

A genetic programming based fuzzy regression approach to modeling manufacturing processes

Department of Industrial and Systems Engineering

The Hong Kong Polytechnic University

Hung Hom, Kowloon, Hong Kong, PRC

Corresponding author

Dr. C.K. Kwong

Tel : (852) 27666610

Fax : (852) 23625267

Email : mfckkong@polyu.edu.hk

Other authors:

Dr. K.Y. Chan

Email: richanky@inet.polyu.edu.hk

Mr. Y.C. Tsim

Email: mfyctsim@inet.polyu.edu.hk

Abstract

Fuzzy regression has demonstrated its ability to model manufacturing processes in which the processes have fuzziness and the number of experimental data sets for modeling them is limited. However, previous studies only yield fuzzy linear regression based process models in which variables or higher order terms are not addressed. In fact, it is widely recognized that behaviors of manufacturing processes do often carry interactions among variables or higher order terms. In this paper, a genetic programming based fuzzy regression GP-FR, is proposed for modeling manufacturing processes. The proposed method uses the general outcome of GP to construct models the structure of which is based on a tree representation, which could carry interaction and higher order terms. Then, a fuzzy linear regression algorithm is used to estimate the contributions and the fuzziness of each branch of the tree, so as to determine the fuzzy parameters of the genetic programming based fuzzy regression model. To evaluate the effectiveness of the proposed method for process modeling, it was applied to the modeling of solder paste dispensing process. Results were compared with those based on statistical regression and fuzzy linear regression. It was found that the proposed method can achieve better goodness-of-fitness than the other two methods. Also the prediction accuracy of the model developed based on GP-FR is better than those based on the other two methods.

Keywords: Fuzzy regression, genetic programming, process modeling, solder paste dispensing

1 Introduction

In today's competitive market, manufacturers need to control variability at each of the many processing steps in a manufacturing line, and all of the variables that control the desired output in a process need to be understood and optimized to maintain tight control. This can be achieved by developing appropriate physical models, or empirical models, to represent the manufacturing process. Analytical models are based on a physical understanding of the process, and rely on physical laws, typically a set of governing partial differential equations. They are attractive because they provide a fundamental understanding of the relationships between the input and output parameters. Various analytical models have been developed for manufacturing processes, such as fluid dispensing (Chen 2002, Li *et al.* 2001), injection molding (Chiang *et al.* 1991), and transfer molding (Han *et al.* 2000), but many manufacturing processes are too complex to model accurately and analytically.

Empirical modeling is a popular approach to the development of process models, based on using experimental data. The classical statistical regression method is a common empirical approach to the development of such process models (Seber 2003). It is well known that the statistical regression models are accurate only in the range in which they are developed. In conventional regression analysis, deviations between the observed values and the estimates are assumed to be due to random errors. Thus, statistical techniques are applied in order to make estimates and inferences in regression analysis. Statistical regression models can be applied only if the given data is distributed according to a statistical model, and the relationship between dependent and independent variables is crisp. However, in some manufacturing processes, it is difficult to find

probability distributions for dependent variables. The deviations in such cases are due to the indefinite structure of the system or to imprecise observations. The uncertainty in this type of process modeling becomes fuzzy.

Artificial neural networks (Simpson 1989) have been used to develop process models for various manufacturing processes, such as resistance spot welding (Li *et al.*, 2007) and transfer molding (Tong *et al.*, 2004). These networks have the capability to transform a non-linear mathematical model into a simplified black-box structure, and have the advantage of learning and generalization abilities, as well as nonlinearity. Previous research has already confirmed that neural networks are powerful tools for modeling nonlinear, complex, and noisy processes. A fuzzy logic modeling technique has been successfully applied to develop models for various manufacturing processes, such as the Flip-Chip bonding process (Kang *et al.*, 1993), vapor phase soldering (Xie *et al.* 1994), and the waterjet depainting process (Babets and Geskin, 2000). The basic elements of a fuzzy logic model are internal functions, membership functions, and outputs. The use of several internal functions accounts for the fuzziness of the model. However, the existing neural networks, and fuzzy logic modeling approaches normally require a large number of experimental data sets to develop models, which are usually not available in process designs. Also due to their lack of transparency, sensitivity studies of process parameters cannot be done easily.

In contrast, fuzzy linear regression has the distinct advantage that a manufacturing process, which has a high degree of fuzziness, can be modeled by using only a few or even incomplete experimental data sets (Tanaka *et al.* 1982, Takagi and Sugeno 1985, Tanaka and Watada 1988). An attempt was made by Schaiable and Lee (1997) to model

the vertical CVD process using the fuzzy linear regression method. Lai and Chang (1994) applied fuzzy linear regression to model the die casting process. Ip et al. (2003a) introduced the fuzzy linear regression to develop a process model for epoxy dispensing. Modeling of transfer molding using fuzzy linear regression was reported by Ip et al. (2003b). Kwong and Bai (2005) have performed process modeling and optimization using both fuzzy linear regression and fuzzy linear programming approaches. Three different approaches of fuzzy linear regression were summarized in Chang and Ayyub (2001). However, the existing fuzzy regression approaches cannot be used to develop models which contain interaction terms or even higher order terms. In fact, interaction among process parameters and nonlinear behavior of manufacturing process commonly exist. If interaction terms or higher order terms are integrated into the approach of fuzzy regression, more accurate models can be developed.

Genetic programming (GP) is an evolutionary method which can be used to generate models with interaction terms or higher order terms (Koza 1992, Koza 1994). Lakshminarayanan et al. (2000), Madar et al. (2005), Gray et al. (1996), McKay et al. (1997) and Willis et al. (1997) have demonstrated how GP can be used to generate models with interaction terms or higher order terms, and the least square algorithm is then used to perform the associated parameter estimation of the models. However, quite a number of manufacturing processes involve uncertainty, due to fuzziness. Therefore the above GP methods together with the least square algorithm may not yield the best modeling results, since the methods do not consider the fuzzy type of uncertainty.

To overcome the above deficiencies, in this paper, a new approach of fuzzy regression namely genetic programming based fuzzy regression GP-FR is proposed.

Fuzzy regression based process models are developed which can take into account interaction terms and higher order terms. The method uses the general outcomes of GP to construct the structures of models based on a tree representation where both the interaction and higher order terms can be considered. Then fuzzy regression is used to estimate the contributions of each branch of the tree, so as to determine the fuzzy parameters of each term of the model. Since interaction and higher order terms can be generated and represented in branches of the tree based on the GP-FR approach, fuzzy regression models with interaction and higher order terms can be produced. To evaluate the effectiveness of the proposed GP-FR approach to modeling manufacturing processes, it was applied to modeling the solder paste dispensing process. Results of the modeling were compared with those based on fuzzy linear regression and statistical regression.

