IEEJ TRANSACTIONS ON ELECTRICAL AND ELECTRONIC ENGINEERING
IEEJ Trans 2014; 9: 71–82
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI:10.1002/tee.21938

Paper

A Combination of Genetic Algorithm-based Fuzzy C-Means with a Convex Hull-based Regression for Real-Time Fuzzy Switching Regression Analysis: Application to Industrial Intelligent Data Analysis

Azizul Azhar Ramli*,**[a], Non-member
Junzo Watada**, Member
Witold Pedrycz***, Non-member

Processing an increasing volume of data, especially in industrial and manufacturing domains, calls for advanced tools of data analysis. Knowledge discovery is a process of analyzing data from different perspectives and summarizing the results into some useful and transparent findings. To address such challenges, a thorough extension and generalization of well-known techniques such as regression analysis becomes essential and highly advantageous. In this paper, we extend the concept of regression models so that they can handle hybrid data coming from various sources which quite often exhibit diverse levels of data quality. The major objective of this study is to develop a sound vehicle of a hybrid data analysis, which helps in reducing the computing time, especially in cases of real-time data processing. We propose an efficient real-time fuzzy switching regression analysis based on a genetic algorithm-based fuzzy C-means associated with a convex hull-based fuzzy regression approach. The method enables us to deal with situations when one has to deal with heterogeneous data which were derived from various database sources (distributed databases). In the proposed design, we emphasize a pivotal role of the convex hull approach, which is essential to alleviate the limitations of linear programming when being used in modeling of real-time systems.

Keywords: convex hull, fuzzy switching regression, genetic algorithm, heterogeneous data, intelligent data analysis, steam generator

Received 22 June 2012; Revised 15 November 2012

1. Introduction

Across a wide variety of fields, data are currently being collected and accumulated at a rapid pace. There is an urgent need for a new generation of computational approaches and tools to assist humans in extracting useful knowledge from the rapidly growing volumes of data [1, 2].

Furthermore, in real-world optimization problems that are widely encountered in engineering, management, economics, medicine, and numerous other disciplines, it is quite common to handle a large amount of data in a limited time [3, 4]. This critical situation imposes demanding requirements on industrial and manufacturing data analysis processes. The need for sophisticated data analysis tools is very noticeable, with an ultimate goal is to efficiently produce high-quality results. Closely related in this context, real-time data processing demonstrates an essential growth in relevance and visibility when dealing with timely managerial tasks, where we encounter evidence of time constraints. This situation supports the emergence of intelligent data analysis (IDA), especially for real-time data analysis which enables us to make decisions within allowable time limits.

Computational intelligence (CI) includes a concept of fuzzy regression, which exhibits advantages and strengths in its ability to deal with non-numeric data [5–7]. Fuzzy regression analysis exploits techniques of linear programming (LP) to explore and describe dependencies among variables [8]. In this context, LP is subject to constraints whose number is proportional to the number of samples (data points) to be considered when designing constructs of fuzzy regression; see Refs [9–11]. An increasing number of attributes and the size of the dataset itself might directly lead to increased computational complexity and the required processing time [12].

As a result, in recent years, various applications of so-called switching regression methods have become available. One visible feature of such models is that the data are generated by several different sources (distributed databases) over some time period [13]. Currently, such data analysis process is pursued mostly in the batch-processing mode.

Our intent is to exploit as well as to combine the concepts and algorithms of genetic-algorithm-based fuzzy C-means (GA-FCM) with a convex hull-based fuzzy regression approach in the implementation of real-time fuzzy switching regression. In addition, the adaptation of a convex hull approach, specifically the so-called Beneath-Beyond algorithm, helps to alleviate the limitations of the 'conventional' switching regression when pursuing real-time data processing. We show its efficient implementation and present results of experimental studies including steam generator industries and randomly generated synthetic datasets. In a nutshell, the ultimate objective here is to decrease, in real-time situations, the time required for data processing as well as to reduce the computational complexity in order to efficiently support the ensuing process of decision making. With

[a] Correspondence to: Azizul Azhar Ramli. E-mail: azizulr@uthm.edu.my

* Department of Software Engineering, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, 86400 Batu Pahat, Johor D/T, Malaysia

** Graduate School of Information, Production and Systems, Waseda University, Waseda University, 2-7 Hibikino, Wakamatsu-ku, Kitakyushu-shi, Fukuoka-ken 808-0135, Japan

*** Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta T6G 2V4, Canada and Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

the intention of demonstrating the optimization and computational details of the proposed method, illustrative examples are presented.

This paper is organized as follows. Section 2 offers a concise literature review, which focuses on the fuzzy switching regression models presented here. These models are constructed on the basis of methods such as FCM clustering and GA approaches, where the combination of both tools is used to cluster data and, in this manner, will reveal the structure in the data. Afterwards, a convex hull approach is discussed. Next, in Section 3, we present the real-time fuzzy switching regression model realized with the use of hybrid combinations of the GA-FCM algorithm and the convex hull-based fuzzy regression approach. Section 4 covers examples concerning real-world industrial data and generated synthetic data, while Section 5 presents further results and offers a pertinent discussion. Finally, Section 6 covers concluding remarks.

2. Related Works

IDA is an important tool to support decision-making activities. Some representative examples deal with problems of risk analysis, targeted marketing, customer retention, portfolio management, and brand loyalty.

Associated with such examples, the tools coming from the area of CI are endowed with an ability to construct models in the presence of noise while enhancing the interpretability of the models themselves. Let us recall that the CI comes as a unified conceptual and algorithmic vehicle embracing neurocomputing, fuzzy sets, and evolutionary optimization. The recent studies reported emphasize the need for forming a consistent methodology that supports the development of high-quality models and offers their further maintenance. One of the main requirements badly needed in successful industrial data analysis is to realize fast computing which implies that the processing to be involved should be characterized by a minimal computational complexity.

Regression models were initially developed as constructs that statistically describe the relationship among variables; that is, they explain one variable by making use of variation of some other (independent) variables [14]. The variables that are used to explain the other variables are called explanatory variables [15, 16].

In what follows, we introduce the notations pertinent to this study:

$i = 1, \ldots, n$	Index of samples/points
n	Number of samples/points
$j = 1, \ldots, K$	Index of attributes
K	Number of attributes
$l = 1, \ldots, m$	Index of vertices
m	Number of vertices
$A = [A_0, \ldots, A_k]$	Coefficient vector of the model
$A_j = (\alpha_j, c_j)$	The jth coefficient with center, α_j and spread c_j
$\alpha = [\alpha_0, \alpha_1, \ldots \alpha_K]$	Vector of center values
$c = [c_0, c_1, \ldots, c_K]$	Vector of spread values
Y_i	The expected fuzzy value of the model
$x_i = [x_{i0}, \ldots, x_{iK}]$	Independent variable vector, where $x_{i0} = 1, (i = 1, \ldots, n)$
y_i or $\tilde{y} = (y_i, d_i)$	Dependent variable with numeric or fuzzy value, $(i = 1, \ldots, n)$
P_{ij}	Points to be added to the convex hull
p_{ij}	Points of constructed convex hull

2.1. Fuzzy switching regression models

A generic regression analysis involves data that originate from a single data source. A single functional relationship between the independent or input variables $x \in \Re^K$ and the dependent or output variable $y \in \Re$ is assumed, and this relationship holds for all of the data collected. A general 'standard' regression model is then described as follows:

$$Y_i = h(x_i) + \varepsilon_i,\ i = 1, \ldots, n \tag{1}$$

where $h(\cdot)$ is a function and ε_i are independent random variables with a zero mean and some variance, $i = 1, \ldots, n$ [17].

An important assumption made in regression analysis is that the dataset to be analyzed is homogeneous in the sense that there is only a single functional relationship between exogenous and endogenous variables [18].

While this assumption may hold in many cases, we also encounter situations involving heterogeneous data. In addition, we might have some prior information as to the split (partition) of the overall dataset into some homogeneous subsets. Therefore, switching regression methods can be considered as a viable design alternative. Interestingly, switching regression has been applied to various fields such as manufacturing, economics, and biocomputing.

