Mathematical Geology, Vol. 39, No. 1, January 2007 (© 2007)
DOI: 10.1007/s11004-006-9066-4

Fuzzy Modeling for Reserve Estimation Based on Spatial Variability[1]

Bulent Tutmez,[2] A. Erhan Tercan,[3] and Uzay Kaymak[4]

This article addresses a new reserve estimation method which uses fuzzy modeling algorithms and estimates the reserve parameters based on spatial variability. The proposed fuzzy modeling approach has three stages: (1) Structure identification and preliminary clustering, (2) Variogram analysis, and (3) Clustering based rule system. A new clustering index approach and a new spatial measure function (point semimadogram) are proposed in the paper. The developed methodology uses spatial variability in each step and takes the fuzzy rules from input-output data. The model has been tested using both simulated and real data sets. The performance evaluation indicates that the new methodology can be applied in reserve estimation and similar modeling problems.

KEY WORDS: reserve estimation, spatial variability, fuzzy modeling, variogram.

INTRODUCTION

Ore reserve estimation, which includes the estimation of characteristics such as ore grade and ore thickness, has crucial importance for evaluation of mineral deposits. Estimated reserve is used in all phases of a mining project, including feasibility, mine planning and production scheduling. Traditional reserve estimation methods can be roughly classified into two main groups: geometric and geostatistical methods (Henley, 2000). Geometrical methods depend on geometrical relationships between sample points while geostatistical methods (Goovaerts, 1997; Henley, 2001) are based on random functions and consider spatial relationship of the sample data. Generally, geostatistical methods are more effective. Bardossy and Fodor (2004) have discussed the advantages and disadvantages of geostatistical methods for reserve estimation and they stressed that geostatistical

[1]Received 17 January 2006; accepted 19 June 2006; Published online: 21 February 2007.
[2]Department of Mining Engineering, Inonu University, Malatya 44280, Turkey; e-mail: btutmez@inonu.edu.tr.
[3]Department of Mining Engineering, Hacettepe University, Ankara 06532, Turkey; e-mail: erhan@hacettepe.edu.tr.
[4]Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands; e-mail: u.kaymak@ieee.org.

methods have some limitations. Geostatistical calculation needs suitable computer programs, and a considerable mathematical background. Further shortcomings of geostatistics were pointed out in detail by Diehl (1997).

Integrating geostatistical concept with fuzzy set theory (Bardossy, Bogardi, and Kelly, 1990) is a novel direction, and the application of fuzzy modeling in reserve estimation is very limited. In the literature, Pham (1997) estimated the ore grades using a linguistic fuzzy model. Galatakis, Theodoridis, and Kouridou (2002) carried out a study for lignite quality estimation using a neural-fuzzy system. The main handicap of these works is that spatial variability of data values can not be used in the algorithms. Recently, Luo and Dimitrakopoulos (2003), Bardossy, Szabo, and Varga (2003), and Bardossy and Fodor (2005) have applied the fuzzy set theory in resource estimation and mathematically evaluated the spatial continuity of ore bodies by fuzzy sets. Similarly, Tutmez (2005) has recently carried out a study that tries to combine fuzzy algorithms and spatial variability in reserve estimation.

This paper considers reserve estimation using spatial variability and fuzzy modeling. In general, existing fuzzy modeling approaches focus on model accuracy. Spatial variability of data is not taken into consideration by them. However, spatial positions of data directly connected with data values (grades) is very important for reserve estimation. The aim of this paper is to propose a systematic data driven fuzzy method based on spatial variability for reserve estimation. The developed model uses fuzzy clustering and Takagi-Sugeno model structure. Spatial relationship is used in each stage of the algorithm.

The paper is organized as follows. Section 2 states the theoretical background of spatial correlation and zone of influence for reserve estimation. Section 3 describes fuzzy model identification and deals with the steps of the proposed fuzzy modeling algorithm. Section 4 gives numerical examples for both simulated and real data sets. Section 5 presents the conclusion.

SPATIAL CORRELATION

Reminder on the Zone of Influence

Spatial correlation can be modeled by the variogram function. In practice, the variogram function is estimated as follows:

$$2\gamma(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \tag{1}$$

where $z(x_i)$ and $z(x_i + h)$ are values of the variable at locations x_i and $x_i + h$ and $N(h)$ is the number of pairs used in calculating the variogram at distance h.

The variogram characterizes the structures of the spatial distribution of the variables considered. Beginning at the origin, $\gamma(0) = 0$, the variogram increases in general with the distance. The manner in which this variogram increases at small distances characterizes the degree of spatial continuity of the variable under study. The variogram may become stable beyond some distance $h = a$ called the range. Beyond this distance a, the mean square deviation between two quantities $z(x)$ and $z(x + h)$ no longer depends on distance h between them and the two quantities are not correlated. The range gives an exact sense to the conventional concept of the zone of influence of a sample (Journel and Huijbregts, 1978).

FUZZY MODEL FOR RESERVE ESTIMATION

In this section, we propose a fuzzy model for reserve estimation. The construction of the proposed fuzzy model proceeds in three main stages: (1) structure identification and preliminary clustering (2) variogram analysis, and (3) clustering based rule system generation.

Structure identification and preliminary clustering comprises of two steps. In the first step, input and output variables of the model are determined. Following the first step, the knowledge representation is developed from input-output dataset by using fuzzy clustering. The main objective of this stage is determining the number of clusters.

Variogram analysis of a data point consists of constructing a variogram model which characterizes, in an operational way, the main features of the spatial variability. Two main steps can be distinguished in this: the construction of an experimental variogram and defining the range (search radius) via the modeled variogram.

Clustering based rule system generation comprises of an application of possibilistic clustering to generate an initial rule base, followed by the identification of optimal consequent parameters in a Takagi-Sugeno fuzzy system. In this way, output of the spatial analysis is used in the generation of the fuzzy rule-based system directly.

Parameter Selection and Preliminary Clustering

In this stage, firstly, the input and output variables of the fuzzy model are chosen. From the data set, a regression matrix X and an output vector y are constructed.

$$\begin{aligned} X &= [x_1, \ldots, x_N]^T, \\ y &= [y_1, \ldots, y_N]^T. \end{aligned} \tag{2}$$

Following the parameter selection, data set is partitioned in the Cartesian product space $X \times y$ by using fuzzy c-means (FCM) clustering (Bezdek, Ehrlich, and Full, 1984) which is the most popular fuzzy clustering algorithm. In the clustering algorithm, both data coordinates and its measured value (ore grade) should be represented and clustering must be carried out in the three dimensional space. This point is very important because the variables used in reserve estimation are characterized by their spatial distribution (Journel and Huijbregts, 1978). The fuzzy c-means (FCM) clustering algorithm partitions 3-dimensional vectors into c fuzzy clusters. The algorithm minimizes an objective function based on distance measures between cluster centers and data points.