2 Genetic Programming Based Fuzzy Regression

Fuzzy linear regression analysis, was first introduced by Tanaka and Watada (1988), in which two factors, namely the degree of fitness and the fuzziness of data sets, are considered. A fuzzy linear regression model is commonly presented as follows:

$$\tilde{y} = \tilde{f}_{LR}(x) = \tilde{A}_0 + \tilde{A}_1x_1 + \tilde{A}_2x_2 + \dots + \tilde{A}_jx_j + \dots + \tilde{A}_Nx_N = \tilde{A}x \quad (1)$$

where $x = [1, x_1, x_2, \dots, x_N]^T$ is a crisp vector of independent variables, and \tilde{y} is the estimated fuzzy output. $\tilde{A} = [\tilde{A}_0, \tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_N]$ is a vector of fuzzy parameters of the fuzzy linear regression model. \tilde{A}_j is presented in the form of symmetric triangular fuzzy numbers denoted by $\tilde{A}_j = (\alpha_j, c_j)$, $j = 0, 1, 2, \dots, N$, where its membership function is shown as below:

$$\mu_{\tilde{A}_j}(a_j) = \begin{cases} 1 - \frac{|a_j - \alpha_j|}{c_j}, & \alpha_j - c_j \leq a_j \leq \alpha_j + c_j, \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where α_j is the central value of the fuzzy number and c_j is the spread. Therefore the fuzzy linear regression model can be rewritten as shown below:

$$\tilde{y} = (\alpha_0, c_0) + (\alpha_1, c_1)x_1 + (\alpha_2, c_2)x_2 + \dots + (\alpha_N, c_N)x_N \quad (3)$$

However, interactions between variables and higher order terms are not included in the fuzzy linear regression defined in (1). In fact, interactions between variables and higher order terms often exist in physical systems.

The general form of the fuzzy regression models, which involves interactions between variables and higher order terms, can be represented as:

$$\tilde{y} = \tilde{f}_0 + \sum_{i=1}^N \tilde{f}_i(x_i) + \sum_{i=1}^N \sum_{j=1}^N \tilde{f}_{ij}(x_i, x_j) + \dots + \tilde{f}_{1,2,\dots,N}(x_1, x_2, \dots, x_N) \quad (4)$$

in which \tilde{f}_0 is a fuzzy bias term and $\tilde{f}_i(x_i)$, $\tilde{f}_{ij}(x_i, x_j)$, ... represent a univariate fuzzy component, and a bivariate fuzzy component, ... respectively (Friedman 1991). A higher order high-dimensional Kolmogorov-Gabor polynomial (Gabor et al 1961) is one of the forms of (4), which can be written as:

$$\tilde{y} = \tilde{f}_{NR}(x) = \tilde{A}_0 + \sum_{i_1=1}^N \tilde{A}_{i_1} x_{i_1} + \sum_{i_1=1}^N \sum_{i_2=1}^N \tilde{A}_{i_1 i_2} x_{i_1} x_{i_2} + \dots + \sum_{i_1=1}^N \dots \sum_{i_d=1}^N \tilde{A}_{i_1 \dots i_d} \prod_{j=1}^d x_j \quad (5)$$

where $\tilde{A}_0 = (\alpha_0, c_0)$, $\tilde{A}_1 = (\alpha_1, c_1)$, $\tilde{A}_2 = (\alpha_2, c_2)$, ...

$\tilde{A}_N = (\alpha_N, c_N)$, $\tilde{A}_{11} = (\alpha_{11}, c_{11})$, $\tilde{A}_{12} = (\alpha_{12}, c_{12})$, ... $\tilde{A}_{NN} = (\alpha_{NN}, c_{NN})$, ...

$\tilde{A}_{N\dots N} = (\alpha_{N\dots N}, c_{N\dots N})$.

The fuzzy regression model (5) can be rewritten as:

$$\tilde{y} = \tilde{A}'_0 x'_0 + \tilde{A}'_1 x'_1 + \tilde{A}'_2 x'_2 \dots \tilde{A}'_{N_{NR}} x'_{N_{NR}} \quad (6)$$

$$\text{or } \tilde{y} = (c'_0, \alpha'_0) x'_0 + (c'_1, \alpha'_1) x'_1 + (c'_2, \alpha'_2) x'_2 + \dots (c'_{N_{NR}}, \alpha'_{N_{NR}}) x'_{N_{NR}} \quad (7)$$

where $1+N_{NR}$ is the number of terms of (5), (6) and (7); $\tilde{A}'_0 = \tilde{A}_0$, $\tilde{A}'_1 = \tilde{A}_1$, $\tilde{A}'_2 = \tilde{A}_2, \dots$

$\tilde{A}'_{N_{NR}} = \tilde{A}_{N \dots N}$; $x'_0 = 1$, $x'_1 = x_1$, $x'_2 = x_2$, \dots $x'_{N_{NR}} = x_1 \cdot x_2 \cdot \dots \cdot x_d$; and $\tilde{A}'_0 = (c_0, \alpha_0)$,

$\tilde{A}'_1 = (c'_1, \alpha'_1)$, \dots $\tilde{A}'_{N_{NR}} = (c'_{N_{NR}}, \alpha'_{N_{NR}})$. \tilde{A}'_i and x'_i are called the fuzzy parameters and

the transformed variables respectively, where $i=0,1,2,\dots,N_{NR}$.

The vectors of the fuzzy parameters are defined as:

$$\tilde{A}' = (\tilde{A}'_0, \tilde{A}'_1, \dots, \tilde{A}'_{N_{NR}}) = ((c'_0, \alpha'_0), (c'_1, \alpha'_1), \dots, (c'_{N_{NR}}, \alpha'_{N_{NR}})), \quad (8)$$

$$c' = (c'_0, c'_1, \dots, c'_{N_{NR}}), \quad (9)$$

$$\text{and } \alpha' = (\alpha'_0, \alpha'_1, \dots, \alpha'_{N_{NR}}). \quad (10)$$

The vector of the transformed variables is defined as:

$$x' = (x'_0, x'_1, x'_2, \dots, x'_{N_{NR}}). \quad (11)$$

Using the vectors of the fuzzy parameter and the vector of transformed variables,

(6) can be rewritten as:

$$\tilde{y} = \tilde{A}' \cdot x'^T \quad (12)$$

Figure 1 shows a fuzzy regression model which contains all samples within the nonlinear polynomial intervals.