An implementation of switching regression is realized for a heterogeneous dataset by forming C homogeneous subsets of data and determining a regression function for each subset $k (k = 1, \ldots, C)$. In other words, a mixed distribution is given aimed at splitting this distribution into C homogeneous sets [19]. The performance criterion quantifies the squared metric distance differences between the estimated values y of the regression function observed in each subset and the corresponding experimental data. The criterion has to be minimized over all data subsets. Additionally, based on the study by Hathaway and Bezdek [20], a switching regression model has been discussed in varying details. Say, $S = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ becomes a set of data where each independent observation $x_k \in \Re^K$ has a corresponding dependent observation $y_k \in \Re$. In switching regression models, assume that the data are drawn from C models such as

$$y = f_i(\vec{x}; \vec{\delta}_i) + \varepsilon_i, 1 \le i \le C \tag{2}$$

where $f_i(\vec{x}, \vec{\delta}_i)$is a polynomial function about $\boldsymbol{x}$, each $\vec{\delta}_i \in \Omega_i \subset \Re^{k_i}, k_i \le n$, and each ε_i is a random vector with a mean vector $\mu_i = 0$ and a covariance matrix δ_i [21].

Generally, the basic idea of switching regression analysis has been mentioned in Hosmor's research, which describes an exemplified application of a fishery area [20]. According to this research, the parameters defining the linear growth curves for the genders can be estimated by treating the data analysis as a switching regression problem. The general models read as follows:

$$y_1 = f_1(x, \vec{\delta}_1) + \varepsilon_1 = \delta_{11}x + \delta_{12} + \varepsilon_1 \tag{3}$$

$$y_2 = f_2(x, \vec{\delta}_2) + \varepsilon_2 = \delta_{21}x + \delta_{22} + \varepsilon_2 \tag{4}$$

With the incorporation of fuzzy sets, an enhancement of the regression model comes in the form of a so-called fuzzy regression or a possibilistic regression, which was originally introduced by Tanaka *et al.* [4] (refer also to Ref. [16]) to reflect the relationship between the dependent variable and independent variables expressed in terms of fuzzy sets. The upper and lower regression boundaries that are used in the possibilistic regression reflect upon the possibilistic distribution of the output values. Associated with the previously discussed methods, fuzzy switching regression has been proposed by several researchers, including Hathway and Bezdek [20], Jajuga [22], and Quandt and Ramsey [23].

Fuzzy switching regression is a technique for estimating multiple fuzzy regression models for a dataset and, as such, it has been used for capturing nonlinear dependencies among selected input and output variables in many data mining applications. Initially, based on Wu *et al.* [24], Hathaway and Bezdek first combined switching regression with FCM and referred to the algorithm as fuzzy C-regression (FCR). FCR is a fuzzy clustering-based switching regression model where regression errors are also used to form a clustering criterion in the FCM clustering, such as the iterative optimization procedure. Moreover, related to Hathaway and Bezdek's approach [20], an estimate of $\{\delta_{ij}\}$ in (3) and (4), using mixture distributions, was employed. Each data point (x_k, y_k) is viewed as coming from regime 1 (corresponding to (3)) with probabilityφ and from regime 2 (corresponding with (4)) with probability $1-\varphi$. An assumption of ε_1 and ε_2 was also made, where both of them are independent for different data and the distributions of ε_1 and ε_2 are normal with mean 0 and unidentified standard deviations σ_1 and σ_2, respectively. Indicating the univariate normal probability density function with mean μ and standard deviation σ by

$$p(\varepsilon;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\varepsilon-\mu}{\sigma}\right)^2} \tag{5}$$

the following log-likelihood function of the samples in N is

$$\begin{aligned} &L(\varphi,\delta_{11},\delta_{12},\delta_{21},\delta_{22},\sigma_1,\sigma_2;S)\\ &=\sum_{k=1}^{n}\left\{\begin{array}{c}\log(\varphi p(y_k-\delta_{11}x_k-\delta_{12};0,\sigma_1)\\ +(1-\varphi)p(y_k-\delta_{21}x_k-\delta_{22};0,\sigma_2))\end{array}\right\} \end{aligned} \tag{6}$$

In addition, an allocation of a data point is shared by several regression models and is given by the fuzzy membership degree, which represents the weight of the data point in each regression model. Then the parameters of the regression models are estimated considering the weights present in each cluster.

The realization of fuzzy switching regression is completed in several phases. We start with heterogeneous data which is divided into several fuzzy sets. For each fuzzy set, a weighted regression is completed, where the weights of the corresponding data are given as the membership degrees of the data to the corresponding subsets. Usually, these models are applied when the results of a regression analysis are very poor for the overall dataset, e.g., the multiple correlation coefficient $\Re^2$ is much smaller than 1. However, it is assumed that a mixed distribution is given with a reasonable regression, which could be formed for each component of this distribution.

2.2. A review of fuzzy C-means clustering models

In real-world applications, there are no well-defined boundaries between clusters (groups of data); as a result, fuzzy clustering has emerged as a viable alternative [25, 26]. The membership degrees assuming values between 0 and 1 are used in fuzzy clustering to quantify partial assignment of data to clusters [27]. In addition, fuzzy clusters of the objects can be represented by a fuzzy partition matrix. The commonly encountered fuzzy clustering algorithm is the FCM [28]. The FCM algorithm is one of the most widely used methods in fuzzy clustering.

The FCM algorithm [29] can be summarized as follows: Let $X=\{x_1,\ldots,x_n\}$ be a set of given data, where each data point $x_k(k=1,\ldots,n)$ is a vector in $\Re^K$, U_{cn} is a set of real $c\times n$ matrices, and c is an integer, $2\le c\le n$. Then the fuzzy c-partition space for $X$is the set defined in the form

$$M_{fcn}=\left\{U\in\Re^{c\times n}|\sum_{i=1}^{c}U_{ik}=1, 0<\sum_{k=1}^{n}U_{ik}<n,\quad \text{and}\quad U_{ik}\in[0,1];1\le i\le c;1\le k\le n\right\} \tag{7}$$

The aim of the FCM algorithm is to find an optimal fuzzy c-partition and the corresponding prototypes minimizing the following objective function [30]:

$$J_m(U,V)=\sum_{i=1}^{c}\sum_{k=1}^{n}(U_{ik})^m D_{ik}^2(v_i,x_k) \tag{8}$$

where $U\in M_{fcn}$ is a fuzzy partition matrix, and $V=(v_1,\ldots,v_c)$ is a matrix of unknown cluster centers (prototype parameters or cluster centers). The notation $v_i\in\Re^K\forall_i$ and $D_{ik}(v_i,x_k)$ is a matrix of prototypes from $x_k$to the ith cluster prototype while m positioned in $[1,\infty)$ is the weighting exponent on each membership function, which influences the membership values. Moreover, Bezdek [29] highlighted that the Euclidean distance and the diagonal structure are used for all FCM results for $m=2$. Taken as a (U,V) minimizer of (8), this algorithm produces a sound cluster structure in X.

In addition, to minimize the criterion J_m by considering the unity constraint in (7), the FCM algorithm is expressed as an alternating minimization algorithm. Choose values for c,m, and ε, a small positive constant, then generate randomly a fuzzy c-partition U^0, and set the iteration number $t=0$. A two-step iterative process works as follows: Given the membership values $\mu_{ik}^{(t)}$, the cluster centers $v_i^{(t)}(i=1,\ldots,c)$ are calculated as

$$v_i^{(t)}=\frac{\sum_{k=1}^{n}(\mu_{ik}^{(t)})^m \mathbf{x}_k}{\sum_{k=1}^{n}(\mu_{ik}^{(t)})^m} \tag{9}$$

Given the new cluster centers $\mathbf{v}_i^{(t)}$, we update the membership values $\mu_{ik}^{(t)}$

$$\mu_{ik}^{t+1}=\left[\sum_{j=1}^{c}\left(\frac{\|\mathbf{x}_k-\mathbf{v}_i^{(t)}\|}{\|\mathbf{x}_k-\mathbf{v}_j^{(t)}\|}\right)^{\frac{2}{m-1}}\right]^{-1} \tag{10}$$

The process stops when $|U^{(t+1)}-U^{(t)}|\le\varepsilon$, or a predefined number of iterations has been reached. Additionally, an FCM algorithm is used for grouping the individuals, and the fitness of each individual is estimated according to membership values. Based on Alata *et al.*, the original FCM proposed by Bezdek is optimized using GA, and other values of the weighting exponent (rather than $m=2$) give less approximation error. Therefore, the least-square error is enhanced in most of the cases handled in this work. Also, the number of clusters is reduced [31].