The fuzzy partition matrix satisfies probabilistic property as

$$\sum_{i=1}^{c} \mu_{ik} = 1, \qquad \forall k = 1, \ldots, N. \tag{3}$$

For partitioning a collection of N data points into c classes, we define an objective function J_m for a fuzzy c-partition,

$$J_m(U, \mathbf{v}) = \sum_{k=1}^{N} \sum_{i=1}^{c} (\mu_{ik})^m (d_{ik})^2 \tag{4}$$

where $m \in [1, \infty]$ is a weighting exponent, μ_{ik} is the membership of the kth data point in the ith class. The term, d_{ik} is a Euclidean distance measure (in 3-dimensional feature space, R^3) between the kth sample data x_k and ith cluster center $\mathbf{v}_i$, given by

$$d_{ik} = d(x_k - \mathbf{v}_i) = \|x_k - \mathbf{v}_i\| = \left[\sum_{j=1}^{3} (x_{kj} - v_{ij})^2 \right]^{1/2}. \tag{5}$$

Cluster means (prototypes) and elements of membership matrix are computed as follows.

$$c_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m x_k}{\sum_{k=1}^{N} \mu_{ik}^m}, \tag{6}$$

and

$$\mu_{ik} = \frac{1}{\sum_{p=1}^{c} \left(\frac{d_{ik}}{d_{pk}} \right)^{2/(m-1)}}. \tag{7}$$

Determining the Number of Clusters

Determining the optimal number of clusters has vital importance in pattern recognition. In the literature, many indices have been proposed by different researchers (Gath and Geva, 1989; Fukuyama and Sugeno, 1989; Kwon, 1998). One of the differentiated cluster validity index was proposed by Xie and Beni (1991). According to Pal and Bezdek (1995), this index provides the best results especially for a small number of clusters, but if the number of clusters becomes very large, the index decreases monotonically. In recent years, other effective methods were proposed based on different theoretical foundations as in (Kaymak and Babuska, 1995).

Because data variability has a special importance in reserve estimation, in the present study a cluster validity approach taking into consideration this parameter (Tutmez, 2005) is used. It is based on reproducing the data variability of the samples in the value of cluster centers with a minimum number of clusters as follows:

$$\text{Minimize } n_c \text{ under,} \qquad \text{Std}[g(x)] \approx \text{Std}[g(c)]. \tag{8}$$

where n_c is the optimal number of clusters, and Std is the standard deviation. In this approach, the number of clusters is plotted against the corresponding standard deviations of the cluster center values and then the number of clusters satisfying constraint (8) is retained as optimum.

Semivariogram Analysis

The purpose of this step is measuring the spatial correlation (Section 2) between cluster centers and data values. Semivariogram analysis in fuzzy model consists of variogram modeling in particular of defining the range parameter of the variogram. Traditionally, spatial variability is calculated by using the difference between sample pairs. In this system all data points are involved in calculation. However, in the present approach we have fixed cluster centers and computation (estimation) is carried out using the differences between data points and the cluster centers. A point semivariogram is used for characterizing the spatial variability in this way.

Point Semivariogram

Point semivariogram (PSV) function was proposed by Şen (1989) and is used in different studies (Şen, 1998; Şen and Oztopal, 2001). It is a modified

version of the classical semivariogram function. The classical semivariogram is not suitable for describing the local variability. PSV was suggested as an alternative measure for evaluating the local behavior of data. It can be used in determining the spatial behavior of any variable around a particular cluster. In other words, it presents the regional effect of all of the other clusters within the study area on the cluster concerned. Şen (1989) presented the basis of the methodology to obtain experimental PSVs for each measurement point, which leads to estimation of the radius of influence around each site (cluster). Its mathematical expression is given as,

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z_m - Z_{(m+h_i)})^2 \tag{9}$$

where N is the number of data values, $N(h)$ is the number of pairs, Z_m is the reference grade value, and $Z_{(m+h)}$ is the grade value at a distance h.

Point Semimadogram

Madograms are particularly useful for establishing range parameter (Deutsch and Journel, 1998). The point semimadogram (PSM), has been proposed by Tutmez (2005). This function is similar to the PSV; instead of squaring the difference between Z_m and Z_{m+h}, the absolute difference is taken. If the experimental variogram includes the outlier values, the PSM is more convenient than the PSV due to the advantages of absolute difference measure (Tutmez, 2005),

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} |Z_m - Z_{(m+h_i)}|. \tag{10}$$

Clustering Based Rule System Generation

This stage consists of the possibilistic clustering for rule base initialization before optimizing rule consequents. First, data clustering is carried out using possibilistic clustering. Second, optimal model parameters are determined by using the membership matrix and cluster information.

Possibilistic Fuzzy Clustering

Possibilistic fuzzy c-means clustering algorithm (PCM) has been proposed by Krishnapuram and Keller (1993). Krishnapuram and Keller (1996) stated that the

possibilistic approach simply means that the membership value of a point cluster represents the typicality of the point in the cluster, or the possibility of the point belonging to the class. They presented a fundamental difference between these two algorithms that FCM is primarily a partitioning algorithm, whereas PCM is a mode-seeking algorithm, i.e., each component generated by the PCM corresponds to a dense region in the data set. In the PCM algorithm, each cluster is independent of the other clusters (Krishnapuram and Keller, 1996). The objective function of this algorithm can be expressed as follows:

$$J_m(U, \mathbf{v}) = \sum_{k=1}^{N} \sum_{i=1}^{c} (\mu_{ik})^m (d_{ik})^2 + \sum_{i=1}^{c} \eta_i \sum_{k=1}^{N} (1 - \mu_{ik})^m, \tag{11}$$

where $\eta_i \succ 0$. The first term corresponds to minimization of the weighted distances, while the second term suppresses the trivial solution and tries to increase the volume of the clusters. The equation for updating the membership degrees that is derived from the objective function is

$$u_{ik} = \frac{1}{1 + \left(\frac{d_{ik}^2}{\eta_i}\right)^{\frac{1}{m-1}}}, \eta_i \succ 0 \tag{12}$$

Determining the η and m using Spatial Variability

It can be shown from the Equations (11–12) that the parameter η_i is related to the scaling. It determines the distance at which the membership degree equals 0.5. Essentially, η_i determines the "zone of influence" of a point. A point x_k will have little influence on estimates of the prototype parameters of a cluster if d_{ik}^2 is large when compared with η_i (Krishnapuram and Keller, 1996). The parameter η_i is chosen for each cluster separately and can be calculated by using the fuzzy intra cluster distance as

$$\eta_i = \frac{K}{N_i} \sum_{k=1}^{n} (\mu_{ik})^m (d_{ik})^2 \tag{13}$$

where $N_i = \sum_{k=1}^{n} (\mu_{ik})^m$. Usually K is chosen equal to 1 (Tim and others, 2004).

We prefer the PCM for clustering due to the functional role of η_i in the algorithm. The "zone of influence" concept is directly related to the range of a variogram and actually it should be determined by means of variogram analysis in our application. Starting from this, the point semivariogram/madogram function

is suggested for defining the η_i (Tutmez, 2005). The proposed approach aims to determine the (η_i) from the semivariogram range.

There is a connection between η_i and weighting exponent m. Note that the spatial correlation is accepted to be 0 for distances beyond the variogram range. However, according to formula (12), membership value of this function does not become 0. For this reason, we try to find an approximate solution. First we consider a termination parameter ε that the membership function (12) should attain at some distance d_{ik}. Then, we can rewrite formula (12) as

$$\frac{1}{1+\left(\frac{d_{ik}^2}{\eta_i}\right)^{\frac{1}{m-1}}} = \varepsilon, \tag{14}$$

$$1 = \varepsilon + \varepsilon\left(\frac{d_{ik}^2}{\eta_i}\right)^{\frac{1}{m-1}}, \tag{15}$$

$$\left(\frac{d_{ik}^2}{\eta_i}\right)^{\frac{1}{m-1}} = \frac{1-\varepsilon}{\varepsilon}, \tag{16}$$

Taking the logarithm of each side of the Equation (16), we obtain

$$\frac{1}{m-1}\ln\frac{d_{ik}^2}{\eta_i} = \ln\left(\frac{1-\varepsilon}{\varepsilon}\right), \tag{17}$$

$$m-1 = \frac{\ln\left(\frac{d_{ik}^2}{\eta_i}\right)}{\ln\left(\frac{1-\varepsilon}{\varepsilon}\right)}. \tag{18}$$

Equation (18) describes the m depending on d_{ik}^2/η and ε. Krishnapuram and Keller (1996) showed that the membership function generated by the PCM is a function of m as shown in Figure 1.

