Modeling Manufacturing Processes Using a Genetic programming-based Fuzzy Regression with Detection of Outliers

K.Y. Chan[1], C.K. Kwong[2] and T.C. Fogarty[3]

[1]Digital Escosystems and Business Intelligence Institute, Curtin University of Technology, Perth, Australia; Email: Kit.Chan@curtin.edu.au

[2]Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong; Email: C.K.Kwong@inet.polyu.edu.hk

Faculty of Business, Computing and Information Management, London South Bank University, London; Email: fogarttc@lsbu.ac.uk

Abstract

Fuzzy regression (FR) been demonstrated as a promising technique for modeling manufacturing processes where availability of data is limited. FR can only yield linear type FR models which have a higher degree of fuzziness, but FR ignores higher order or interaction terms and the influence of outliers, all of which usually exist in the manufacturing process data. Genetic programming (GP), on the other hand, can be used to generate models with higher order and interaction terms but it cannot address the fuzziness of the manufacturing process data. In this paper, genetic programming-based fuzzy regression (GP-FR), which combines the advantages of the two approaches to overcome the deficiencies of the commonly used existing modeling methods, is proposed in order to model manufacturing processes. GP-FR uses GP to generate model structures based on tree representation which can represent interaction and higher order terms of models, and it uses an FR generator based on fuzzy regression to determine outliers in experimental data sets. It determines the contribution and fuzziness of each term in the model by using experimental data excluding the outliers. To evaluate the effectiveness of GP-FR in modeling manufacturing processes, it was used to model a non-linear system and an epoxy dispensing process. The results were compared with those based on two commonly used FR methods, Tanka's FR and Peters' FR. The prediction accuracy of the models developed based on GP-FR was shown to be better than that of models based on the other two FR methods.

Keywords: genetic programming, fuzzy regression, outlier detection, epoxy dispensing process

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 variables controlling 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 to represent the manufacturing process. Physical models [5, 9, 12 and 29] are based on a physical understanding of the process, and they typically consist of 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. However, physical models are usually too complex to be generated accurately for many manufacturing processes.

Statistical regression is a common approach to develop empirical process models [39], but the resulting models are accurate only within the ranges of data from which they are developed. 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 many manufacturing processes, it is difficult to find probability distributions for dependent variables. Artificial neural networks [4, 15, 20, 30, 40 and 45] and fuzzy logic modeling techniques [1, 10, 18, 19, 33 and 47] have been used to develop process models in various manufacturing processes. These approaches normally require a large amount of experimental data to develop models, which are sometimes not available in manufacturing processes. Genetic programming (GP) has been commonly used to develop polynomial models with interaction terms or higher order terms [11, 14, 24, 25, 26, 27, 31, 32, 44 and 46], but quite a number of manufacturing processes involve uncertainty due to fuzziness that cannot be addressed by GP.

In contrast, a fuzzy linear regression approach in modeling manufacturing processes, which have a high degree of fuzziness, has the distinct advantage of being able to generate models using only a small number of experimental data sets [2, 6, 21, 41-43]. An attempt was made by Schaiable and Lee [38] to model the vertical CVD process using the fuzzy linear regression method. Lai and Chang [28] applied fuzzy linear regression to model the die casting process. Ip et al. [16] used fuzzy linear regression to develop a process model for epoxy dispensing. Modeling of transfer molding using fuzzy linear regression was also reported by Ip et al. [17]. Kwong and Bai [22] 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 [3]. However, existing fuzzy regression (FR) approaches cannot be used to develop models that contain interaction terms or higher order terms. In fact, behavior of many manufacturing processes is non-linear. If interaction terms or higher order terms could be considered in FR, models which provide more accurate prediction of manufacturing processes would be developed. Furthermore, it is widely recognized that the quality of model development declines when outliers in experimental data exist, but very few studies have attempted to detect outliers when developing FR models. Chen [8] proposed a method to detect outliers involving crisp inputs and fuzzy outputs. The method detects the difference in width between the spread of fuzzy data and the spread of fuzzy output. However, experimental data and the results of manufacturing processes involve crisp values of experimental settings and crisp values of experimental responses. Therefore, the method cannot be applied to manufacturing processes.