2.3. Genetic-based clustering technique

Several researchers have employed genetic-based clustering techniques to solve various types of problems [32, 33]. More specifically, we exploit GAs to determine the prototypes of the clusters in the Euclidean space $\Re^K$. At each generation, a new set of prototypes is created through the process of selecting individuals according to their level of fitness. Subsequently, the individuals are affected by running genetic operators [33]. This process leads to the evolution of a population of individuals who become more suitable given the corresponding values of the fitness function.

There are a number of studies that utilize the advantages of the GA-enhanced FCM. We focus here on the genetically guided clustering algorithm proposed by Hall *et al.* Based on [30], in any generation, element i of the population is V_i, a $c\times s$ matrix of cluster centers (prototypes). The initial population of size I is constructed by a random assignment of real numbers to each of the$s$features of the $c$centers of the clusters. The initial values are constrained to be in the range of the attributes to which they were assigned (determined from the experimental datasets) but are otherwise random [30].

Additionally, there are several advantages due to the synergy of the GA and FCM. The time needed to reach an optimum through GA is less than the time needed by running the iterative approach. Also, GA provides higher resolution capability compared to the iterative search because of the fact that the precision depends on the step value in the 'for loop function'. So GA gives better performance and has less approximation error with less time [31]. Furthermore, GA is generally able to find the lowest known J_m value or a J_m associated with a partition very similar to that associated with the lowest value of J_m.

On datasets with several local extrema, the GA approach always avoids the less desirable solutions as well as help avoiding deterioration of partitions, which provides an effective method for optimizing clustering models whose objective function can be represented in terms of cluster centers. The time cost of genetically guided clustering is shown to make a series of random initializations of fuzzy/hard c-means, where the partition associated with the lowest J_m value is chosen, and is an effective competitor for many clustering domains [31].

It can be concluded that the time needed for the GA to optimize an objective function depends on the number and the length of the individual in the population and the number of parameters to be optimized [31].

In addition, as Vs will be used within the GA, it is necessary to reformulate the objective function for the FCM for optimization purposes [30]. Thus, the way the assignments are made to clusters is based on the following:

$$R_1(V) = \sum_{k=1}^{n} \min\{D_{1k}, D_{2k}, \ldots, D_{ck}\} \tag{11}$$

which is equivalent to a reformulation of J_1 that eliminates U.

Subsequently, (8) can be written down in terms of distances from the prototypes (as is being done in the FCM method) [30]. Specifically, for $m > 1$, as long as $D_{jk}(\mathbf{v}_j, \mathbf{x}_k) > 0\ \forall j, k$, we have

$$U_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{D_{ik}(\mathbf{v}_i, \mathbf{x}_k)}{D_{jk}(\mathbf{v}_j, \mathbf{x}_k)} \right)^{\frac{2}{m-1}}} \tag{12}$$

$$\text{for } 1 \le i \ge c;\ 1 \le k \le n$$

Hall *et al.* substituted (12) into (8); this gives rise to the FCM functional, reformulated as follows:

$$R_m(V) = \sum_{k=1}^{n} \left(\sum_{i=1}^{c} D_{ik}^{1/(1-m)} \right)^{1-m} \tag{13}$$

The intent is to optimize R_m with a genetically guided algorithm (GGA) [30].

In general, GGA consists of selecting parents for reproduction, performing crossover with the parents, and applying mutations to the children; in this case, Hall *et al.* used a binary gray code representation in which any two adjacent numbers are different by 1 bit. This method may yield a faster convergence and improved performance over a straightforward binary encoding. The complete process of the GGA is summarized as follows:

1) Choose m, c, and D_{ik}.
2) Randomly initialize I sets of c cluster centers. Constrain the initial values within the space defined by the data to be clustered.
3) Calculate R_m by using (13) for each population member and apply a modified objective function $R'_m(V) = R_m(V) = R_m(V) + b \times R_m(V)$, where $b \in [0, c]$ is the number of empty clusters.
4) Convert the population members to their binary equivalents (using the Gray code).
5) For $i = 1$ to number of generations, do

   a. Use k-fold tournament selection (default $k = 2$) to select $I/2$ parent pairs for reproduction;
   b. Complete a two-point crossover and bitwise mutation for each feature of the parent pairs;
   c. Calculate R_m by using (13) for each population member and apply a modified objective function $R'_m(V) = R_m(V) = R_m(V) + b \times R_m(V)$, where $b \in [0, c]$ is the number of empty clusters;
   d. Create a new generation of size I, which is selected from the two best members of the previous generation and the best children that are generated by using crossover and mutation.

6) Provide the cluster centers to the terminal population with the smallest R'_m value and report this smallest value.

In addition, an automatic setting of crossover and mutation rates is also used in this GGA. Hall *et al.* considered notation in which $f_{\max}$ is the maximum fitness in a population, $\bar{f}$ becomes the average fitness in a population, f is the fitness of an individual child about to have a mutation applied to it, and f' is the larger of two fairness values of individuals about to have crossover applied to them. Then, the probability of crossover, p_c, is given as

$$p_c = k_1(f_{\max} - f')/(f_{\max} - \bar{f}),\ f' \ge \bar{f}, \tag{14}$$

$$p_c = k_3,\ f' < \bar{f} \tag{15}$$

Moreover, the probability of mutation is specified in the form

$$p_m = k_2(f_{\max} - f)/(f_{\max} - \bar{f}),\ f \ge \bar{f}, \tag{16}$$

$$p_m = k_4,\ f < \bar{f} \tag{17}$$

where $k_1, k_2, k_3, k_4 \le 1.0$ [30].

Furthermore, closely related to the proposed idea, the implementation of the fuzzy switching regression requires fuzzy clustering to 'translate' the problem into a series of subproblems to be handled for the individual subsets of data [34].

2.4. Affine, convex hull, and supporting hyperplane

Let us recall that the affine hull of set S in the Euclidean space $\Re^K$ is the smallest affine set containing S or, equivalently, the intersection of all affine sets containing S. Here, an affine set is defined as the translation of a vector subspace. The affine hull aff(S) of S is the set of all affine combinations of elements of S, namely,

$$\mathrm{aff}(S) = \left\{ \sum_{j=1}^{K} \alpha_j x_j | x_j \in S, \alpha_j \in \Re, \alpha_j \ge 0, \sum_{j=1}^{K} \alpha_j = 1. \right\} \tag{18}$$

The convex hull of set S of points, hull(S), is defined to be a minimal convex set containing S. We say that point $P \in S$ is an extreme point of S if $P \notin \mathrm{hull}(S - P)$. In general, if S is finite, then hull(S) is a convex polygon while the extreme points of S are the corners of this polygon and the edges of this polygon will be referred to as the edges of the hull(S). Figure 1 shows how a convex hull polygon can be constructed.

We say that the extreme points of S have been identified, and hence hull(S) has been identified, if

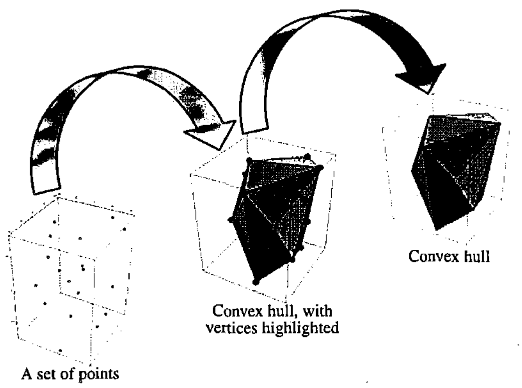

Fig. 1. Construction steps of a convex hull polygon

1) for each P_i containing a point of S, P_i has a Boolean variable 'extreme' that is true if and only if the point contained in P_i is an extreme point of S, and
2) for each P_i containing an extreme point of S, P_i contains the position of its point in the clockwise ordering, the total number of extreme points, and its adjacent extreme points in the clockwise order.

In general, this term is concerned with a concept of geometry. A hyperplane divides a space into two half-spaces. A hyperplane is said to support a set S in Euclidean space $\Re^K$ if it meets the following conditions [35]:

1) S is entirely contained in one of the two closed half-spaces of the hyperplane, and
2) S has at least one point located on the hyperplane.

