl, J. and Dollinger, F. (eds.): *Angewandte Geographische Informations-verarbeitung. Beiträge zum Agit-Symposium Salzburg 1998*, pp. 279-288. Heidelberg: Wichmann.

Riedl, L. (1999): Leop's MapModels. User Manual. Wien: Institut für Stadt- und Regionalforschung, Technische Universität Wien.

Sachs, L. (1999): *Angewandte Statistik: Anwendung statistischer Methoden*. Berlin, Wien and New York: Springer.

Spies, M. (1993): *Unsicheres Wissen: Wahrscheinlichkeit, Fuzzy-Logik, neuronale Netze und menschliches Denken*. Heidelberg, Berlin and Oxford: Spektrum Akademischer Verlag.

Stadelhofer, T. (2000): *Geomarketing in der Fertighausbranche.* Wien: Diplomarbeit, Abteilung für Wirtschaftsgeographie und Geoinformatik, Wirtschaftsuniversität Wien.

Staufer-Steinnocher, P. (2000): Business-GIS und Geomarketing. Die Geographie der Märkte differenziert den Unternehmenserfolg. *Proceedings GIS/SIT 2000*, Fribourg 12.-14. April 2000, pp. 14.1-14.10. Basel: GISWISS.

Staufer-Steinnocher, P. and Stadelhofer, T. (2000): Marktforschung und Marktbearbeitung im Fertighaussektor, in Fally, M. and Strobl, J. (eds.): *Business Geographics: GIS in der Wirtschaft*, pp. 130-137. Heidelberg: Wichmann.

Staufer-Steinnocher, P. and Wanek, D. (2002): Kundenprofilanalysen und Räumliche Kundenpotentialanalysen – Crisp- und Fuzzy-Verfahren für Geomarketing. in Strobl, J., Blaschke, T. and Griesebner, G. (eds.): *Angewandte Geographische Informationsverarbeitung XIV*, pp. 579-588. Heidelberg: Wichmann.

Zadeh, L. (1968): Probability Measures of Fuzzy-Events. *Journal of Mathematics, Animation and Application* 23, pp. 421-427.