In the PCM, as the value of m increases, the clusters become less localized and they extend much beyond the normalized squared distance of 1. Note that the membership is 0.5 at a distance of $\sqrt{n_i}$. We now impose that the membership should be sufficiently small at a distance $2\sqrt{\eta_i}$ (twice the effective radius). Because of this, the membership function should almost completely decay to zero at $d_{ik}^2/\eta_i = 4$. We can now determine range distances (a_i) for each cluster from semivariogram models, and set this equal to $2\sqrt{\eta_i}$. These values can be transferred into the PCM as follows

$$a_i^2 = (2\eta_i)^2 = 4\eta_i \tag{19}$$

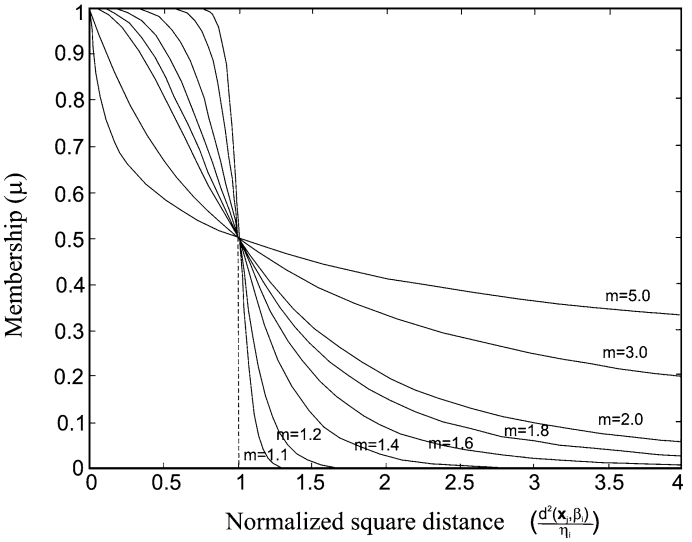

Figure 1. Relationship between distance and memberships for different values of m in PCM (After Krishnapuram and Keller, 1996).

when $d_{ik}^2 = a_i^2 = 4\eta$, Eq. (18) can be rewritten as

$$m_i = 1 + \frac{\ln(4)}{\ln\left(\frac{1-\varepsilon}{\varepsilon}\right)}. \tag{20}$$

If we set $\varepsilon = 0.1$, we find $m = 1.63$.

Alternatively, using a selected m value, a suitable η distance can be calculated. For this, the following is obtained from (16):

$$\eta_i = d_{ik}^2 \left(\frac{\varepsilon}{1-\varepsilon}\right)^{m-1}. \tag{21}$$

For basic solution, $m = 2$, $d_{ik}^2 = a_i^2$ and $\varepsilon = 0.1$ can be selected. Hence, one obtains $\eta_i = a_i^2/9$. Note that the value of η decreases, as the value of m increases.

Rule Based System

In fuzzy modeling, the fuzzy rule-based structure should represent the complex-uncertain system for effective approximation. In the present approach, Takagi-Sugeno (TS) system based on spatial variability is proposed. In the TS

fuzzy model (Takagi and Sugeno, 1985), the rule consequents are crisp functions of the model inputs:

$$R_i: \text{ If } x \text{ is } A_i \text{ then } y_i = f(x),\ i = 1, 2, \ldots, K. \tag{22}$$

The selected TS model depends on the output structure. In this study, 1st order TS fuzzy model is used. This selection is based on the smoothing problem of the general fuzzy models. Smooth estimations, in which the variability of the model output is less than the variability in the data, are undesirable for many systems such as geology based decisions and environmental pollution estimates. For example, mine project planning uses this variability as input. A 1st order TS fuzzy model can capture the variability in the data better than a zero-order TS fuzzy model. The rules in a 1st order TS fuzzy model look

$$R_i: \text{If } x_1 \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2} \text{ then } g_i = a_{i1}x_1 + a_{i2}x_2 + b_i, \quad i = 1, 2, \ldots, K \tag{23}$$

where R_i is the ith rule, g_i is the the rule output and a_{i1}, a_{i2} and b_i are estimated parameters for each rule i. In the rule base, K denotes the number of rules. The aggregated output of the model is calculated by taking the average of the rule contributions (Setnes, Babuska, and Verbruggen, 1998).

Representing the spatial variability in a rule mechanism using linguistic expressions is a necessary step for this approach. In the geostatistical approach, which is based on the spatial variability, functional relationship relies on the distances between data values. As a general hypothesis, if the distance between data values is small, grades (measured data values) are close each other. On the other hand, great distance implies dissimilar grade values. In the rule base, we can represent this relationship by the linguistic term "close."

Membership Function Assignment

When spatial coordinates of sample data are close to the spatial coordinates of the cluster center, it is expected that grade values for both sample and cluster center are also close. For modeling the fuzzy term "close," Yager and Filev (1994) suggested a Gaussian-type function

$$\mu_X = \exp\left[-(\ln 2)\left(\frac{x - c}{\sigma}\right)\right] \tag{24}$$

where σ is a constant which determines the spread of the function, x is sample data value and c is the cluster center (Fig. 2). Gaussian membership functions are more common in fuzzy systems literature, as they provide both simplicity and flexibility.

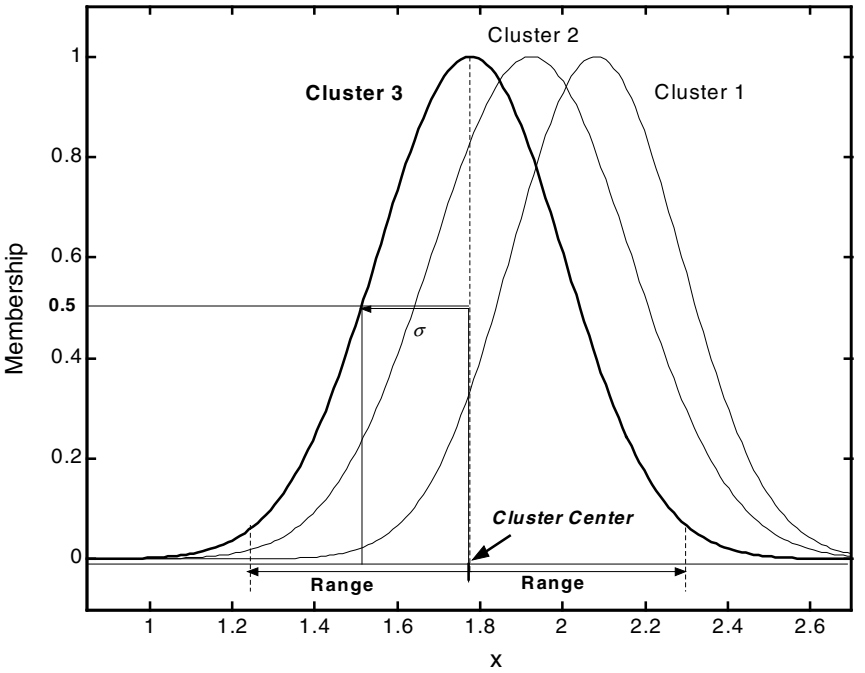

Figure 2. Gaussian memberships and model distance parameters.