In addition, if the dimension of the supporting line is higher than three, the related relationship can be written down as

$$S = \left(x \in \Re^K \mid \sum_{j=1}^{K} \alpha_j x_j = b \right) \tag{19}$$

where $\alpha = [\alpha_1, \dots, \alpha_K]$ denotes a unit vector, $\mathbf{x} = [x_1, \dots, x_K]$ is an arbitrary point, and b assumes any arbitrary real value.

$$S^+ = \left(x \in \Re^K \mid \sum_{j=1}^{K} \alpha_j x_j \geq b \right) \tag{20}$$

$$S^- = \left(x \in \Re^K \mid \sum_{j=1}^{K} \alpha_j x_j \leq b \right) \tag{21}$$

In the case when the following conditions

$$S \bigcap P \neq \phi \text{ and } P \subset S^+ \text{ or } P \subset S^-, \tag{22}$$

are satisfied, we say that the supporting hyperplane S supports the set P.

Therefore, using this definition, a convex hull, conv(P), can be expressed as follows:

$$\text{conv}(P) = \bigcap_{S^+:\text{upper-supporting-hyperplane}} S^+ \tag{23}$$

$$\text{conv}(P) = \bigcap_{S^-:\text{lower-supporting-hyperplane}} S^- \tag{24}$$

Beneath-Beyond Algorithm: This algorithm incrementally builds up the convex hull by keeping track of the current convex hull P_i using an incidence graph. For instance, the Beneath-Beyond algorithm consists of the following steps:

1) Select and sort points along one direction, say x_1. Let $s = P_0, P_1, \dots, P_{n-1}$ be input points after sorting. Process the points in an increasing order.
2) Take the first n points, which define a facet as the initial hull.
3) Let P_i be the point to be added to the hull at the ith stage. Let $P_i = \text{conv}(P_0, P_1, \dots, P_{i-1})$ be the convex hull polytope built so far. This step includes two kinds of hull updates:

   a. A pyramidal update is done when $P_i \notin \text{aff}(P_0, P_1, \dots P_{i-1})$, i.e. when P_i is not on the hyperplane defined by the current hull. A pyramidal update consists of adding a new node representing P_i to the incidence graph and connecting this node to all existing hull vertices by new edges.
   A non-pyramidal update is done when the above condition is not met, i.e. P_i is in the affine subspace defined by the current convex hull. In this case, faces that are visible from P_i are removed and new facets are created.

In addition, by processing a point in Quickhull [36], the randomized incremental algorithm is equivalent to using an implementation of the simplified Beneath-Beyond theorem based on Grunbaum's Beneath-Beyond Theorem proposed in 1961 [36].

Theorem 1. Let H be a convex hull in $\Re^K$ and let p be a point in $\Re^K - H$. Then the faces $f = \text{conv}(p \cup H)$ are as follows:

1) f is also a face of H if there is a facet F and H such that f is in F and p is below F.
2) f is not a face of H if $f = \text{conv}(p \cup f')$ with $f' \in H$, and either

   a. p is a linear combination of vertices of f', or
   b. p is above one facet of H containing f' and below another facet containing f'.

The rationale behind the first condition is straightforward. The second condition describes a face of the cone that is to be created if p is at least above one face. The ridge with one incident facet below and the other above p is the equivalent of the edge between visible and invisible faces for the discussed incremental algorithm above.

3. Real-Time Fuzzy Switching Regression Analysis in a Heterogeneous Dataset

Fuzzy switching regression applied to real-time scenarios deals with dynamic changes of data size. An adaptation of the GA-FCM [30] with the convex hull-based fuzzy regression approach called the Beneath-Beyond algorithm [37] in particular becomes of interest here. There are two major phases that are involved in the implementation of real-time fuzzy switching regression analysis: GA-FCM, which is concerned with the determination of the clusters in first phase, and the utilization of convex hull-based fuzzy regression approach, see Fig. 2.

Proceeding with more details, the complete procedure can be outlined as a series of the following steps.

Step 1. (Sample Selection)

Select raw data samples which are retrieved from distributed locality resources.

Step 2. (Build Clusters)

Given the data, construct the $C(1 < C < n)$ clusters. The number of clusters depends on the nature of the data. Figure 3 illustrates the structure of the data revealed through clustering. If the number of clusters is 'not sufficient' (inadequate/poor clusters

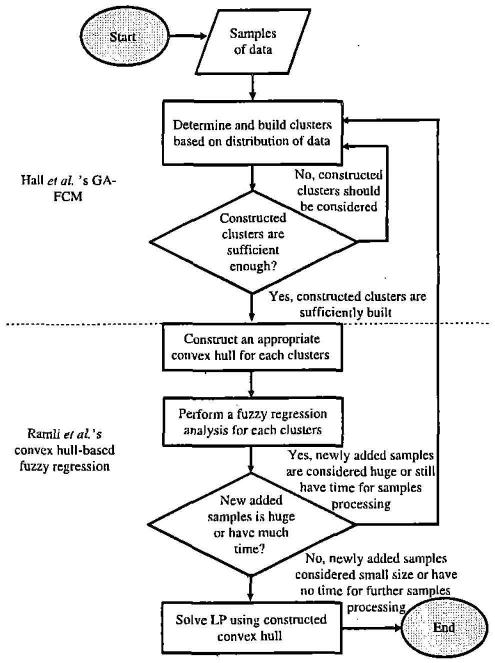

Fig. 2. Main processing phases of the proposed approach

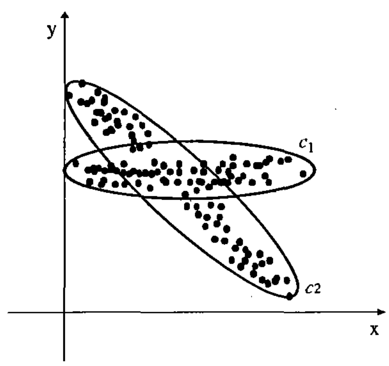

Fig. 3. Example of clusters formed on a basis of available data

fitting) to represent a structure in the data, it has to be increased. Given the relevance of the clustering phase, we use an augmented GA-based version of the FCM. In our case, we concentrate on Hall *et al.* [30] in which the GGA approach was considered.

Step 3. (Construction of Convex Hull)

The adaptation of the convex hull approach is realized by considering the outside points that were obtained when running the previous process. Such selected points will become vertices and bond each other to produce convex edges. The connected edges form a convex hull for the selected data.

Step 4. (Fuzzy Regression Analysis)

Fuzzy regression is realized on a basis of the constructed clusters. In other words, each cluster comes with its own fuzzy regression models. The example of fuzzy regression is illustrated in Fig. 4.

Step 5. (Process Newly Added Samples)

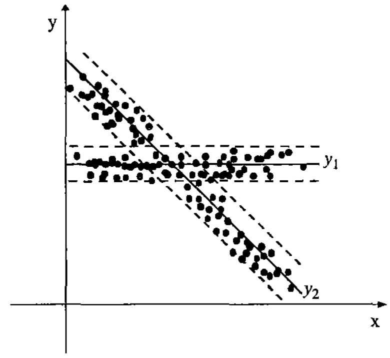

Fig. 4. Realization of the fuzzy regression based on fuzzy clusters

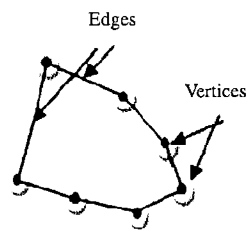

Fig. 5. Adaption of convex hull approach to graph

Related to real-time scenario, new data are added to the currently processed data. Therefore, if this group of data becomes huge or we require more processing time, the procedure will restart from Step 2; else it moves to the next step.

Step 6. (Solve LP)

By utilizing the convex hull developed so far, the implementation of the LP becomes easier because of the slightly changed (either increasing or decreasing) number of vertex points that must be considered for further computing. In other words, the analysis takes into consideration selected vertex points which were used for the convex hull polygon construction. These vertices are treated as constraints in the LP formulation and used to generate each of the regression models.