Since some terms in (6) may be redundant, prudent selection of significant terms or orders is advisable if a more parsimonious and adequate model is desired. In this paper,

the genetic programming based fuzzy regression (GP-FR), is proposed and descriptions of it are given.

The pseudocode of the genetic programming based fuzzy regression (GP-FR) is shown below.

```

t=0

Initialize  $\Omega(t)=[\theta_1(t), \theta_2(t), \dots \theta_{POP}(t)]$ 

Assign fuzzy parameters to all  $\theta_i(t)$ 

//  $\Omega(t)$  is the population of the  $t$ -th generation.

//  $\theta_i(t)$  is the  $i$ -th individual of  $\Omega(t)$ .

Evaluate all  $\theta_i(t)$  according to a fitness function

while (Terminational condition not fulfilled) do {

    Parent Selection  $\Omega(t+1)$ 

    Crossover  $\Omega(t+1)$ 

    Mutation  $\Omega(t+1)$ 

    Determine fuzzy parameters in all  $\theta_i(t+1)$  by using

        Tanaka's fuzzy regression

    Evaluate all  $\theta_i(t+1)$ 

     $\Omega(t) = \Omega(t+1)$ 

    t=t+1

}

```

The GP-FR first starts with creating a random initial population $\Omega(t)$ with POP individuals $\theta_i(t)$, while $t=0$. Each individual $\theta_i(t)$ is in a form of tree structure, that can be used to represented the structure of the fuzzy regression model as defined in (5). Then the

fuzzy parameters are assigned to each individual $\theta_i(t)$ by applying Tanaka and Watada's (1988) fuzzy regression. All individuals are evaluated according to a defined fitness function which is aimed at evaluating the goodness-of-fitness of the fuzzy regression model. The parent selection process uses the goodness-of-fitness of each individual to determine the selection of potential individuals for performing crossover or mutation. Finally, the new individuals with the determined fuzzy parameters are evaluated using the fitness function in order to create a new population $\Omega(t+1)$. The process continues until the pre-defined termination condition is fulfilled. Major aspects of applying the GP-FR on modeling the functional relationships are described below:

2.1 Model Representation

In the GP-FR, one of the most popular methods to represent structures is by using hierarchical trees which are composed of functions F and terminals T (Koza 1992). The fuzzy regression model (5) contains only the three arithmetic operations, +, - and *, thus F is represented as $F = \{+, -, *\}$. The set of terminals $T = \{\mathbf{x}, \tilde{\mathbf{p}}\}$ contains the variable set $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$ of the fuzzy regression model and the fuzzy parameter set $\tilde{\mathbf{p}} = \{\tilde{p}_0, \tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_{N_{NS}}\}$ of the fuzzy regression model, where n is the number of variables and N_{NS} is the number of terms of the fuzzy regression model. A potential solution is depicted as a labeled tree with ordered branches. In the tree, operations from the function set F are used as internal nodes, and arguments from the terminal set T are used as terminal nodes. For example, Figure 2 shows an example of a hierarchical tree that expresses the following formulation:

$$(x_1 * x_1) - (x_2 * x_2) + (x_1 * x_2 * x_4)$$

which is equivalent to:

$$x_1^2 - x_2^2 + x_1 x_2 x_4$$

The fuzzy parameters set $\tilde{\mathbf{p}} = \{\tilde{p}_0, \tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_{N_{NS}}\}$ can be obtained after determining the structure of the function from the tree. In Figure 2, the number of fuzzy parameters of the fuzzy regression model is 4. Therefore, the completed fuzzy regression model can be represented as follows:

$$\tilde{p}_0 + \tilde{p}_1 \cdot x_1^2 - \tilde{p}_2 \cdot x_2^2 + \tilde{p}_3 \cdot x_1 x_2 x_4,$$

It can also be represented by:

$$\tilde{p}_0 + \tilde{p}_1 \cdot x'_1 - \tilde{p}_2 \cdot x'_2 + \tilde{p}_3 \cdot x'_3,$$

where $x'_1 = x_1^2$, $x'_2 = x_2^2$ and $x'_3 = x_1 \cdot x_2 \cdot x_4$.

In this research, the fuzzy parameters, $\tilde{p}_0, \tilde{p}_1, \dots, \tilde{p}_{N_{NS}}$, are determined according to Tanaka and Watada's (1988) fuzzy regression. The following linear programming problem is formulated for the fuzzy regression problem with reference to Tanaka *et al.* (1982). By solving the linear programming problem, the fuzzy parameters, $\tilde{p}_0, \tilde{p}_1, \dots, \tilde{p}_{N_{NS}}$, can be determined.

$$\text{Minimize } J = \sum_{j=0}^{N_{NR}} \left(c'_j \sum_{i=1}^M |x'_j(i)| \right) \quad (13)$$

where M is the number of data sets, and $x'_j(i)$ is the j -th transformed variable of the fuzzy polynomial model of the i -th data set, subject to:

$$\sum_{j=0}^{N_{NR}} \alpha'_j x'_j(i) + (1-h) \sum_{j=0}^{N_{NR}} c'_j |x'_j(i)| \geq y(i) \quad (14)$$

$$\sum_{j=0}^{N_{NR}} \alpha'_j x'_j(i) + (1-h) \sum_{j=0}^{N_{NR}} c'_j |x'_j(i)| \leq y(i) \quad (15)$$

$$c'_j \geq 0, \quad \alpha'_j \in R, \quad x'_0(i) = 1 \text{ for all } i, \quad (16)$$

$$0 \leq h \leq 1, \quad i = 1, 2, \dots, M, \quad j = 0, 1, 2, \dots, N_{NR}. \quad (17)$$

J in (13) is the total fuzziness of the regression model. The value of h in (14) and (15) is between 0 and 1. h refers to the degree to which the fuzzy linear model fits the given data sets, and is subjectively chosen by decision makers. Constraints (14) and (15) impose the restriction that the observation of the i -th data set $y(i)$ has at least h degree of belonging to $\tilde{y}(i)$ as $\mu_{\tilde{y}(i)}(y(i)) \geq h$ ($i = 1, 2, \dots, M$). Therefore, the objective of solving the linear programming problem (13-17) is to determine the fuzzy nonlinear parameters $\tilde{A}'_j = (c'_j, \alpha'_j)$ such that the total vagueness J is minimized subject to $\mu_{\tilde{y}(i)}(y(i)) \geq h$ ($i = 1, 2, \dots, M$).