Also, fuzzy rule bases typically use membership functions of a single variable, and Gaussian membership functions are suitable in this sense, since multi-dimensional Gaussian functions can be decomposed into single-dimensional Gaussians naturally. Hence, in our rule based system, we apply Gaussian membership functions after the possibilistic clustering step.

Parameter Optimization

A Takagi-Sugeno fuzzy model consists of a set of fuzzy rules, each of which can be described as a local linear input-output relationship (Babuska, 1998). The consequent parameters for each rule can be obtained from a weighted least square estimation.

Let X_e denote a matrix with its elements consisting of input values and a column vector with ones: $X_e = [X; \mathbf{1}]$ and Γ_i be a diagonal matrix in $\Re^{N*N}$ having the normalized weighted memberships $\gamma_i(x_k) = \beta_i(x_k)/\sum_{j=1}^{K} \beta_j(x_k)$ as its kth diagonal element, where the degree of fulfillment of the ith rule β_i can be derived from the fuzzy partition matrix U using point-wise projection (Höppner

and others, 1999). Let X' denote a matrix composed of matrices Γ_i and X_e (Setnes, Babuska, and Verbruggen, 1998).

$$X' = [(\Gamma_1 X_e); (\Gamma_2 X_e); \cdots; (\Gamma_K X_e)] \tag{25}$$

The unknown parameters a_i and b_i in the fuzzy model are lumped into a single parameter vector $\boldsymbol{\theta} \in \Re^{K^*(n+1)}$;

$$\boldsymbol{\theta} = \left[a_1^T, b_1, a_2^T, b_2, \ldots, a_K^T, b_K\right]^T \tag{26}$$

The resulting weighted least squares problem has the solution:

$$\left[a_i^T, b_i\right]^T = \left[X_e^T \Gamma_i X_e\right]^{-1} X_e^T \Gamma_i y \tag{27}$$

NUMERICAL EXAMPLES

The proposed modeling approach is illustrated using two data sets. In the first case study, the fuzzy modeling approach is used for estimating simulated iron-ore grades. In this application performance comparison has been also carried out with a past study. In the second case study, a real data set is used and this time, the proposed modeling approach is considered for the purpose of lignite reserve estimation.

Example 1: A Comparative Study with Simulated Data

Pham (1997) introduced the grade estimation using fuzzy set algorithms and used a simulated data set (Table 1) which was taken from Clark (1979). In this example, the same data set is used and the performance of proposed fuzzy model is compared to the results of (Pham, 1997). The data set used in the example is shown in Figure 3. In clustering literature, it is often suggested that the data should be appropriately normalized before clustering (Jain and Dubes, 1988). Samples taken from iron-ore deposit were standardized by using a linear transformation between maximum and minimum grade values [24.4, 41.5].

Determining the Number of Clusters

Considering three features (Easting, Northing, Iron Data Value), clustering is carried out using FCM. Optimal number of clusters is defined using the validity approach presented in Section 3.1. Standard deviations of the cluster grades are plotted against the number of clusters (Fig. 4) and cluster number 5 is obtained as the optimum (Fig. 3).

Table 1. Initial Data Set from Iron-Ore Deposit (Clark, 1979)

Sample No	Easting	Northing	% Cu	Sample No	Easting	Northing	% Cu
1	10	40	35.5	12	15	135	28.6
2	55	145	29.4	13	125	20	41.5
3	175	50	36.8	14	260	115	33.2
4	235	15	33.7	15	365	60	34.3
5	285	110	35.3	16	345	115	31
6	20	105	32.5	17	25	155	29.6
7	50	40	30.6	18	155	15	40.4
8	145	125	30.1	19	220	90	28.5
9	205	0	40.1	20	265	65	24.4
10	390	65	31.6	21	325	105	39.5
11	310	150	34.8	12	15	135	28.6

Semivariogram Analysis

Figure 5 represents the experimental and model madograms which show the differences between spatial data points and the cluster centers. Calculating the experimental madograms, four lag distances were used for each cluster. Zone of

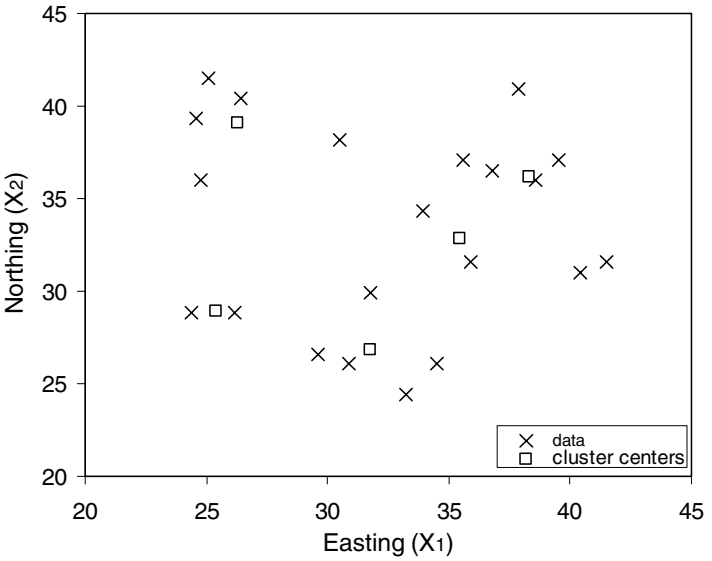

Figure 3. Map of sampling points of a simulated iron ore deposit (After Clark, 1979).

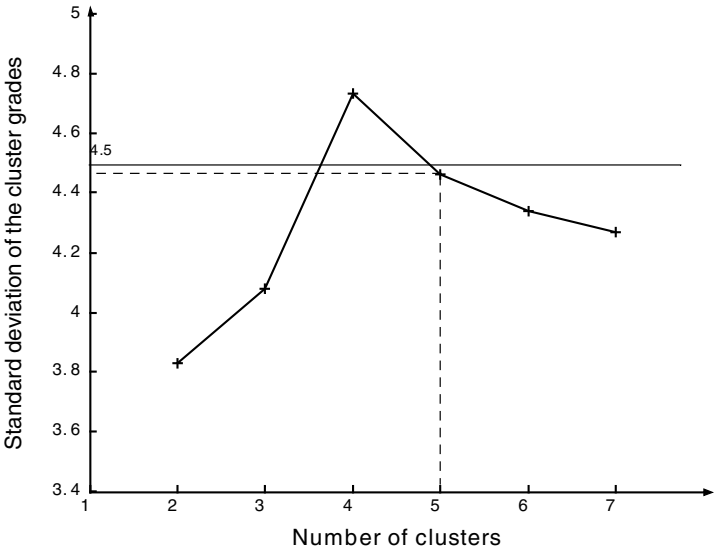

Figure 4. Clustering validity for iron-ore data.

influences of clusters were determined from semimadogram range distances. The range values can be shown as a vector: [5 7 4 10 4].

Clustering and Parameter Optimization

Parameter optimization step consists of the PCM clustering and local least square estimation. In the first stage, Gaussian partition matrix is calculated and then its components are used in the least square estimation problem as the weight factor. Possibilistic clustering algorithm used the following input.