3.1. Solution to a problem with a convex hull approach The weaknesses of the implementation of multidimensional points in the fuzzy regression analysis can be addressed by the adaptation of a convex hull approach [37]. In the suggested modification, real-time data processing becomes realizable based on the constructed vertices of the convex hull using related points on the graph. The real-time implementation mostly deals with a large number of data. Each particular data will be represented as dynamic convex points and the related edges will be constructed as well. The adaptation of a convex hull approach is illustrated in Fig. 5.

The reconstruction process of the convex hull is also concerned with a new convex point or vertices which are located inside or outside the already developed convex hull. Certain related vertices may be removed and some new vertices will connect to the new points of the convex hull. Figure 6 shows an example of new edges being the result of the reconstruction process of the convex hull.

In order to obtain a suitable regression model based on the convex hull constructed in this way, the connected vertex points serve as constraints in the formulation of the LP problem. Considering this process, we note that the limited number of selected vertices will directly reduce the computational complexity

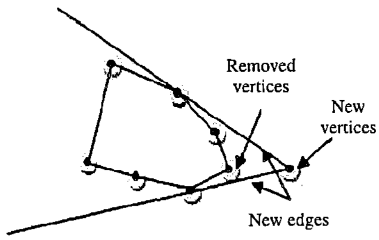

Fig. 6. Reconstruction of convex hull of graph

when forming a sound model while reducing the overall processing time.

Let us recall that the main purpose of fuzzy regression analysis is to form the upper and lower bounds of the linear regression model. Both the upper line Y^U and the lower line Y^L for fuzzy regression are expressed in the form

$$Y^U = \{A_0 + A_1x_1 + \ldots + A_Kx_K\}^U : \{\mathbf{A}\mathbf{x}_i^t\}^U = \alpha\mathbf{x}_i^t + \mathbf{c}|\mathbf{x}_i^t| \quad (25)$$

$$Y^L = \{A_0 + A_1x_1 + \ldots + A_Kx_K\}^L : \{\mathbf{A}\mathbf{x}_i^t\}^L = \alpha\mathbf{x}_i^t - \mathbf{c}|\mathbf{x}_i^t| \quad (26)$$

The related changes of the corresponding relationships can be expressed as follows:

1) Evaluation function

$$\min_{\alpha,c} \sum_{i=1}^{n} \sum_{j=2}^{K} c_j|P_{ij}|. \quad (27)$$

2) Constraints

$$P_{i1} \in Y_i \Leftrightarrow \begin{cases} P_{i1} \le \alpha_0 + c_0 + \sum_{j=2}^{K} \alpha_j P_{ij} + \sum_{j=2}^{K} c_j|P_{ij}| \\ P_{i1} \ge \alpha_0 - c_0 + \sum_{j=2}^{K} \alpha_j P_{ij} - \sum_{j=2}^{K} c_j|P_{ij}| \\ \qquad (i = 1, \ldots, n) \end{cases} \quad (28)$$

The above expression can be further rewritten as follows:

$$\begin{aligned} Y^U &= \{Y_i^U | i = 1, \ldots, n\} \\ Y^L &= \{Y_i^L | i = 1, \ldots, n\}. \end{aligned} \quad (29)$$

We also arrive at the following simple relations:

$$P_{i1} \le Y_i^U, \; P_{i1} \ge Y_i^L \; (i = 1, \ldots, n) \quad (30)$$

In addition, we know that any discrete topology is a topology that is formed by a collection of subsets of a topological spaceχ. The smallest topology has two open sets, the empty set ϕ and the universe χ. The largest topology contains all subsets as open sets and is called the discrete topology. In particular, every point in χ is an open set in the discrete topology. The discrete metric ρ on χ is defined by

$$\rho(x, y) = \begin{cases} 1 & \text{if } x \ne y \\ 0 & \text{if } x = y \end{cases} \quad (31)$$

for any$x, y \in X$. In this case (X, ρ) is called a discrete metric space or a space of isolated points.

According to the definition of discrete topology, (30) is rewritten as follows:

$$\begin{aligned} S(Y^U) &= \sum_{j=1}^{K} \{Y_i P_{ij}\}^U \ge 0 \\ S(Y^L) &= \sum_{j=1}^{K} \{Y_i P_{ij}\}^L \le 0 \end{aligned} \quad (32)$$

where we assume that $P_{i1} = 1$.

This expression corresponds to the definition of the support hyperplane. Under the consideration of the range below

$$S \bigcap P \ne \phi \text{ and } P \subset S^+ \text{ or } P \subset S^-, \quad (33)$$

the following relation is valid:

$$\bigcap S(Y^U) = \bigcap S(Y^L). \quad (34)$$

This is explained by the fact that the regression equations $Y^U$and $Y^L$are formulated by vertices of a convex hull. Therefore, it is explicit that the convex hull approach (or its vertices) can clearly define the discussed constraints of fuzzy mathematical programming, which is more reliable and accurate.

The convex hull is a smallest convex unit that contains given points. Let us denote the set of points given as input data as P, and the set of vertices of the convex hull as P_C where $P_C \subseteq P$. Therefore, a convex hull satisfies the following relationship:

$$\mathrm{conv}(P) = \mathrm{conv}(P_C) \quad (35)$$

Let us introduce the following set:

$$P_C = \{x_{Cl} \in \Re^K | l = 1, \ldots, m\} \subseteq P \quad (36)$$

where m is the number of vertices of the convex hull. Substituting this relation into dimension of the supporting line which is higher than 3, we encounter the following constraints:

$$P_{i1} \in Y_i \Leftrightarrow \begin{cases} P_{i1} \le \alpha_0 + c_0 + \sum_{j=2}^{K} \alpha_j P_{ij} + \sum_{j=2}^{K} c_j|P_{ij}| \\ P_{i1} \ge \alpha_0 - c_0 + \sum_{j=2}^{K} \alpha_j P_{ij} - \sum_{j=2}^{K} c_j|P_{ij}| \\ \qquad (i = 1, \ldots, m) \end{cases} \quad (37)$$

Using these, the constraints of the LP of the problem of fuzzy regression can be written down in the following manner:

$$y_i \in Y_i \Leftrightarrow \begin{cases} y_i \le \alpha x_i^t + c|x_i^t| \\ y_i \ge \alpha x_i^t - c|x_i^t| \\ (i = 1, \ldots, m). \end{cases} \quad (38)$$

Expression (38) is rewritten based on (28) where the index of samples $(i = 1, \ldots, n)$was changed into the index of vertices $(i = 1, \ldots, m)$. This occurs by virtue of (36). As previously mentioned, m here represents the number of selected points, which become the vertices of constructed convex hull polygon. In addition, (38) composes LP constrains in order to obtain the optimum model as regression models: upper and lower boundaries. Associated to the fuzzy switching regression analysis, this formulation of LP constraints will be employed when constructing clusters.

In other words, in order to form a suitable regression model based on the constructed convex hull, the connected vertex points will be used as the constraints in the LP formulation of the fuzzy regression. Considering this process, the limited number of selected vertices minimizes the computing complexity required to develop the model.

3.2. An efficient formulation of the real-time fuzzy switching regression through the convex hull approach

Based on the original method of the determination of the convex hull, the process has to be realized using all analyzed data. Here, we propose a new method by building a convex hull for such processed data. Focusing on the construction of an appropriate convex hull for each cluster process, the following substeps are completed:

Table I. Comparison between previous studies and the proposed approach of fuzzy switching regression implementation

No.	Previous studies of switching regression [20]	Newly proposed approach of fuzzy switching regression
i.	It is a huge constraint to deal with large-scale data volume, especially in real-time data processing implementations.	It is efficient in dealing with a large-scale data volume in a real-time data processing implementation.
ii.	With the new arrival of data, the past performance analysis is totally ignored and the analysis process has to restart again.	Decrease of computational time; it is not necessary to delicately compute the past performed analysis. The analysis can be built on the basis of the past results.
iii.	It is necessary to recalculate the LP performance considering the initially analyzed data and also the newly added data.	Reconstruction of convex hull edges is affected whether related newly analyzed data are selected or not selected as new convex hull vertices—either inside or outside of the remaining constructed convex hull.

1) Select the outsider points among the distributed analyzed data points. These points will become vertices of a convex polygon.
2) Connect each of the selected potential vertex points for constructing convex edges.
3) Connect constructed edges for producing boundaries of convex hull H_β.
4) Omit points that are included in the convex hull H_β.
5) Perform the Beneath-Beyond method to formulate the convex hull H_0 using one of the selected vertex points that were chosen for building the convex hull.