These modeling methods ignore both the interaction terms (or higher order terms) in manufacturing processes as well as the fuzzy nature of data. Moreover, they produce black-box models not usually recommended by process engineers, and they include outliers in model development or require a large amount of data to produce models, that are usually not available in real situations. These modeling methods cannot address the entire range of characteristics of the manufacturing process. To overcome these deficiencies, we propose genetic programming-based fuzzy regression (GP-FR), which can be used to generate models with interaction or higher order terms. GP-FR uses the general outcomes of GP to construct models based on a tree structure representation in which both the interaction and higher order terms can be considered. The FR generator is also proposed to detect the outliers from experimental data sets based on an indicator of outliers. The FR generator then estimates the contribution of each branch of the tree in order to determine the fuzzy coefficient of each term of the model by using the experimental data sets excluding the outliers. As interaction and higher order terms can be generated and represented in the branches of the tree based on the GP-FR approach, FR models in fuzzy polynomial form with interaction and higher order terms can be generated as explicit models. Furthermore, as the FR generator is used to determine fuzzy coefficients of the model, only a small amount of data is required to generate the process models, which is practical in the manufacturing process.

The effectiveness of the proposed GP-FR approach is evaluated by modeling simple non-linear systems and the epoxy dispensing process for electronic packages, which is used in various electronic packaging processes such as integrated circuit (IC) encapsulation, die-bonding, and placement of surface mount components [30]. In today's competitive market, the process parameters of the epoxy dispensing process, which directly affects the quality of electronic packaging products, need to be understood and optimized. However, epoxy dispensing is a highly non-linear process that involves extremely complex inter-relationships among the epoxy properties, process conditions and overall encapsulation quality [13]. GP-FR is used to develop models for this manufacturing process. Modeling results based on GP-FR is compared with those based on the fuzzy linear regression methods of Tanaka [17] and Peters [16], which have been employed to model the epoxy dispensing processes.

2 Fuzzy regression

The FR model can be developed based on M experimental data sets $\{(y_1, \boldsymbol{x}_1), (y_2, \boldsymbol{x}_2), \cdots, (y_i, \boldsymbol{x}_i), \cdots, (y_M, \boldsymbol{x}_M)\}$. $\boldsymbol{x}_i$ is the ith experimental data set of the explanatory variable, $\boldsymbol{x}_i = (x_{i0}, x_{i1}, \cdots, x_{ij}, \cdots, x_{iN})$, where $x_{i0} = 1$ for all i, and x_{ij} is the observed value of the j-th variable in the i-th experimental data set and is always crisp. y_i is the i-th observation of the explained variable, $i = 1, 2, \cdots, M$, and it is a crisp value. In particular, the fuzzy linear regression model can be represented as follows:

$$\tilde{y} = \tilde{f}(\boldsymbol{x}) = \tilde{A}_0 + \tilde{A}_1 x_1 + \cdots \tilde{A}_j x_j \cdots + \tilde{A}_N x_N \tag{1}$$

where $\tilde{y}$ is the estimated observation after adjusting $\tilde{A}_0, \tilde{A}_1, \cdots, \tilde{A}_N$. In FR models, the disturbance is not introduced as a random addend in the linear relation, but it is incorporated into the fuzzy coefficients $\tilde{A}_j$ ($j = 0, 1, \cdots, N$). The FR problem is to determine the fuzzy coefficients $\tilde{A}_j = (a_j^c, a_j^s)$ ($j = 1, 2, \cdots, N$) with the central point a_j^c and the spread a_j^s of $\tilde{A}_j$ such that the total systematic fuzziness is minimized, while the given input-output pairs should be included in their h-level set described as $y \in [\tilde{y}]_h$ [43]. It can be formulated as a linear programming (LP) problem as follows:

$$\min \quad Z = \sum_{j=0}^{N} \left(a_j^{\mathrm{S}} \sum_{i=1}^{M} |x_{ij}| \right) \tag{2a}$$

subject to

$$(1-h)\sum_{j=1}^{N} a_j^{\mathrm{S}} |x_{ij}| - \sum_{j=0}^{N} a_j^{\mathrm{C}} x_{ij} \geq -y_i, \ i = 1, 2, \cdots, M \tag{2b}$$