Number of cluster: 5

η parameter vector: $\begin{bmatrix} 2.5 \\ 3.5 \\ 2 \\ 5 \\ 2 \end{bmatrix}_{5 \text{ by } 2}$

Weighting exponent: $m = 1.6$

According to clustering based least square optimization, model parameters are determined. Clusters centers and model parameters are listed in Table 2.

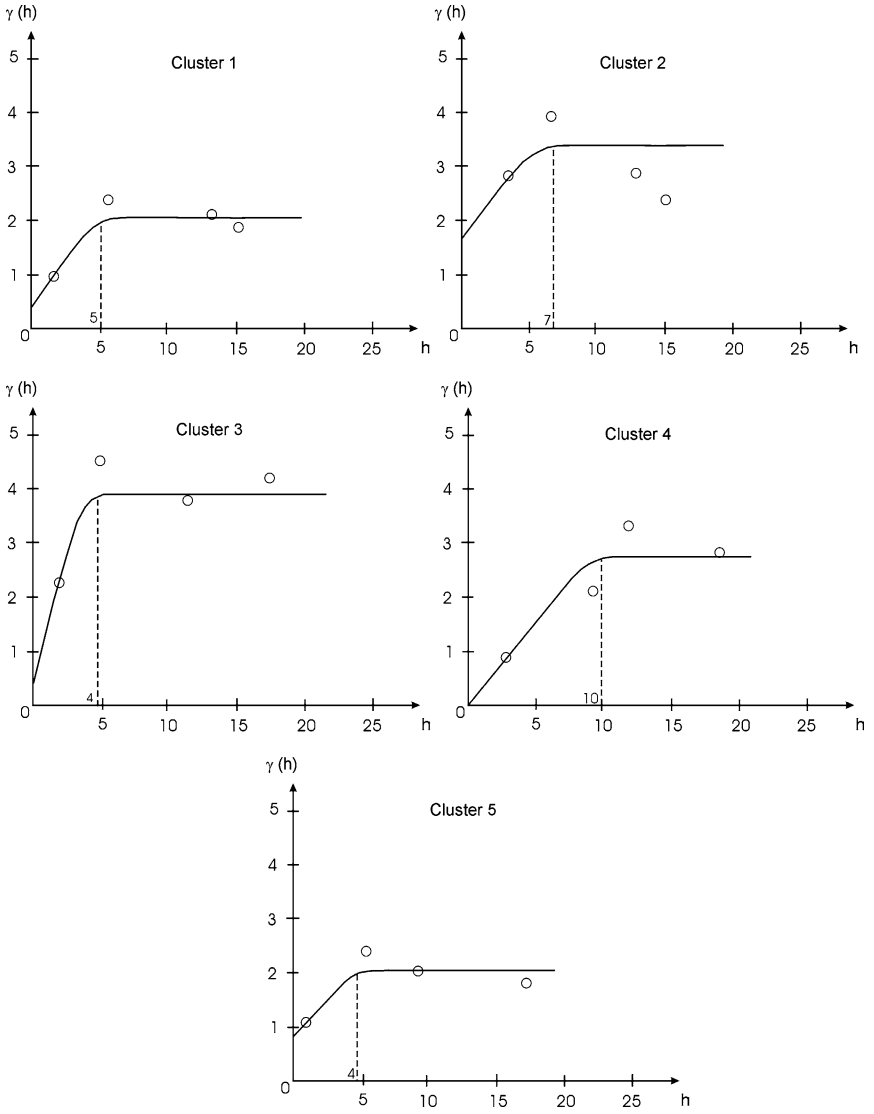

Figure 5. Experimental and model point madograms.

Fuzzy Rule Based System

Gaussian type membership functions are used in the rules. Functions were modified in accordance with the cluster centers and the zone of influence. For this

Table 2. Cluster Centers and Optimized Model Parameters

Clusters	Coordinate x_1	Coordinate x_2	a_1	a_2	b
Cluster 1	31.80	26.80	−1.883	−0.056	102.024
Cluster 2	26.27	39.10	−0.361	−0.505	58.864
Cluster 3	25.40	28.90	−2.950	30.757	−778.373
Cluster 4	38.30	36.20	−0.063	−0.089	32.655
Cluster 5	35.47	32.80	9.255	7.981	−551.100

purpose, the function centers were selected same as the cluster center values. Input membership functions are depicted in Figure 6. The rules are outlined below:

Rule 1: **If** x_1 coordinate is close to coordinate of C_{1X_1} and x_2 coordinate is close to coordinate of C_{1X_2}, **then** Grade $= -1.883x_1 - 0.056x_2 + 102.024$
Rule 2: **If** x_1 coordinate is close to coordinate of C_{2X_1} and x_2 coordinate is close to coordinate of C_{2X_2}, **then** Grade $= -0.361x_1 - 0.505x_2 + 58.864$
Rule 3: **If** x_1 coordinate is close to coordinate of C_{3X_1} and x_2 coordinate is close to coordinate of C_{3X_2}, **then** Grade $= -2.950x_1 + 30.757x_2 - 778.373$
Rule 4: **If** x_1 coordinate is close to coordinate of C_{4X_1} and x_2 coordinate is close to coordinate of C_{4X_2}, **then** Grade $= -0.063x_1 - 0.089x_2 + 32.655$
Rule 5: **If** x_1 coordinate is close to coordinate of C_{5X_1} and x_2 coordinate is close to coordinate of C_{5X_2}, **then** Grade $= -9.255x_1 + 7.981x_2 - 565.415$

where C_1-C_5 are the cluster centers. Each cluster is represented by a linear equation which is obtained from the weighted regression estimation (Table 2).

Performance Evaluation

A scatter plot of actual values against predicted values of fuzzy model is depicted in Figure 7. In the very successful estimation, predicted and actual values should lie on 45° line. For this case study, the coefficient of correlation "r" shows that the proposed fuzzy model has extremely high estimation capacity ($r = 0.94$) and it performs better than the Pham (1997). In addition to the "r" values, performance evaluations have been carried out based on root mean square of prediction error (RMSE) and standard deviations. For model validation, using lower-upper (LU) decomposition technique (Deutsch and Journel, 1998), the same data set evaluated in training was conditionally simulated and 81 data values were derived from this application. The conditioning data have the same variogram models of Clark's data (Clark, 1979). The index values are calculated and compared to the result of Pham (1997) in Table 3. We note that the RMSE of the proposed model is lowest, while the variability (standard deviation) of predicted values is closest to the variability of measured values.