2.2 Fitness function

GP-FR evaluates the goodness-of-fitness of each individual by using a fitness function, which is based on the mean absolute error (MAE), and can reflect the differences between the predicted values of the model and the actual values of the data sets. The MAE of the j -th individual can be calculated based on (18).

$$MAE_j = 100\% \times \sqrt{\frac{1}{M} \sum_{k=1}^M \left| \frac{y(k) - F_j(x(k))}{y(k)} \right|}, \quad (18)$$

where F_j is the fuzzy regression model represented by the j -th individual, $(y(k), x(k)) = (y(k), (x_1(k), x_2(k), \dots, x_N(k)))$ is the k -th training data set, N is the number of

variables of the training data set, and M is the number of training data sets used for developing the fuzzy regression model.

(18) is commonly known as an indicator of training errors in a model. It reflects how well the model can fit the training data sets. However, a model may contain a lot of unnecessary and complex terms. A complex over-parameterized model with a large number of parametrical terms reduces the transparency and ease of interpretation of the model. To avoid the GP-FR from generating models which are too complex, a fitness function is designed to balance the tradeoff between the reduction of complexity and model accuracy. In this research, penalty terms are introduced into the fitness function of the GP-FR (McKay et al. 1997), and the fitness of the j -th individual is denoted as:

$$fitness_j = \frac{1 - MAE_j}{(1 + \exp(c_1(L_j - c_2)))} \quad (19)$$

where $fitness_j$ is the fitness value, L_j is the number of nodes of the j -th individual, and c_1 and c_2 are both penalty terms.

2.3 Crossover and mutation

Like other evolutionary algorithms, the two main evolutionary operators are crossover and mutation. The crossover operation produces a pair of offspring that inherit characteristics from both parents by selecting a random node in each of the hierarchical tree structures of the parents (as shown in Figure 3a) and exchanging the associated sub-expressions of the hierarchical tree structures (as shown in Figure 3b). Because of the dynamic representation used in GP-FR, typically the parents are different in size, shape and content. The process of mapping the genotype onto the phenotype does not correspond to a one-to-one relationship. Therefore the resulting offspring can be

expressed by more than one different tree structure and some diversification of the population is allowed.

Mutation is performed by randomly selecting a node that is an internal or terminal node, and by replacing the associated sub-expression with a randomly generated sub-expression. For example, Figure 4 shows that the node of a minus is selected and is mutated to a sum.

2.4 Selection and convergence

After the operations of crossover and mutation, individuals from the current population with relatively better fitness are selected to serve as parents for the next generation. The approach of a roulette-wheel, which is one of the most common selection methods used for selecting individuals to perform reproduction operations in evolutionary algorithms (Goldberg 1989), is used for the selection of individuals. Regarding the j^{th} individual, its fitness is assigned a value, $fitness_j$, and the selection probability value, $prob_j$, is defined as:

$$prob_j = \frac{fitness_j}{\sum_{j=1}^{POP} fitness_j} \quad (20)$$

where POP is the population size of the GP-FR. Equation (20) shows that the individual with a larger fitness value has a higher probability of being selected.

After the selection, the population evolves and improves iteratively until a stopping condition is met. In genetic programming, there are several stopping conditions. In this research, the stopping criterion is met when the number of generations is equal to

a pre-defined number of generations. Otherwise, the GP-FR goes on to the next evolutionary iteration.

3 Validation of genetic programming based fuzzy regression approach to modeling manufacturing processes

To validate the effectiveness of the GP-FR approach to modeling manufacturing processes, it is employed to model a solder paste dispensing process used in electronic manufacturing. The modeling results are compared with those based on the fuzzy linear regression and statistical regression.

3.1 Solder Paste Dispensing Process

In electronic manufacturing, solder dispensing machines are so controlled that they automatically place a certain amount of solder paste on a printed circuit board. The solder paste spots must be positioned correctly and must contain the specified amount of solder paste such that when a surface mounted IC component is placed onto the solder paste, all of the IC's leads are aligned correctly with each solder spot.

The motion of the dispensing machine has three axes. The x -axis and y -axis are used to place the solder paste dispenser over the desired location, and the z -axis is used to position the tip of the solder paste dispenser at the desired height above the board. The dispensing machine works as follows:

A series of x , y and z coordinates describing the locations of where the solder paste is to be deposited is entered into the system's computer program. For each of these locations two programmable process parameters are specified: 1) the amount of solder

paste, which is governed by the amount of time that the pump is engaged, and 2) the dwell time, which is the length of time the dispensing system remains over the location after the pump has been disengaged. A schematic diagram of a solder paste dispensing system is shown in Figure 5. The continuously running motor is connected to a clutch. The output of the clutch drives a screw pump. The amount of time, that the clutch is engaged, determines the amount of solder paste deposited, and is called the shot size. The solder paste exits through the interchangeable needle. The different solder pastes come prepackaged in tubular containers, which are inserted in the receptacle adjacent to the motor and clutch structure.

In the process of solder paste dispensing, the key quality characteristic is the diameter of the circular solder pads. The four significant operating parameters (factors) for the solder paste dispensing process to be studied are pressure, needle inner diameter, shot size and dwell time which are represented by x_1 , x_2 , x_3 and x_4 respectively. In the experimental plan, each factor has two levels. Table 1 shows the setting of each level of the factors. Table 2 shows the experimental results.

4.2 Model Development

For model development whether using GP-FR, statistic regression or fuzzy linear regression, the four operating parameters x_1 , x_2 , x_3 and x_4 need to be normalized to [0,1], and their resulting values are shown in Table 2. Evaluation of the effectiveness of the models can be carried out by investigating the mean of training errors and the variance of training errors, which are defined as Re and Rv respectively as follows:

$$\text{Re} = \frac{1}{M} \sum_{i=1}^M \frac{|y(i) - P(x_1(i), x_2(i), x_3(i), x_4(i))|}{y(i)} \times 100\% \quad (21)$$

and

$$Rv = \frac{1}{M-1} \sum_{i=1}^M \left[\frac{|y(i) - P(x_1(i), x_2(i), x_3(i), x_4(i))|}{y(i)} \times 100\% - Re \right] \quad (22)$$

In (21) and (22), M is the number of experimental data sets; $y(i)$ is the i -th measured value of the solder spot diameter; $x_1(i)$, $x_2(i)$, $x_3(i)$ and $x_4(i)$ are parameter values of the i -th experimental data set; and $P(x_1(i), x_2(i), x_3(i), x_4(i))$ is the predicted value for the i -th experiment.