$$(1-h)\sum_{j=1}^{N} a_j^{\mathrm{S}} |x_{ij}| + \sum_{j=0}^{N} a_j^{\mathrm{C}} x_{ij} \geq y_i, \ i = 1, 2, \cdots, M \tag{2c}$$

$$a_j^{\mathrm{S}} \geq 0, \ j = 0, 1, \cdots, N \tag{2d}$$

where Z is the total fuzziness of the system, and $h \in [0,1)$ is referred to as the degree to which the fuzzy linear model fits with the given data, and is subjectively chosen by decision makers. Notice that the constraints (2b) and (2c) are the consequences of the requirements $\mu(y \subseteq \tilde{y}) \geq h$, while the last constraint (2d) ensures that $a_j^S \forall j$ are non-negative.

Although the approach is widely used, in this paper we intend to overcome two of its limitations. First, it has been mentioned by several investigators that the approach is sensitive to outliers [35, 37] that could affect the results of FR analysis. As a result, because of the existence of the outliers, the model has more unnecessary uncertainties than the system should have. Second, the approach cannot yield models that contain interaction terms or even higher order terms. Interaction among process parameters and the nonlinear behavior of manufacturing processes commonly exists. If interaction terms or higher order terms can be generated in FR models, prediction accuracy of the models could be improved.

To overcome these two limitations of the FR, a genetic programming-based fuzzy regression (GP-FR) is proposed in Section 3. It has two main components: a FR generator, discussed in Section 3.1 and a genetic programming algorithm, discussed in Section 3.2. First, the genetic programming is used to generate the structure of the FR model that includes both higher order terms and interaction terms. The FR generator is then used to detect from experimental data sets. After that, it is used to determine the fuzzy coefficients of the GP-FR models using experimental data sets which exclude the outliers.

3 Genetic programming-based fuzzy regression

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

$$\tilde{y} = \tilde{f}_0 + \sum_{i=1}^{N} \tilde{f}_i(x_i) + \sum_{i=1}^{N}\sum_{j=1}^{N} \tilde{f}_{ij}(x_i, x_j) + \ldots + \tilde{f}_{1,2,\ldots N}(x_1, x_2, \ldots x_N) \quad (3)$$

where $\tilde{f}_0$ is a fuzzy bias term and $\tilde{f}_i(x_i)$, $\tilde{f}_{ij}(x_i, x_j)$, … represent univariate fuzzy components and bivariate fuzzy components respectively. A higher order high-dimensional Kolmogorov-Gabor polynomial is one of the forms of (3), which can be written as follows:

$$\tilde{y} = \tilde{f}_{NR}(x) = \tilde{A}_0 + \sum_{i_1=1}^{N} \tilde{A}_{i_1} x_{i_1} + \sum_{i_1=1}^{N}\sum_{i_2=1}^{N} \tilde{A}_{i_1 i_2} x_{i_1} x_{i_2} + \ldots \sum_{i_1=1}^{N} \ldots \sum_{i_d=1}^{N} \tilde{A}_{i_1 \ldots i_d} \prod_{j=1}^{d} x_j \quad (4)$$

where $\tilde{A}_0 = (a_0^c, a_0^s)$, $\tilde{A}_1 = (a_1^c, a_1^s)$, $\tilde{A}_2 = (a_2^c, a_2^s)$, … $\tilde{A}_N = (a_N^c, a_N^s)$, $\tilde{A}_{11} = (a_{11}^c, a_{11}^s)$, $\tilde{A}_{12} = (a_{12}^c, a_{12}^s)$,… $\tilde{A}_{NN} = (a_{NN}^c, a_{NN}^s)$,… $\tilde{A}_{N\ldots N} = (a_{N\ldots N}^c, a_{N\ldots N}^s)$.