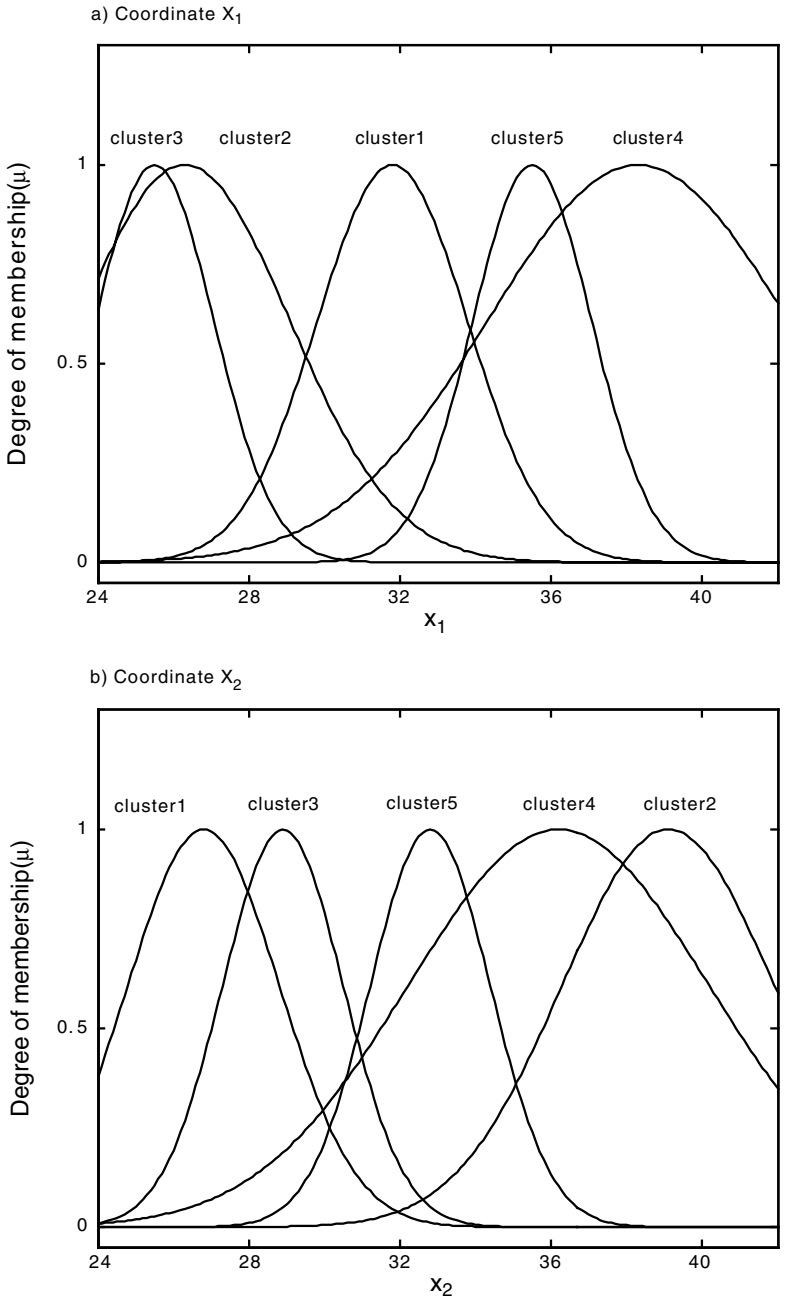

Figure 6. Input membership functions.

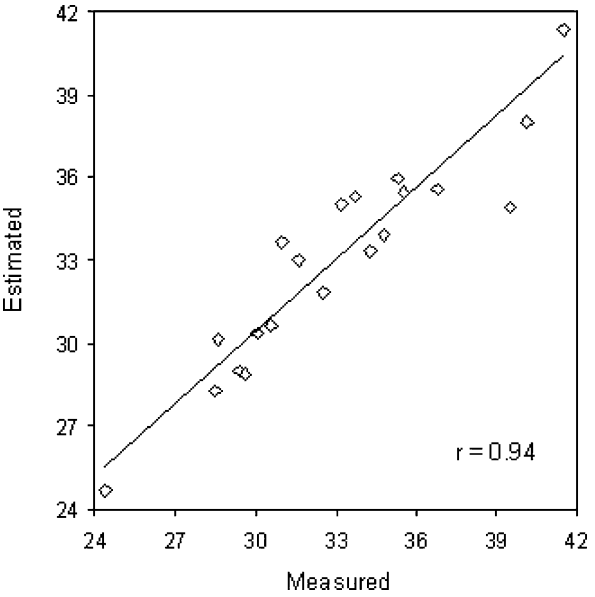

Figure 7. Scatter plot for measured and estimated values.

Example 2: A Case Study for Lignite Thickness Estimation

This section presents the application of the proposed fuzzy modeling approach to thickness estimation of a lignite deposit. Lignite reserve estimation refers to estimates for tonnage and averages of lignite quality parameters over an entire deposit. Estimation of lignite thicknesses, which is an important step for calculating the tonnage, has been carried out by the proposed fuzzy modeling approach.

Data Set

Sivas-Kalburcayiri field which is considered in this case study is one of the most productive basins in Anatolia, Turkey. Lignite seams in the field are used to

Table 3. Performance Comparisons for Simulated Data Set

INDEX	Measured	Model training	Model testing	Pham[a] 3 clusters	Pham 4 clusters	Pham 5 clusters	Pham 6 clusters
Standard deviation	4.51	4.22	3.70	0.65	0.88	1.83	1.85
RMSE	–	1.55	1.18	4.13	4.26	4.84	4.97

[a] Pham (1997).

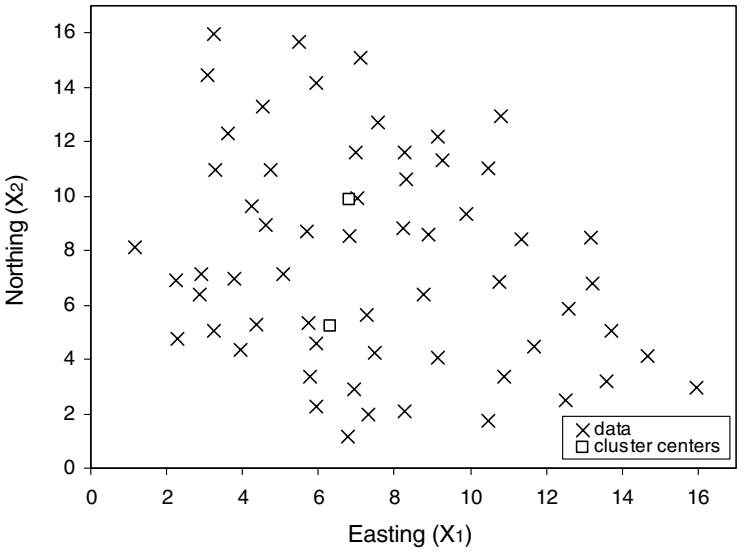

Figure 8. The Sivas-Kalburcayiri lignite deposit data set.

feed coal to a power plant. The locations of the 60 records for the upper sector of the field (Tercan and Karayigit, 2001) are shown in Figure 8. The lignite thickness data were scaled by using a linear transformation between 1.16 and 15.98.

Determining the Number of Clusters

Data were clustered by using FCM. Three dimensional clustering space was comprised of the Cartesian coordinates and thickness values. Optimal number of clusters was determined using the validity method presented in Section 3.1. Standard deviations of the cluster thickness values are plotted against the number of clusters (Fig. 9) and two clusters are retained as optimum (Fig. 8).

Point Semimadogram Analysis

The semimadogram analysis is used for defining the range of thickness around a cluster concerned. Using the semimadogram model, the zone of influence for each cluster was calculated. In this application, zone of influence (range) values were determined as 7 m and 9 m, respectively. Figure 10 shows the spherical model madograms.