Using the 16 experimental data sets and their results shown in Table 2, the proposed GP-FR was implemented using Matlab to relate solder paste diameter and the operating parameters x_1 , x_2 , x_3 and x_4 . The GP parameters are set as shown in Table 3 with reference to (Madar et al. 2005). Since GP-FR is a stochastic method, different results will be obtained from different runs. To evaluate its overall performance, 30 runs on the GP-FR were carried out, and the mean of the 30 runs was calculated. The mean fitness along generations of the 30 runs is shown in a convergence curve in Figure 6.

The model with the smallest mean training error among the 30 runs is shown below as:

$$y = (1.1887, 0.6985) + (0.1473, 0.0058) \cdot x_2 + (0.8763, 0.0071) \cdot x_3 + (0.1397, 0.0005) \cdot x_4 - (0.1854, 0.0008) \cdot x_1 \cdot x_2 - (0.2032, 0.0003) \cdot x_2 \cdot x_4 \quad (23)$$

Re and Rv were found as 3.2580% and 0.1285 respectively.

Using the same experimental data sets as shown in Table 2, the following statistical regression model was determined.

$$y = 1.2929 - 0.0914 \cdot x_1 - 0.0483 \cdot x_2 + 0.8458 \cdot x_3 + 0.0483 \cdot x_4 \quad (24)$$

The R^2 value of the model (24) is 96.7%. Re is 4.4575% and Rv is 0.1551. Based on the same set of data, the following fuzzy linear regression model was also determined.

$$y = (1.2649, 1.4605) + (-0.0711, 0.0152) \cdot x_1 + (-0.0457, 0.0152) \cdot x_2 + (0.8433, 0.0051) \cdot x_3 + (0.0584, 0.0025) \cdot x_4 \quad (25)$$

Re and Rv were found as 4.9874% and 0.1418 respectively.

Table 4 summarizes all the Re and Rv of the three methods, the proposed GP-FR, statistical regression and fuzzy regression. From Table 4, it can be seen that both the Re and Rv of the proposed GP-FR are smaller than those of the statistical regression and fuzzy linear regression. This indicates that the proposed GP-FR can fit the data sets with the smallest mean of errors and the smallest variance of errors.

To further validate the modeling performance of the GP-FR, four data sets were randomly selected from the 16 data sets, as shown in Table 2, as testing data sets and the remaining 12 data sets were used to develop a GP-FR model, a statistical regression model and a fuzzy linear regression model. Their prediction errors were calculated. The validations were repeated 12 times. We ran the GP-FR 30 times in each validation test and the mean of the 30 runs was calculated. Table 5 summarizes the prediction errors of the three methods. From the table, it can be seen that GP-FR yields the smallest mean of prediction errors and variance of prediction errors. Results of the 12 validation tests are shown in Figure 7 from which it can be seen that the prediction errors of GP-FR, for all the tests, are the smallest.

4 Conclusion

The existing fuzzy regression approaches are not able to model with interaction terms and higher order terms. In this paper, a genetic programming based fuzzy regression (GP-FR)

approach is proposed for modeling manufacturing processes, by which, models can be developed with proper interaction terms and higher order terms. The proposed GP-FR uses the general outcomes of GP to construct the structure of a model based on a tree representation. Then, a fuzzy regression algorithm is used to estimate the contributions and the fuzziness of each of the branches of the tree so as to determine the fuzzy parameters of each term of the model.

To validate the proposed GP-FR approach to modeling manufacturing processes, it has been applied to the modeling of the solder paste dispensing process, and has been compared with the other commonly used explicit modeling methods, statistical regression and fuzzy linear regression. The result shows that the smallest number of training errors can be achieved by GP-FR. This indicates that GP-FR is more capable to fit the data sets than the other two tested methods. Also, a comparison of the validation result shows that the smallest number of prediction errors and errors in variance can be achieved by GP-FR. The achievement of better results can be explained by the introduction of interaction terms in GP-FR, but the two most commonly used methods ignore them.

Acknowledgements

The work described in this paper was supported by a grant from Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University.

References

[Babets and Geskin 2000] Babets K. and Geskin E.S., Application of fuzzy logic for modeling of waterjet depainting, *Machining Science and Technology*, vol. 4, no. 1, pp. 81-100, 2000.

[Bates and Watts 1988] Bates D.M. and Watts D.G., *Nonlinear Regression Analysis and Its Applications*, New York: Wiley, 1998.

[Chang and Ayyub 2001] Chang Y.H.O. and Ayyub B.M., Fuzzy regression methods – a comparative assessment, *Fuzzy Sets and Systems*, vol. 119, pp. 187-203, 2001.

[Chen 2002] Chen D.X., *Modeling and off-line control of fluid dispensing for electronics packaging*. PhD thesis, University of Saskatchewan, 2002.

[Chen et al 2004] Chen Y., Tang J., Fung R.Y.K. and Ren Z., Fuzzy regression based mathematical programming for QFD, *International Journal of Production Research*, vol. 42, no. 5, pp. 3583-3604, 2004.

[Chiang et al 1991] Chiang H. H., Hieber C. A. and Wang K.K., A unified simulation of the filling and postfilling stages in injection molding, Part 1: formulation. *Polymer Engineering and Science*, vol. 31, pp. 116-124, 1991.

[Friedman 1991] Friedman J.H., Multivariate adaptive regression splines, *The Annals of Statistics*, vol. 19, no. 1, pp. 1-141, 1991.

[Gabor 1961] Gabor D., Wildes W. and Woodcock R., A Universal non-linear filter, predictor and simulator which optimizes itself by a learning process, *Proceedings of IEE*, vol. 108B, pp. 422-438, 1961.

[Goldberg 1989] Goldberg D.E, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, 1989.

[Gray et al. 1996] Gray G.J., Murray-Smith D.J., Li Y., and Sharman K.C., *Nonlinear model structure identification using genetic programming and a block diagram oriented simulated tool*, *Electronic Letters*, vol. 32, pp. 1422-1424, 1996.

[Han et al 2000] Han R., Shi L. and Gupta M., Three-dimensional simulation of microchip encapsulation process, *Polymer Engineering and Science*, vol. 40, no. 3, pp. 776-785, 2000.