The FR model (4) can be rewritten as follows:

$$\tilde{y} = \tilde{A}'_0\, x'_0 + \tilde{A}'_1\, x'_1 + \tilde{A}'_2\, x'_2 \ldots \tilde{A}'_{N_{NR}}\, x'_{N_{NR}} \quad (5)$$

or $\tilde{y} = (a_0^c{}', a_0^s{}')x'_0 + (a_1^c{}', a_1^s{}')x'_1 + (a_2^c{}', a_2^s{}')x'_2 + \ldots (a_{N_{NR}}^c{}', a_{N_{NR}}^s{}')x'_{N_{NR}}$ (6)

where $1+N_{NR}$ is the number of terms of (4), (5) and (6); $\tilde{A}'_0 = \tilde{A}_0$, $\tilde{A}'_1 = \tilde{A}_1$, $\tilde{A}'_2 = \tilde{A}_2$,… $\tilde{A}'_{N_{NR}} = \tilde{A}_{N\ldots N}$; $x'_0 = 1$, $x'_1 = x_1$, $x'_2 = x_2$, … $x'_{N_{NR}} = x_1 \cdot x_2 \cdot \ldots x_d$; and $\tilde{A}'_0 = (a_0^c{}', a_0^s{}')$, $\tilde{A}'_1 = (a_1^c{}', a_1^s{}')$, … $\tilde{A}'_{N_{NR}} = (a_{N_{NR}}^c{}', a_{N_{NR}}^s{}')$. $\tilde{A}'_i$ and x'_i are called the fuzzy coefficients and the transformed variables respectively, where i=0,1,2,…N_{NR}.

The vectors of the fuzzy coefficients are defined as follows:

$$\tilde{A}' = (\tilde{A}'_0, \tilde{A}'_1, \ldots \tilde{A}'_{N_{NR}}) = ((a_0^c{}', a_0^s{}'), (a_1^c{}', a_1^s{}'), \ldots (a_{N_{NR}}^c{}', a_{N_{NR}}^s{}')), \quad (7)$$

$$a^{c}{}' = \left(a_0^{c}{}', a_1^{c}{}', \ldots a_{N_{NR}}^{c}{}'\right), \tag{8}$$

and

$$a^{s}{}' = \left(a_0^{s}{}', a_1^{s}{}', \ldots a_{N_{NR}}^{s}{}'\right). \tag{9}$$

The vector of the transformed variables is defined as follows:

$$x' = \left(x'_0, x'_1, x'_2, \ldots x'_{N_{NR}}\right). \tag{10}$$

Using the vectors of the fuzzy coefficients and the vector of the transformed variables, (5) can be rewritten as follows:

$$\tilde{y} = \tilde{A}' \cdot x'^{T} \tag{11}$$

3.1 Fuzzy regression generator

The FR generator is proposed to determine the fuzzy coefficients of FR models which are structured in the form of (4). In the FR generator, the dependent data is no longer inside or outside the interval as is one of Tanaka's FR, but belongs to a certain range based on the mechanism of Peters' FR [36]. Outliers are compensated for by data that lies within the interval, and the estimated interval is determined by using all of the data rather than using only the "worst" data. Therefore, a new variable is introduced to represent the degree to which the solution belongs to the set of "good solutions" (i.e. degree of membership). Based on Peters' FR, the fuzzy coefficients of the FR model are determined by solving the linear programming (LP) problem formulated as (12a) to (12f):

$$\max \bar{\lambda} = \frac{1}{M}\sum_{i=1}^{M}\lambda_i \tag{12a}$$

$$\text{s.t.} \quad \left(1-\bar{\lambda}\right)p_0 - \sum_{i=1}^{M}\sum_{j=0}^{N} a_j^{S}{}'\left|x_{ij}\right| \geq -d_0 \tag{12b}$$

$$(1-\lambda_i)p_i - \sum_{j=0}^{N} a_j^c{}' x_{ij} - \sum_{j=0}^{N} a_j^S{}' |x_{ij}| \le y_i, i = 1,2,...,M \tag{12c}$$

$$(1-\lambda_i)p_i + \sum_{j=0}^{N} a_j^c{}' x_{ij} + \sum_{j=0}^{N} a_j^S{}' |x_{ij}| \ge y_i, i = 1,2,...,M \tag{12d}$$

$$\lambda_i \le 1, i = 1, 2, \cdots, M \tag{12e}$$

$$a_j^S{}' \ge 0, j = 0,1,..., N_{NS} \tag{12f}$$

where d_0 represents the desired value of the objective function, p_0 is the tolerance of the desired lower bound and p_i is the width of the tolerance interval of y_i. $\bar{\lambda}$ is the arithmetic mean of all λ_i.