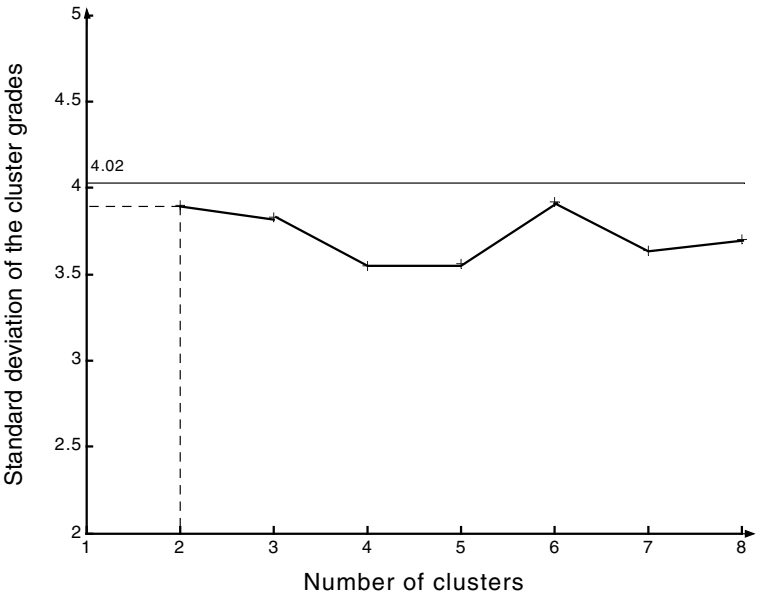

Figure 9. Cluster validity application.

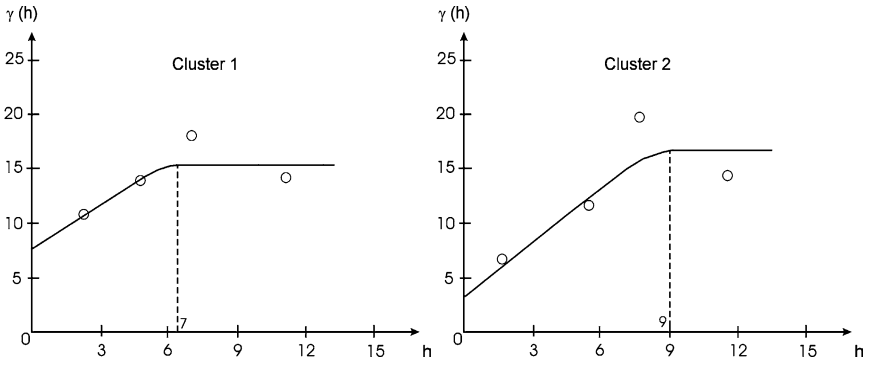

Figure 10. Semimadogram models for each cluster.

Table 4. Cluster Centers and Parameter Vector

Clusters	Coordinate x_1	Coordinate x_2	x_1	x_2	Constant
Cluster 1	6.31	5.22	2.119	3.762	−14.065
Cluster 2	6.82	9.86	0.157	0.749	−9.827

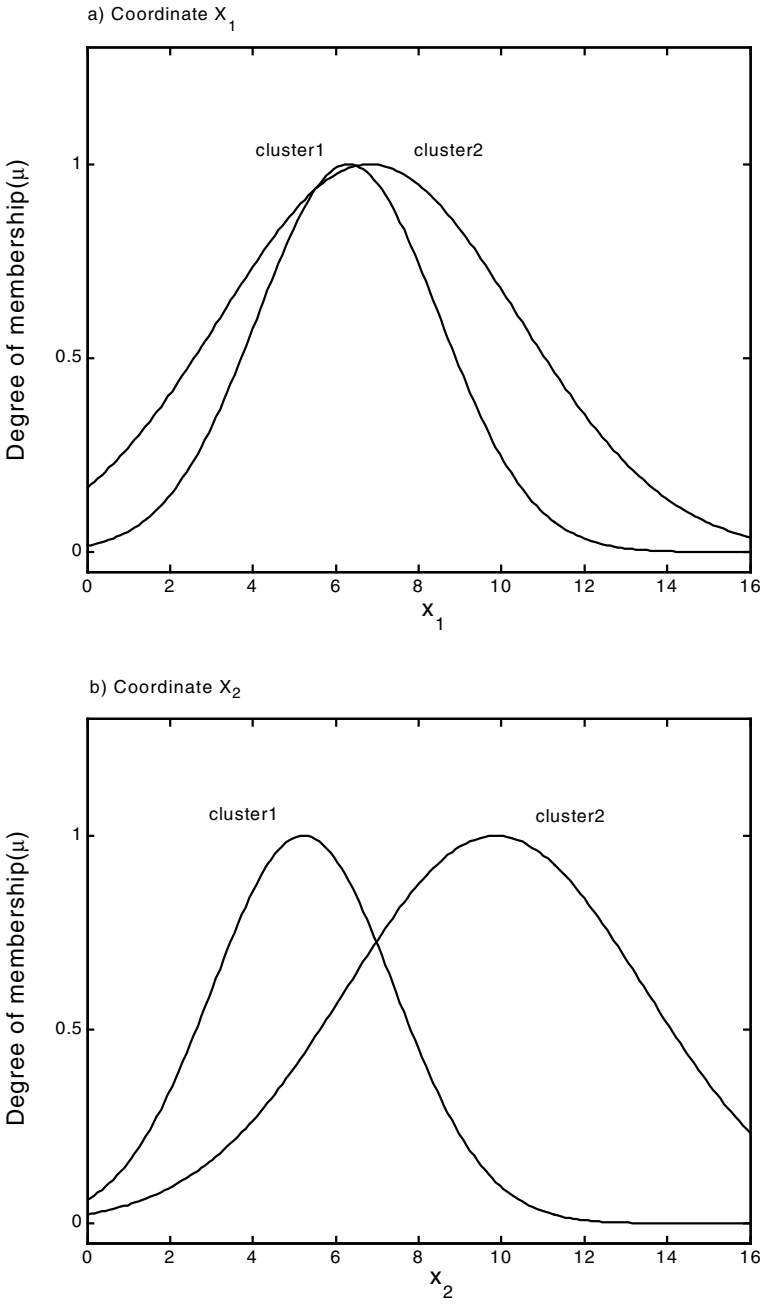

Figure 11. Input membership functions.

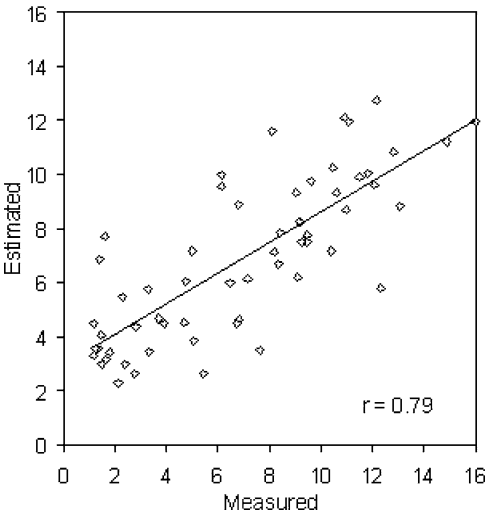

Figure 12. Scatter plot for measured and estimated values.

Clustering and Parameter Optimization

The PCM algorithm applied in this study used an initial partition matrix obtained from FCM. Clustering parameters were outlined below

$$\text{Number of cluster: } 2$$

$$\eta \text{ parameter vector: } \begin{bmatrix} 3.5 \\ 4.5 \end{bmatrix}_{2 \text{ by } 1}$$

$$\text{Weighting exponent: } m = 1.5$$

By using the input parameters listed above, clustering and local least square parameter optimization are carried out. Cluster centers and parameter estimations are listed in Table 4.