[Holland 1975] Holland J.H., *Adaptation in Natural and Artificial Systems*, Michigan Press, 1975.

[Ip et al 2003a] Ip C.K.W, Kwong C.K., Bai H. and Tsim Y.C., The process modeling of epoxy dispensing for microchip encapsulation using fuzzy linear regression with fuzzy intervals, *International Journal of Advanced Manufacturing Technology*, vol. 22, pp. 417-423, 2003.

[Ip et al 2003b] Ip K.W., Kwong C.K. and Wong Y.W., Fuzzy regression approach to modeling transfer moulding for microchip encapsulation. *Journal of Materials Processing Technology*, vol. 140, pp. 147-151, 2003.

[Kang et al 1993] Kang S. Y., Xie H. and Lee Y. C., Physical and fuzzy logic modeling of a flip-chip thermo-compression bonding process, *Journal of Electronic Packaging*, vol. 115, pp. 63-70, 1993.

[Khalil 2002] Khalil H.K., *Nonlinear Systems*, Prentice Hall, 2002.

[Koza 1992] Koza J., *Genetic Programming: On the Programming of Computers by Means of Natural Evolution*, MIT Press: Cambridge, 1992.

[Koza 1994] Koza J., *Genetic Programming II: automatic discovery of reusable programs*, MIT Press, 1994.

[Kwong and Bai 2005] Kwong C.K. and Bai H., Fuzzy Regression Approach to Process Modeling and Optimization of Epoxy Dispensing, *International Journal of Production Research*, vol. 43, no. 12, pp. 2359-2375, 2005.

[Lai and Chang 1994] Lai Y.J. and Chang S.I., A fuzzy approach for Multiresponse optimization: an off-line quality engineering problem, *Fuzzy Sets and Systems*, vol. 63, pp. 117-129, 1994.

[Lakshminarayanan et al 2000] Lakshminarayanan S., Fujii H., Grosman B., Dassau E., and Lewin D.R., New product design via analysis of historical databases, *Computers and Chemical Engineering*, vol. 24, pp. 671-676, 2000.

[Li et al 2001] Li H.X., Tso S.K. and Deng H., A concept approach to integrate design and control for the epoxy dispensing process, *International Journal of Advanced Manufacturing Technology*, vol. 17, pp. 677-682, 2001.

[Li et al 2007] Li H.L., Chou T. and Chou C.P., Optimization of resistance spot welding process using Taguchi method and a neural network, *Experimental Techniques*, vol. 31, no. 5, pp. 30-36, 2007.

[Madar et al. 2005] Madar J., Abonyi J. and Szeifert F., Genetic programming for the identification of nonlinear input – output models, *Industrial and Engineering Chemistry Research*, vol. 44, pp. 3178 – 3186, 2005.

[McKay et al. 1997] McKay B., Willis M.J., and Barton G.W., Steady-state modeling of chemical processes using genetic programming, *Computers and Chemical Engineering*, vol. 21, no. 9, pp. 981-996, 1997.

[Schaible and Lee 1997] Schaible B. and Lee Y.C., Fuzzy logic based regression models for electronics manufacturing applications, *Advances in Electronic Packaging*, vol. 1, pp. 147-155, 1997.

[Seber 2003] Seber G.A.F., *Linear regression analysis*, Wiley, 2003.

[Simpson 1989] Simpson P.K., *Artificial neural systems*, New York: Pergamon Press, 1989.

[Takagi and Sugeno 1985] Takagi T. and Sugeno M., Fuzzy identification of systems and its application to modeling and control, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 15, no. 1, pp. 116-132, 1985.

[Tanaka et al. 1982] Tanaka H., Uejima S. and Asai K., Linear regression analysis with fuzzy model, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 12, pp. 903-907, 1982.

[Tanaka and Watada 1988] Tanaka H. and Watada J., Possibilistic linear systems and their application to the linear regression model, *Fuzzy Sets and Systems*, vol. 272, pp. 275-289, 1988.

[Tong et al 2004] Tong K.W., Kwong C.K., Yu K.M., Intelligent process design system for the transfer moulding of electronic packages, *International Journal of Production Research*, vol. 42, no. 10, pp. 1911-1931, 2004.

[Willis et al. 1997] Willis M.J., Hiden H., Hinchliffe M., McKay B., and Barton G.W., Systems modeling using genetic programming, *Computers and Chemical Engineering*, vol. 21, pp. 1161-1166, 1997.

[Xie et al 1994] Xie H. and Lee Y. C., Process optimization using a fuzzy logic response surface method, *IEEE Transactions on Components, Packaging and Manufacturing Technology – Part A*, Vol. 17, No. 2, pp. 202-210, 1994.

[Zimmermann 1985] Zimmermann H.J., *Fuzzy sets theory and its applications*, Kluwer, Boston, 1985.

Table 1 Factors and their levels

Factors	Factor levels	
	Level 1 (0)	Level 2 (1)
x_1 , Pressure (bar)	0.276	0.827
x_2 , Needle inner diameter (mm)	0.041	0.584
x_3 , Short size (ms)	150	500
x_4 , Dwell time (ms)	0	500

Table 2 2^4 factorial experimental plan and results

	Normalization				Solder spot diameters (mm)
	$x_1(j)$	$x_2(j)$	$x_3(j)$	$x_4(j)$	$y(j)$
j -th experimental data set					
1	0	0	0	0	1.1176
2	1	0	0	0	1.1176
3	0	1	0	0	1.2954
4	1	1	0	0	1.2192
5	0	0	1	0	2.1082
6	1	0	1	0	2.1336
7	0	1	1	0	2.2098
8	1	1	1	0	1.9558
9	0	0	0	1	1.3970
10	1	0	0	1	1.3716
11	0	1	0	1	1.2700
12	1	1	0	1	1.1430
13	0	0	1	1	2.1590
14	1	0	1	1	2.1336
15	0	1	1	1	2.1590
16	1	1	1	1	1.9304

Table 3 GP parameters implemented in the GP-FR

Population size	50
Maximum number of evaluated individuals	5000
Generation gap	0.9
Probability of crossover	0.5
Probability of mutation	0.5
Probability of changing terminal via non-terminal	0.25
Penalty factors in the fitness function (20)	$c_1=0.5$ and $c_2=30$

Table 4 Comparisons of the modeling results

	GP-FR	Statistical regression	Fuzzy regression
Mean of training errors Re (%)	3.2580	5.0291	4.9874
Variance of training errors Rv	0.1285	0.1551	0.1418