It is useful to choose the Beneath-Beyond algorithm to realize the convex hull approach considering that no extra computing is required for the construction of the facet structure. This algorithm may reduce the computational time required to obtain the best solution for an equivalent problem. In addition, the main criterion that distinguishes the different variants of this algorithm is related to dealing with a search for the visible facets. That is, we can find a visible facet among the facets added in the previous stage; therefore, we may simply search through all of the latter facets until a visible facet has been found. Then we examine adjacent facets and repeat the process on those that are visible.

Table I gives a general comparison between the previously studied approaches [20] and our newly proposed idea related to fuzzy switching regression implementation.

Referring to the explanation offered above, it becomes apparent that real-time fuzzy switching regression analysis can be improved by adopting a convex hull approach, or, more specifically, a Beneath-Beyond algorithm. Considering that the changes of edges depend on the outer plot positions, we can show that the proposed approach is more suitable to deal with real-time data implementations.

4. Illustrative Examples

To demonstrate the process of real-time data processing, we have chosen a set of real-world data that were obtained from a certain heavy industry company and a set of randomly generated synthetic data. Each of the selected samples of data is divided into two groups. This process is considered to show the simulation of a real-time situation, in which the first group is analyzed at the beginning of the procedure and the remaining data were added next, which mimics the real-time scenario. In addition, we also developed the conventional switching regression [20] for the same selected data. We discuss the efficiency of our method to construct a fuzzy switching regression model for one real-world case as well as randomly generated synthetic sample cases. Some comparative analysis is also provided.

Example 1: Performance monitoring of a steam generator involves a continuous evaluation of the plant's efficiency over time using data supplied by sensors. These evaluations are repeated in constant time intervals using data coming from online instrumentation. According to Buljubaic and Delalic, the objective of performance monitoring is a continuous evaluation of degradation such as decrease in performance of the steam generator [38]. In addition, these results require additional information that is helpful in problem identification, improvement of plant performance, and making economic decisions about maintenance schedules.

We use a set of data resulting from a simulation of a steam generator (SG) and, in this case, we focus on steam boiler samples. The steam, which drives the steam turbine, is generated by heating water. In addition, depending on the power of the reactor, two, three or four steam generators are provided and, together with the reactor, are installed in the hermetically sealed doubled-walled reactor building. The efficiency of a steam boiler depends on the heat transfer properties from the primary to the secondary fluid [39].

The data set consists of 400 samples forming an initial group that consists of three variables describing the input–output relationship; drum pressure: the pressure inside the steam drum (PSI), x_1; oxygen exhaust: excess oxygen in the exhaust gas (%), x_2; and steam flow: the rate of steam flow from the steam drum (kg/s), y. Additionally, this analysis will identify a cluster or group of steam flows that achieved the standard and vice versa.

We ran GA-FCM clustering with two clusters and constructed convex hull-based fuzzy regression for each of them. This process gives rise to fuzzy regression boundaries that cover the entire set of analyzed data. Therefore, the result will be the optimal convex hull polygon that can be constructed using selected outside vertices.

In general, the first cluster shows that 14 vertices were selected to be connected convex edges to construct a convex hull. Another cluster (second cluster) comes with 12 vertices, which were selected to establish convex edges and their combination generates a convex hull.

Given a real-time processing scenario, let us assume that 100 samples were added to the previously analyzed data and that all of the data need to be processed again; the entire 500 samples of the data set are further processed.

We may anticipate that the newly added data are distributed inside the remaining convex hull. We have found that only minor changes occurred, resulting in a total of 15 vertices formed during the reconstruction process of the convex hull for the first cluster. On the other hand, the second cluster in this example also has been impacted, and in this case we have arrived at 11 vertices.

Finally, the selected vertices, which are used for the convex hull polygon for each cluster, will become a constraint in the formulation of the LP used to construct the fuzzy regression.

Example 2: In general, the synthetic sample sets are generated to meet specific needs or certain conditions that may not be found in the original real-world scenario. This process can be useful when designing any type of system because the synthetic data are used as a simulation or as a theoretical value or situation. Moreover, this condition allows us to take into account unexpected results and to have a basic solution or remedy, if the results prove to be unsatisfactory [40].

Given the capabilities discussed in this example, we consider carrying out random generation of synthetic data to produce a set of sample data within the assumption of heterogeneous sources and locations. In addition, the data samples generated will consist of two inputs, x_1 and x_2, variables with the intention of forecasting an output, y, variable.

To demonstrate the dynamic changes of data representation in the context of real-time data processing, as explained above, we initially ran the proposed approach for 750 samples. Based on the results obtained, we noted that the constructed convex hull for the first cluster consisted of 52 convex vertices while another one (second cluster) had 54 convex vertices. Therefore, by utilizing these vertex points as constraints in the LP formulation, we can directly form an appropriate regression model for each cluster.

Next, we add 300 more samples and process them in the same manner as discussed in the previous example. Based on the results obtained, we notice that there are several changes that affect the initially constructed convex hull. These changes are to the constructed convex hull due to the distribution of the newly added samples. The reconstructed convex hull for the first cluster was modified and comes with a small number of decrements, which happen as a change is made from 52 (as previously highlighted) to 51 vertices points. The second convex hull, corresponding to the second cluster, has been slightly changed, and now the total number of vertices becomes 53. In addition, more than 50% of the vertices were considered in the reconstruction process of the convex hulls for both clusters, which consider newly added samples of data that are reused from the earlier process.

Therefore, referring to both numerical examples, we conclude that the employment of the proposed method is highly beneficial, especially for the dynamic database environment due to the minimal time usage as well as the lower computational overhead.

5. Discussion

The increase in sample size might cause computational difficulties in the implementation of the LP problem. Another issue might emerge when changes occur with regard to the variables themselves. Thus, the entire set of constraints must be reformulated, which increases the overall computational overhead. The increase in the computing complexity has been alleviated by the use of the proposed method.

Before going further into this precious section, we indicate here the computer specification that has been used to perform the whole processes. The specifications of machine are as follows: a personal notebook PC with Intel(R) Pentium CORE(TM) Duo 2 CPU (2.00 GHz) processors combined with 2 GB DDR2 type of RAM. Moreover, Windows Vista Business Edition (32 bit) was the operating system installed in this machine.

For comparative purposes, Table II summarizes the results provided by the proposed fuzzy switching regression and the conventional switching regression approach for the two numerical examples covered in the previous section. In addition, the approach by Hathaway and Bezdek [20] was considered here as the conventional one.

We may conclude that most of the obtained proposed fuzzy switching regression models are not very much different, which indicates that there were only slight changes in analyzed data as well as in the constructed convex hull polygons. Therefore, the newly added data will not influence the regression models excessively and the produced models become more accurate because the reconstructed convex hull automatically covers all points of the analyzed data. In other words, we do not have to consider the complete data to construct the regression models; we only utilize the selected vertices that are used for the construction of the convex hull. Therefore, this situation will lead to a decrease in computational load.

Moreover, based on the reported computing time, we note a substantial reduction in time required to construct regression when using the method introduced in this study. Referring to Table II, focusing on the *Time Required* column, we highlighted that the differentiation of the time interval between two completed iteration cycles is only 0.24 s or 0.093% for the total of essential duration used while initial group of data processes for the steam generator sample sets. Conversely, merely 0.49 s or 0.138% of 3.49 s (total processing time of the initial set of synthetic samples) is added. Compared to the conventional approach, both the analyzed samples required more time, along with an increase in the sample volume. We may anticipate that the differences could be become substantially more when dealing with larger data sets.

Closely related with computing time, we also present here *Inference Speed* criteria. Based on this comparison criterion, we assumed a conventional switching regression approach as 100% inference speed; therefore we noticed that an inference speed for the steam generator dataset is 129% faster through the proposed approach compared to the selected conventional one. Conversely, 134% greater inference speed has been achieved for the synthetic dataset. Thus, we can simply generalize here that, with more data to be processed, thus, the percentage of inference speed greatly increases.