The parameters p_0 and p_i are determined in a context-dependent way according to the decision maker's experience and knowledge. A very low value of the special λ_i in a Peters' fuzzy linear regression model indicates that the corresponding data set y_i is far outside the interval and can be treated as an outlier. Therefore λ_i ($i = 1, 2, \cdots, M$) can be employed to determine whether or not the *i*th data set is an outlier [23].

With a threshold value of λ_0, the y_i is defined as an outlier if $\lambda_i \le \lambda_0$ for $i = 1, 2, \cdots, M$, and the outliers are removed one by one from the training data sets during the process of developing FR models. The pseudocode of the FR generator is shown below:

Determine the values of d_0, p_0 and p_i;

Select an appropriate threshold value λ_0;

Solve the LP problem (14a)-(14f) using the 1-st to the M-th data sets;

While ($\min_{i=1,\ldots,M} \lambda_i < \lambda_0$) **do {**

Solve the LP problem (12a)-(12f) using the 1-st to the M-th data sets

excluding the k–th data set;

// where λ_k is the small one among all the λ_i with i=1, 2, …M;

Remove the k-th data set from the M data sets

M=M-1;

}

Return the final fuzzy coefficients

In the FR generator, the first step is to determine the values of d_0, p_0 and p_i [36], which can be used to control the spread of the interval of the FR model. A threshold value λ_0 is then defined as follows:

$$\lambda_0 = \frac{1}{2}\left(\max_{i=1,\cdots,M} \lambda_i + \min_{i=1,\cdots,M} \lambda_i \right) \tag{13}$$

where $\max_{i=1,\cdots,M} \lambda_i$ and $\min_{i=1,\cdots,M} \lambda_i$ are calculated using all the data sets. In the while-loop of the FR generator, all λ_i with i = 1,2, …M, are found by solving the LP problem (12a)-(12f). If all λ_i are larger than λ_0, the fuzzy coefficients are returned as the final solution, and the FR generator is terminated. Otherwise, the LP problem (12a)-(12f) is solved by excluding the k-th data set, where λ_k are the smallest among all the λ_i with i=1,…M. The k-th data set is then removed from the M data sets, and the number of data sets becomes M-1. The operations in the while-loop continues until all λ_i with i=1,…M are larger than λ_0.

3.2 Genetic Programming

GP is proposed to determine the structures of process models, and its pseudocode is shown as follows.

t=0

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

Assign fuzzy coefficients 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 (termination condition not fulfilled) **do {**

Parent Selection $\Omega(t+1)$

Crossover $\Omega(t+1)$

Mutation $\Omega(t+1)$

Determine fuzzy coefficients in all $\theta_i(t+1)$ by using FR generator discussed in Section 3.1

Evaluate all $\theta_i(t+1)$

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

$t=t+1$

}

The GP-FR begins by creating a random initial population $\Omega(t)$ with POP individuals $\theta_i(t)$, while t=0. Each individual $\theta_i(t)$ is in the form of a tree structure [24, 25], which can be used to represent the structure of the FR model as defined in (4). Hierarchical trees are composed of functions *F* and terminals *T* [24, 25]. The FR model (4) contains only the three arithmetic operations, +, - and *, thus *F* is represented as *F* = {+, -, *}. The set of terminals is defined as $T = \{ x_1, x_2, \ldots x_N\}$. 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, a hierarchical tree can be expressed as:

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

which is equivalent to:

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

The FR-generator as described in Section 3.1 is used to assign the fuzzy coefficients to each individual $\theta_i(t)$. All individuals are evaluated according to a defined fitness function, which is aimed at evaluating the goodness-of-fitness of the FR model. The mean absolute error (MAE) of the *j-th* individual can be calculated based on (14).

$$MAE_j = 100\% \times \sqrt{\frac{1}{M}\sum_{k=1}^{M}\left|\frac{y_k - F_j(\mathbf{x_k})}{y_k}\right|}, \quad (14)$$

where F_j is the FR model represented by the *j-th* individual, $(y_k, \mathbf{x_k}) = (y_k, (x_{k1}, x_{k2}, \ldots x_{kN}))$ is the *k-th* training data set which excludes the outliers detected by the FR generator, and *M* is the number of training data sets excluding the outliers used to develop the FR model.