Table 5 Prediction errors of the three methods

Validations	Testing data set	GP-FR	Statistical regression	Fuzzy linear regression
1	1 13 15 16	7.7065	8.2668	9.5705
2	1 5 6 12	8.7606	8.9494	9.7422
3	6 12 15 16	8.9340	9.6433	9.6136
4	3 5 10 13	6.4408	6.8848	6.3889
5	4 6 13 16	6.4891	7.0259	7.2404
6	2 8 15 16	6.5076	7.8022	7.9411
7	1 10 11 14	6.7059	12.3621	9.8317
8	1 2 4 15	6.6578	15.0754	14.6359
9	8 12 14 16	9.2822	11.9544	12.0100
10	1 7 11 14	6.4753	9.0530	8.0198
11	4 7 13 16	5.5274	7.4860	6.4493
12	4 6 15 16	5.8662	6.4916	6.6645
Mean of prediction errors		7.1128	9.2496	9.0090
Variance of prediction errors		1.5586	6.0657	6.8675

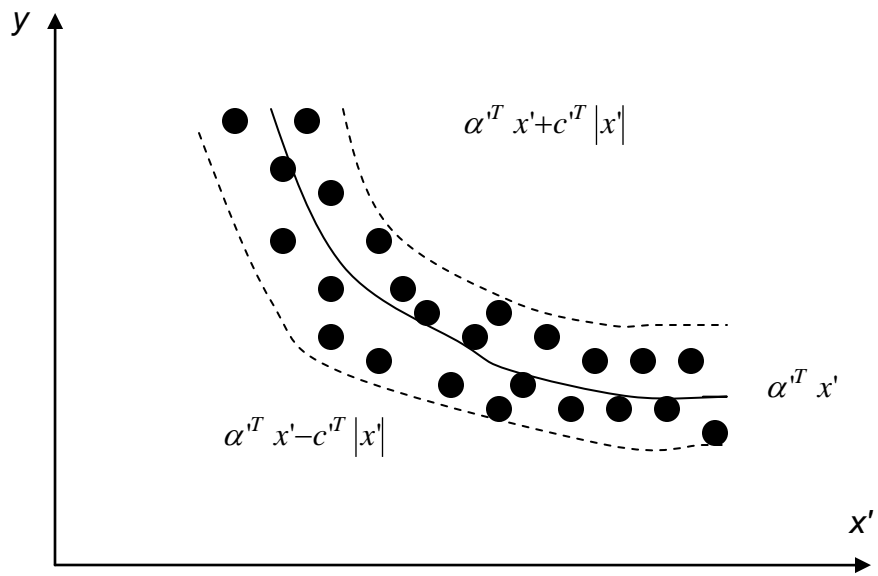


Figure 1 The fuzzy regression model