Apart from Table II, we highlight also some interesting features in term of *Accuracy Level* viewpoint. Related to this aspect, suppose that those produced conventional switching regression models are considered 100% accurate and are taken as evaluation benchmark. Generally, we can clearly distinguish that those regression models that are obtained from our proposed approach are very much similar to the models produced through a conventional one. Hence, the qualities of accuracy for those models formed via the proposed approach are considered to be superior. In addition, the most important feature that relates to the proposed approach is with regard to the utilization of the fuzzy concept along with switching regression implementation process. As was clarified in the previous section, this combination produced a great implication especially associated with the accuracy intensity of the produced models even in real-time circumstances.

On the other hand, related to the computational complexity factor for the subsequent iterations, our new approach will only consider the newly added samples of data together with the selected vertices of the previous convex hull. For that reason, this computing scenario will reduce the computational complexity because of the smaller number of the analyzed samples of data used for the subsequent processing of regression models. These results illustrate the potential of our proposed model to solve switching regression problems that demand incremental adaptability.

6. Concluding Remarks

In this study, we have reported on the development of fuzzy switching regression models that can be regarded as a potential IDA tool for an array of essential problems in real-time data processing, especially problems encountered in the industry and manufacturing fields.

We have developed the enhancement of fuzzy switching regression, which comes as a hybrid GA-FCM with the convex hull-based fuzzy regression approach, specifically the Beneath-Beyond algorithm. In real-time processing, where we are faced with dynamically modified data, the proposed algorithm realizes fuzzy switching regression by reconstructing particular edges and considering new vertices for which the recomputing has to be realized.

Table II. Fuzzy switching and conventional regression models: details of performance comparisons

Regression approach	Dataset	Group sample	Obtained regression	Time required models (s)	Inference speed	Accuracy level
Proposed fuzzy switching regression approach	Steam Generator	400 samples	$y_1 = (23.4, 2.4) + (3.5, 1.9)x_1 + (1.0, 1.1)x_2$ $y_2 = -(21.1, 0.0) - (3.5, 0.0)x_1 - (12.2, 0.0)x_2$	2.58	126%	99%
		500 samples	$y_1 = (22.3, 2.4) + (4.0, 1.6)x_1 + (1.1, 1.5)x_2$ $y_2 = -(19.5, 0.0) - (4.0, 0.0)x_1 - (12.4, 0.0)x_2$	2.82	129%	99%
	Synthetic data	750 samples	$y_1 = (13.8, 2.1) + (5.5, 2.4)x_1 + (7.6, 2.2)x_2$ $y_2 = -(15.3, 0.0) - (6.2, 0.0)x_1 - (9.2, 0.0)x_2$	3.52	120%	98%
		1050 samples	$y_1 = (15.6, 2.3) + (5.0, 2.6)x_1 + (8.6, 2.3)x_2$ $y_2 = -(20.7, 0.0) - (6.0, 0.0)x_1 - (10.3, 0.0)x_2$	4.01	134%	98%
Conventional switching regression approach [15]	Steam Generator	400 samples	$y_1 = 23.7 + 1.0x_1 + 0.1x_2$ $y_2 = -20.2 - 3.0x_1 - 10.8x_2$	3.49		
		500 samples	$y_1 = 23.0 + 1.0x_1 + 0.2x_2$ $y_2 = -19.2 - 3.2x_1 - 11.9x_2$	3.97	100%	100%
	Synthetic data	750 samples	$y_1 = 15.0 + 5.3x_1 + 6.1x_2$ $y_2 = -16.4 - 5.3x_1 - 9.3x_2$	4.38		
		1050 samples	$y_1 = 16.1 + 5.5x_1 + 7.1x_2$ $y_2 = -18.6 - 4.8x_1 - 11.0x_2$	6.05		

$x_1 =$ input variable for drum pressure, and $x_2 =$ input variable for oxygen exhaust for both implemented regression approach. Note that $y_1 = (23.4, 2.4) + (3.5, 1.9)x_1 + (1.0, 1.1)x_2$ represents the upper line, while $y_2 = -(21.1, 0.0) - (3.5, 0.0)x_1 - (12.2, 0.0)x_2$ is the lower line obtained for steam generator and synthetic dataset via implementation of proposed fuzzy switching regression approach. This is valid for those obtained regression models in Table II.

The implementation of the proposed method clearly highlights that the constructed convex hull becomes a data boundary inside which the analyzed data points are found. Essentially, this proposed hybrid approach becomes an alternative to real-time fuzzy switching regression analysis.

Furthermore, we have demonstrated that the convex hull-based fuzzy regression approach performs efficiently for real-time data processing for fuzzy switching regression problems; this method could recompute and reconstruct the related edges by taking into account newly added data to form suitable regression models. On the other hand, we also showed that the number of obtained vertices of the convex hull edifice will not drastically change (either increase or decrease), thereby retaining the computing effort relatively constant in spite of an increasing number of samples.

These results suggest that the proposed method can be applied to real-world large-scale systems, especially those operating in real-time computing environments. In addition, the proposed method will not lead to repetition in computing, as it focuses only on newly arriving data, which potentially could become new vertices. This strategy becomes suitable for the implementation of real-time fuzzy switching regression where the convex hull can effectively handle new data with a low computational overhead. Due to this promising observation, the resulting real-time fuzzy switching regression will help reduce the computational overhead, thus decreasing the overall processing time. Although in this paper we dealt with small samples of datasets (and this has been done for illustrative purposes), it is worth noting that the method scales up quite easily.

In conclusion, our goal was to establish a practical approach to solving fuzzy switching regression by implementing the hybrid strategy of the GA-FCM algorithm as well as convex hull-based fuzzy regression. We offered some evidence showing that this method performs as a randomized incremental algorithm that is truly output sensitive to the number of vertices. In addition, the approach requires less space compared to most of the randomized incremental algorithms and executes faster for inputs with non-extreme points, especially when dealing with real-time data processing. We may envision that such fuzzy switching regression could become an efficient vehicle for analyzing real-world data where ambiguity or fuzziness cannot be avoided.

As future directions for this proposed approach, issues that could be pursued include cut-off time for dynamic data situations and the consideration of data preprocessing procedures for real time data analysis implementations. Furthermore, the proposed approach might be able to enhance nonlinear regression analysis with adaptation to other intelligent techniques of data analysis.

Acknowledgments

The first author was supported by Malaysian Ministry of Higher Education (MOHE) under Skim Latihan Akademik IPTA-UTHM (SLAI-UTHM) scholarship program at the Graduate School of Information, Production and Systems (IPS), Waseda University, Fukuoka, Japan.

References

(1) Suzuki Y, Kawamura A, Saga S, Maeda J. Perceptual clustering with fuzzy encoding. *IEEJ Transactions on Electronics, Information and Systems* 2003; **123(1)**:144–149.

(2) Jalal C, Mahmoud TS, Rezaei PH. An interval-based approach to fuzzy regression for fuzzy input-output data. *IEEE International Conference on Fuzzy Systems*, Taipei, Taiwan, June 2011; 27–30.