Rule-based Model

For designing the rule-based model, Gaussian type membership functions were modified in terms of the cluster centers and zone of influence. Input membership functions are shown Figure 11. The general rule system is described by the structure outlined below:

Table 5. Performance Evaluation for Real Data

Index	Measured	Estimated
Standard deviation	4.03	3.19
RMSE	–	2.46

Rule 1: **If** x_1 coordinate is close to coordinate of C_{1X_1} and x_2 coordinate is close to coordinate of C_{1X_2}, **then** Thickness $= 2.119x_1 + 3.762x_2 - 14.065$

Rule 2: **If** x_1 coordinate is close to coordinate of C_{2X_1} and x_2 coordinate is close to coordinate of C_{2X_2}, **then** Thickness $= 0.157x_1 - 0.749x_2 - 9.827$

where C_1-C_2 are the cluster centers. In the system, each cluster is associated with a linear regression equation which is obtained from the weighted least square estimation (Table 4).

Figure 12 shows a scatter plot of actual values against predicted values of fuzzy model. In addition, RMSE and standard deviations are calculated and given in Table 5. The variability (standard deviation) of predicted values can be accepted as close to the variability of the measured values.

CONCLUSIONS

This paper presented a systematic fuzzy modeling algorithm based on spatial variability for estimating mineral reserves. In the paper, point semimadogram function has been proposed for measuring the spatial variability, and a new cluster validity index has been introduced for determining the appropriate number of cluster. Spatial variability, which has crucial importance for reserve estimation, was taken into consideration all over the model.

The proposed modeling approach was tested by both simulated and real data. The results have shown the capability and the performance measures demonstrated the power of the proposed algorithm. For the future work, neuro-fuzzy modeling based on spatial variability could be considered.

ACKNOWLEDGMENTS

The authors would like to extend their appreciation to the Scientific and Technical Research Council of Turkey (Tubitak) and Münir Birsel Foundation.

REFERENCES

Babuska, R., 1998, Fuzzy modeling for control: Kluwer Academic Publication, Boston, 260 p.

Bardossy, A., Bogardi, I., and Kelly, E., 1990, Kriging with imprecise (fuzzy) variograms: Theory I., Application II.: Math. Geol., v. 22, p. 63–79.

Bardossy, Gy., Szabo, I. R., and Varga, G., 2003, A new method of resorce estimation for bauxite and other solid mineral deposits: Jour. Hungn Geomathematics, v. 1, p. 14–26.

Bardossy, Gy., and Fodor, J., 2004, Evaluation of uncertainties and risks in geology: Springer, 221 p.

Bardossy, Gy., and Fodor, J., 2005, Assessment of the completeness of mineral exploration by the application of fuzzy arithmetic and prior information: Acta Polytechnica Hung., v. 2, no. 1.

Bezdek, J. C., Ehrlich, R., and Full, W., 1984, FCM: the fuzzy c-means clustering algorithm: Computers & Geosciences, v. 10, no. 2–3, p. 191–203.

Clark, I., 1979, Practical Geostatistics: Applied Science Publishers, London, 281 p.

Deutsch, C. V., and Journel A. G., 1998, GSLIB: Geostatistical Software Library, Oxford, University Press, New York, 370 p.

Diehl, P., 1997, Quantification of the term geological assurance in coal classification using geostatistical methods: Schriftenreihe der GDMB Klassif. von Lagerstatten, v. 79, p. 87–203.

Fukuyama, Y., and Sugeno, M., 1989, A new method of choosing the number of cluster for fuzzy c-means method: In Proc. 5th Fuzzy System Symposium, p. 247–250.

Galatakis, M., Theodoridis, K., and Kouridou, O., 2002, Lignite quality estimation using ANN and adaptive neuro-fuzzy inference systems (ANFIS): APCOM, p. 425–431.

Gath, I., and Geva, A., 1989, Unsupervised optimal fuzzy clustering: IEEE Trans. Pattern Analysis and Machine Intelligence, no. 7, p. 773–781.

Goovaerts, P., 1997, Geostatistics for natural resources evaluation: Oxford University Press, New York, 483 p.

Henley, S., 2000, Resources, reserves and reality: Earth Science Computer Applications, v. 16, p. 1–2.

Henley, S., 2001, Geostatistics, cracks in the foundation: Earth Science Computer Applications, v. 17, p. 1–3.

Höppner, F., Klawonn, F., Kruse, R., and Runkler, T., 1999, Fuzzy cluster analysis: John Wiley, Chichester, 290 p.

Jain A., and Dubes, R., 1988, Algorithms for clustering data: Prentice Hall, Englewood Cliffs, 310 p.

Journel, A. G., and Huijbregts, Ch, J., 1978. Mining Geostatistics: Academic Press, London, 600 p.

Kaymak, U., and Babuska, R., 1995, Compatible cluster merging for fuzzy modeling: In Proceedings FUZZ-IEEE/IFES'95, Japan, p. 897–904.

Krishnapuram, R., and Keller, J. M., 1993, A possibilistic approach to clustering: IEEE Trans. Fuzzy Sys., v. 1, p. 98–110.

Krishnapuram, R., and Keller, J. M., 1996, The possibilistic C-means algorithm: insights and Recommendations: IEEE Trans. Fuzzy Syst., v. 4, no. 3, p. 385–393.

Kwon, S. H., 1998, Cluster validity index for fuzzy clustering: Electronics Letter, v. 34, no. 22, p. 2176–2177.

Luo, X., and Dimitrakopoulos, R., 2003, Data-driven fuzzy analysis in quantitative mineral resource assessment: Computers & Geosciences, v. 29, p. 3–13.

Pham, T. D., 1997. Grade estimation using fuzzy-set algorithms: Math. Geol., v. 29, no. 2, p. 291–305.

Pal, N. I., and Bezdek, J. C., 1995, On cluster validity for the fuzzy c-means model: IEEE Trans. Fuzzy Syst., v. 3, no. 8, p. 370–379.

Setnes, M., Babuska, R., and Verbruggen, H. B., 1998, Rule-based modeling: precision and Transparency: IEEE Trans. Syst, Man, and Cybernetics-Part C, v. 28, no. 1.

Şen, Z., 1989, Cumulative semivariogram models of regionalized variables: Math. Geol., v. 21, no. 8, p. 891–903.

Şen, Z., 1998, Point cumulative semivariogram for identification of heterogeneities in regional seismicity of Turkey: Math. Geol., v. 30, no. 7, p. 767–787.

Şen, Z., and Öztopal, A., 2001, Assessment of regional air pollution variability in Istanbul: Environmetrics, no. 12, p. 401–420.

Takagi, T., and Sugeno, M., 1985, Fuzzy Identification of systems and its applications to modeling and control: IEEE Trans. Systems, Man, and Cybernetics, v. 15, p. 116–132.

Tercan, A. E., and Karayigit, A. I., 2001, Estimation of lignite reserve in the Kalburcayiri field, Kangal Basin, Sivas, Turkey: Int. Jour. Coal Geol., v. 47, p. 91–100.

Timm, H., Borgelt, C., Döring, C., and Kruse, R., 2004, An extension to possibilistic fuzzy cluster analysis: Fuzzy Sets Syst., v. 147, p. 3–16.

Tutmez, B., 2005, Reserve Estimation Using Fuzzy Set Theory, Ph.D. Thesis: Hacettepe University, Ankara, 168 p.

Xie, X. L., and Beni, G. A., 1991, Validity measure for fuzzy clustering: IEEE Trans. Pattern Analysis and Machine Intelligence, v. 3, no. 8, p. 841–846.

Yager, R., and Filev, D. P., 1994, Essentials of fuzzy modeling and control: John Wiley, 388 p.