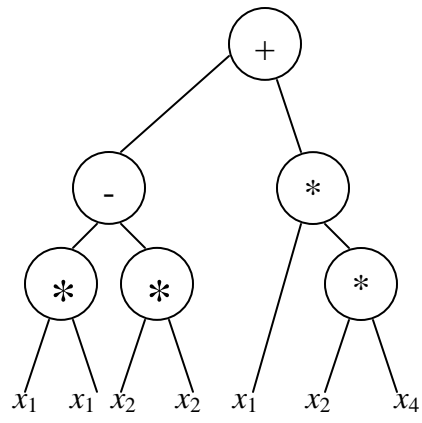


Figure 2 An example of a hierarchical tree

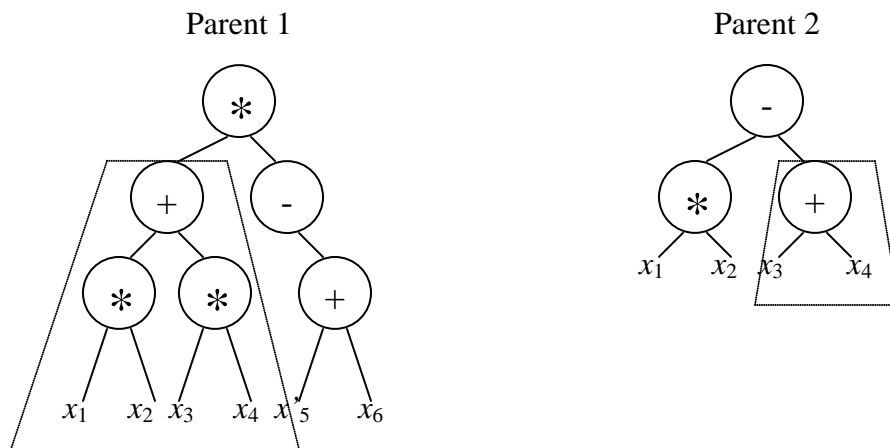


Figure 3 (a) Random selection of a sub-expression before crossover

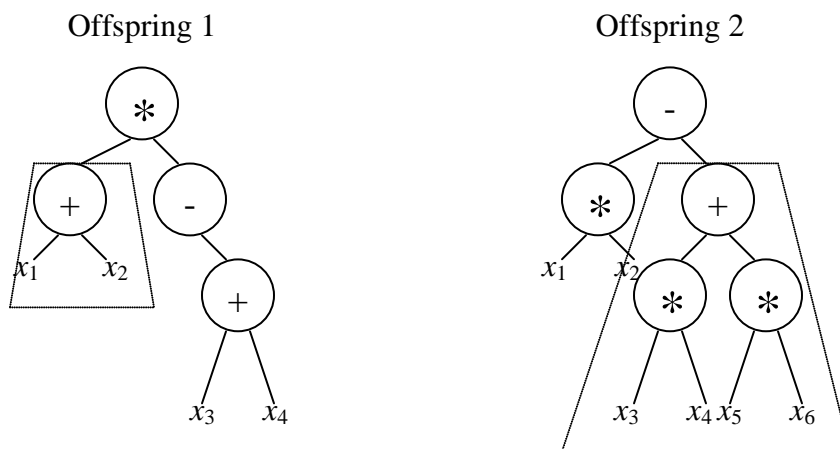


Figure 3 (b) Offspring produced by the crossover

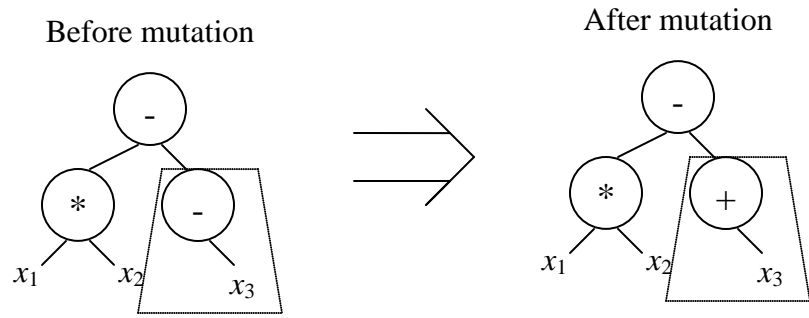


Figure 4 Offspring produced by the mutation

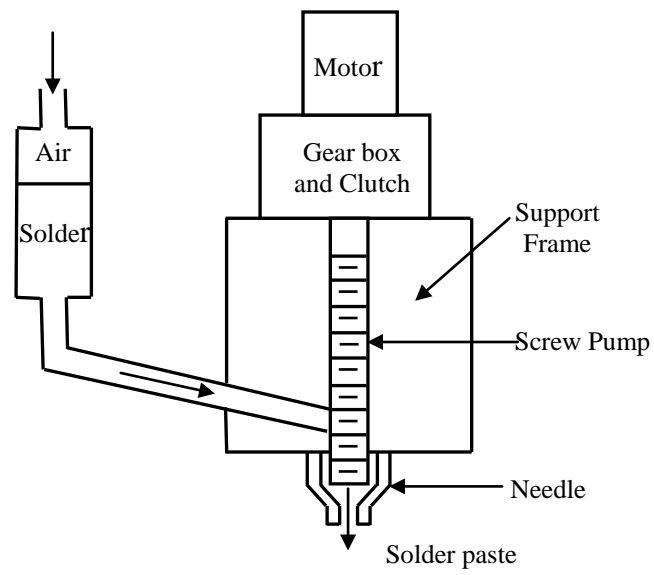


Figure 5 Solder paste dispensing system

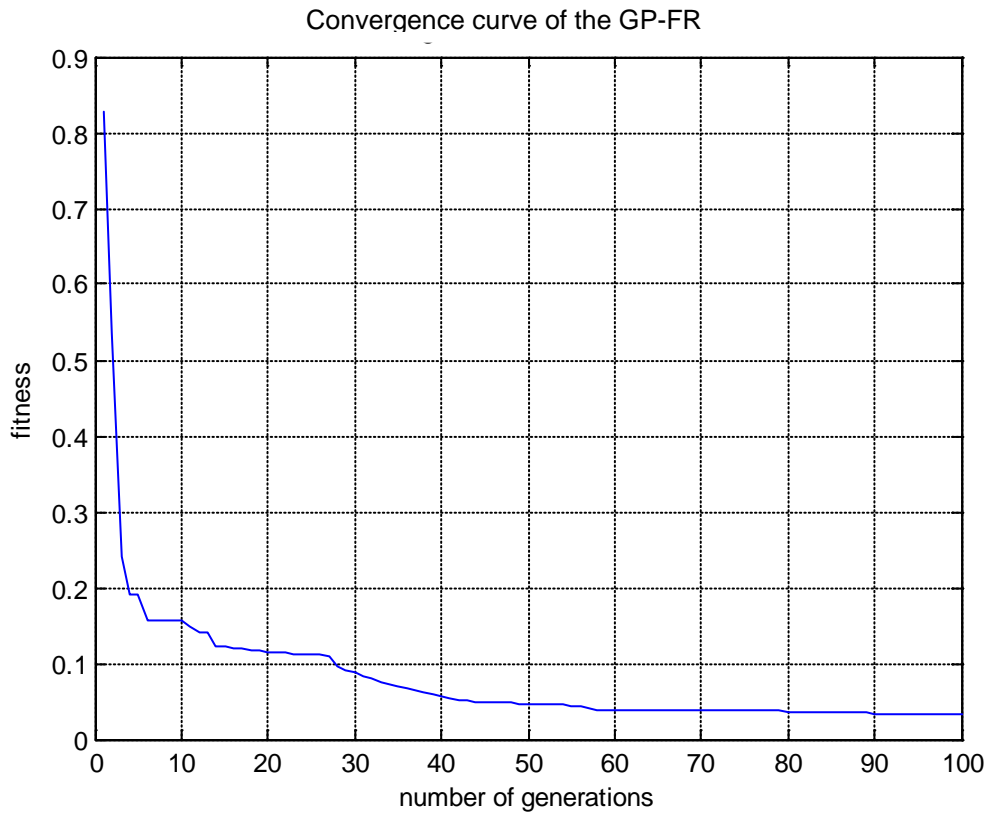


Figure 6 Convergence curve of the GP-FR

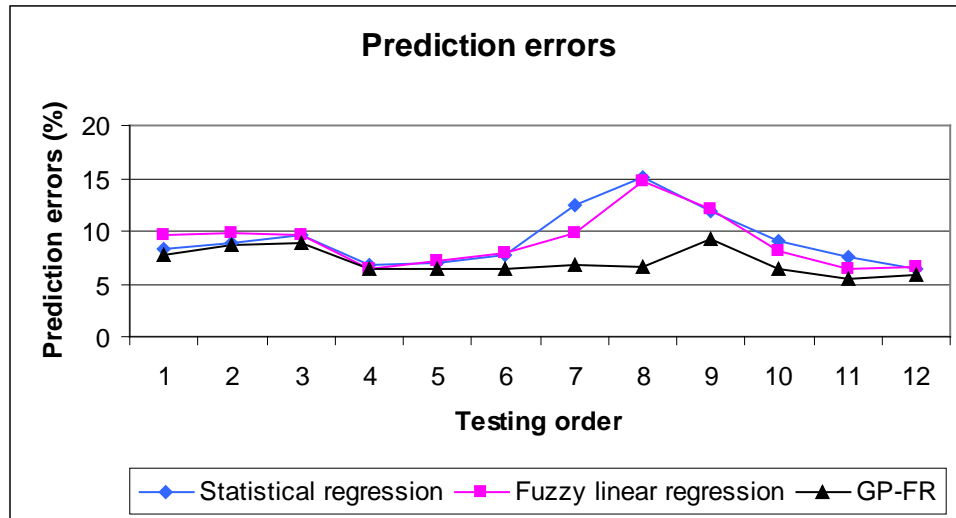


Figure 7 Prediction errors for each testing order using statistical regression, fuzzy linear regression and GP-FR