(3) Kahraman C, Beşkese A, Bozbura FT. Fuzzy regression approach and applications. In *Studies in Fuzziness and Soft Computing 201*. Kahraman C (ed). Springer-Verlag: Berlin, Heidelberg; 2006; 589–615.
(4) Tanaka H, Uejima S, Asai K. Linear regression analysis with fuzzy model. *IEEE Transaction on Systems, Man, and Cybernetics* 1982; **12(6)**:903–907.
(5) Stahl C. A strong consistent least-squares estimator in a linear fuzzy regression model with fuzzy parameters and fuzzy dependent variables. *Fuzzy Sets and Systems* 2006; **157(19)**:2593–2607.
(6) Wang H-F, Tsaur R-C. Insight of a fuzzy regression model. *Fuzzy Sets and Systems* 2000; **112(3)**:355–369.
(7) Watada J, Pedrycz W. A fuzzy regression approach to acquisition of linguistic rules. In *Handbook of Granular Computing*. Pedrycz W, Skowron A, Kreinovich V (eds). John Wiley & Sons: Chichester; 2008; 719–732.
(8) Bekir C, A. Firat Ö. LP methods for fuzzy regression and a new approach. In *Synergies of Soft Computing and Statistics for Intelligent Data Analysis: Advances in Intelligent Systems and Computing*, vol. 190. Kruse R, Berthold MR, Moewes C, Gil MÁ, Grzegorzewski P, Hryniewicz O (eds). Springer-Verlag: Berlin, Heidelberg; 2013; 183–191.
(9) Yao C-C, Yu P-T. Fuzzy regression based on asymmetric support vector machines. *Applied Mathematics and Computation* 2006; **182(1)**:175–193.
(10) Sakawa M, Yano H. Fuzzy linear regression analysis for fuzzy input-output data. *Information Science* 1992; **63(3)**:191–206.
(11) D'Urso P, Massari R, Santoro A. Robust fuzzy regression analysis. *Information Science* 2011; **181(19)**:4154–4174.
(12) Lee CL, Liu A, Chen W. Pattern discovery of fuzzy time series for financial prediction. *IEEE Transactions of Knowledge and Data Engineering* 2006; **18(5)**:613–625.
(13) Shen H, Yang J, Wang S. Outlier detecting in fuzzy switching regression models. In *Lecture Notes in Computer Science*, vol. 3192. Springer; 2004; 208–215.
(14) Yabuchi Y, Watada J. Fuzzy robust regression model by possibility maximization. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 2011; **15(4)**:479–484.
(15) Bargiela A, Pedrycz W, Nakashima T. Multiple regression with fuzzy data. *Fuzzy Sets and Systems* 2007; **158(19)**:2169–2188.
(16) Tanaka H, Guo P. Portfolio selections based on upper and lower exponential possibility distributions. *European Journal of Operational Research* 1999; **114(1)**:115–126.
(17) Nong X. A new fuzzy linear regression model for least square estimate. *Information and Business Intelligence: Communications in Computer and Information Science* 2012; **268(5)**:709–715.
(18) Bisserier A, Boukezzoula R, Galichet S. A revisited approach to linear fuzzy regression using trapezoidal fuzzy intervals. *Information Sciences* 2010; **180(19)**:3653–3673.
(19) Castano S, De Antonellis V, De Capitani di Vimercati S. Global viewing of heterogeneous data sources. *IEEE Transactions on Knowledge and Data Engineering* 2001; **13(2)**:277–297.
(20) Hathaway RJ, Bezdek JC. Switching regression models and fuzzy clustering. *IEEE Transactions on Fuzzy Systems* 1993; **1(3)**: 195–204.
(21) Wang S-T, Jiang H-F, Lu H-J. Integrated fuzzy clustering algorithm GFC for switching regressions. *Ruan Jian Xue Bao/Journal of Software* 2002; **13(10)**:1905–1914.
(22) Jajuga K. Linear fuzzy regression. *Fuzzy Sets and Systems* 1986; **20(3)**:343–353.
(23) Quandt RE, Ramsey JB. Estimating mixtures of normal distributions and switching regressions. *Journal of the American Statistical Association* 1978; **73(364)**:730–738.
(24) Wu K-L, Yang M-S, Hsieh J-N. Alternative fuzzy switching regression. *International Multiconference of Engineers and Computer Scientists (IMECS 2009)*, vol. 1, Hong Kong, March 2009; 18–20.
(25) Peters G. Rough clustering and regression analysis. In *RSKT 2007, Rough Sets and Knowledge Technology, Lecture Notes in Computer Science*, vol. 4481. Yao JT, Lingras P, Wu W-Z, Szczuka M, Cercone NJ, Ślęzak D (eds). Springer-Verlag: Berlin, Heidelberg; 2007; 292–299.
(26) Li MJ, Ng MK, Cheung Y-M, Huang JZ. Agglomerative fuzzy K-means clustering algorithm with selection of number of clusters. *IEEE Transactions on Knowledge and Data Engineering* 2008; **20(11)**: 1519–1534.
(27) Tanaka F, Suzuki Y, Maeda J. Autoregressive model using fuzzy c-regression model clustering for traffic modeling. ***IEEJ** Transactions on Electronics, Information and Systems* 2003; **123(3)**: 631–632.
(28) Ruspini EH. A new approach to clustering. *Information Control* 1969; **15(1)**:22–32.
(29) Bezdek JC. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic Publishers: Norwell, MA; 1981.
(30) Hall LO, Ozyurt IB, Bezdek JC. Clustering with a genetically optimized approach. *IEEE Transactions on Evolutionary Computation* 1999; **3(2)**:103–112.
(31) Alata M, Molhim M, Ramini A. Optimizing fuzzy C means clustering algorithm using GA. *World Academy of Science, Engineering and Technology* 2008; **39(2)**:224–229.
(32) Lin HJ, Yang FW, Kao YT. An efficient GA-based clustering technique. *Tamkang Journal of Science and Engineering* 2005; **8(2)**: 113–122.
(33) Wang Y. Fuzzy clustering analysis by using genetic algorithm. *ICIC Express Letters* 2008; **2**(4):331–337.
(34) Maulik U, Bandyopadhyay S. Genetic algorithm-based clustering technique. *The Journal of the Pattern Recognition* 2000; **33(9)**: 1455–1465.
(35) Boissonnat J-D, Cerezo A, Devillers O, Duquesne J, Yvinec M. An algorithm for constructing the convex hull of a set of spheres in dimension d. *Computational Geometry Theory Application* 1996; **6(2)**:123–130.
(36) Bradford Barber C, Dobkin David P, Huhdanpaa H. The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software* 1996; **22**(4):469–483.
(37) Ramli AA, Watada J, Pedrycz W. Real-time fuzzy regression analysis: a convex hull approach. *European Journal of Operational Research* 2011; **210(3)**:606–617.
(38) Buljubai I, Delalic S. Improvement of steam boiler plant efficiency based on results of on-line performance monitoring. *Technical Gazette* 2008; **15(3)**:29–33.
(39) de Mesquita RN, Ting DKS, Cabral ELL, Upadhyaya BR. Classification of steam generator tube defects for real-time application using eddy current test data and self-organizing maps. *Real-Time Systems* 2004; **27(1)**:49–70.
(40) Abowd JM, Lane JI. New approaches to confidentiality protection: synthetic data, remote access and research data centers. In *Privacy in Statistical Databases PSD 2004, Lecture Notes in Computer Science 3050*. Domingo-Ferrer J, Torra V (eds). Springer-Verlag: Berlin; 2004; 282–289.

Azizul Azhar Ramli (Non-Member) received the B.Sc. degree in computer science from Universiti Teknologi Malaysia, Malaysia, and the M.Sc. degree in intelligence system form Universiti Utara Malaysia, Malaysia in 2002, and 2004, respectively. He is currently pursuing the Ph.D. degree at Graduate School of Information, production and Systems (IPS), Waseda University, Japan. His research interests include management information system, data mining, expert system, fuzzy modeling, and real-time analysis.

Junzo Watada (Member) received the B.Sc. and M.Sc. degrees in electrical engineering from Osaka City University, Japan, and the Ph.D. degree from Osaka Prefecture University, Japan. Currently he is a Professor of Management Engineering, Knowledge Engineering and Soft Computing at Graduate School of Information, Production, and Systems, Waseda University, Japan. His professional interests include soft computing, tracking system, knowledge engineering, and management engineering.

Witold Pedrycz (Non-Member) received the M.Sc., Ph.D., and D.Sci. degrees from the Silesian University of Technology, Gliwice, Poland. He is currently a Professor and Canada Research Chair (CRC) in Computational Intelligence in the Department of Electrical, and Computer Engineering, University of Alberta, Edmonton, Canada. He is also with the Polish Academy of Sciences, Systems Research Institute, Warsaw, Poland. He is the author or co-author of numerous papers and 13 research monographs. His current research interests include computational intelligence, fuzzy modeling, knowledge discovery, and data mining, fuzzy control including fuzzy controllers, pattern recognition, knowledge-based neural networks, granular and relational computing, and software engineering. Prof. Pedrycz has been a member of numerous program committees of IEEE conferences in the area of fuzzy sets and neuro-computing. He is an Associate Editor of the *IEEE Transactions on Fuzzy Systems*. He is the Editor-in-Chief of *IEEE Transactions on Systems, Man, and Cybernetics, Part A, Information Sciences*, and the past president of the International Fuzzy Systems Association (IFSA), and North American Fuzzy Information Society (NAFIPS). He is a recipient of the Norbert Wiener Prize of the IEEE Systems, Man, and Cybernetics Society.