(14) is commonly known as an indicator of training error in a model. It reflects how well the model fits the training data sets. However, a model may contain many unnecessary and complex terms. A complex over-parameterized model with a large number of parametrical terms reduces the transparency and interpretation of the model.

To prevent the GP-FR from generating models that are too complex, a fitness function is designed to balance the trade-off between the reduction of complexity and the model accuracy. In this research, penalty terms are introduced into the fitness function of the GP-FR [32]. The fitness of the *j-th* individual is denoted as follows:

$$fitness_j = \frac{1 - MAE_j}{\left(1 + \exp\left(c_1\left(L_j - c_2\right)\right)\right)} \tag{15}$$

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.

The parent selection process then uses the goodness-of-fitness of each individual to determine the selection of potential individuals to perform crossover or mutation. Finally, the new individuals with the determined fuzzy coefficients are evaluated using the fitness function to create a new population $\Omega(t+1)$. The process continues until the pre-defined termination condition has been fulfilled.

4. Evaluation of the genetic programming based fuzzy regression

In this section, the evaluation of the performance of the proposed GP-FR is illustrated by modeling a simple non-linear input-output model and by modeling a fluid dispensing process [34] for the encapsulation of IC chips for electronic packages. The modeling results are compared with those based on Tanaka and Watada's fuzzy regression (FR-Tanaka) [43] and Peters' fuzzy regression (FR-Peters) [36]. All the algorithms, GP-FR, FR-Tanaka and FR-Peters were implemented using Matlab. The GP parameters in GP-FR are set as shown in Table 1 with reference to [31].

Table 1 The GP parameters implemented in 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

4.1 Non-linear input-output model

The following simple but non-linear input-output model in which interaction exists between input variables is considered.

$$y = 10 \cdot x_1 \cdot x_2 + 5 \cdot x_3 + 2 \cdot x_3 \cdot x_4 + x_5^2$$

Here, x_1, x_2, x_3, x_4 and x_5 are the input variables and y is the output variable of the model. The aim of the experiment is to identify the model from a set of training data in which outliers are introduced. The training data consists of 90 data sets simulated by the model, and 10 data sets generated randomly by

$$y = \left(10 \cdot x_1 \cdot x_2 + 5 \cdot x_3 + 2 \cdot x_3 \cdot x_4 + x_5^2\right) \cdot \left(1 + rand\left(-1,1\right)\right)$$

The latter are considered outliers of the training data sets. Another 10 independent data sets simulated by the model are employed as the testing data sets for the validation of the developed model. Both the training data sets and testing data sets are shown in Table 2 in the Appendix. In the GP-FR, the function set F contains the basic arithmetic operations F = {+,-,*}, and the terminal set T contains the following arguments T = {x_1, x_2, x_3, x_4, x_5}.

The FR model developed by FR-Tanaka with the 90 training data sets is as follows:

$$y=(-7.1502, 43.612) + (8.6767, 0.0000)\, x_1 + (7.0221, 0.0000)\, x_2 + (4.448, 0.0000)\, x_3 + (1.3817, 0.0000)\, x_4 + (5.6052, 0.0000)\, x_5 \quad (16)$$

Using the same experimental data sets shown in Table 2, the following FR model was determined by FR-Peters:

$$y=(-4.3449,1.1) + (5.4541, 0.0000)\, x_1 + (6.0718, 0.0000)\, x_2 + (7.1213, 0.0000)\, x_3 + (0.85327, 0.0000)\, x_4 + (1.0780, 0.0000)\, x_5 \quad (17)$$

The following FR model was also determined by the proposed GP-FR:

$$y=(10.651, 0.94951) \cdot x_1 \cdot x_2 + (5.907, 1.1373) \cdot x_3 + (0.85714, 0.037744) \cdot x_3 \cdot x_4 + (1.1701, 0.53891) \cdot x_5^2 \quad (18)$$

From the (16) and (17), it can be observed that only the linear terms could be generated by the FR-Tanaka and FR-Peters. As GP is a stochastic optimization algorithm, 30 independent runs were executed on the GP-FR. It has been found that the correct structure of the non-linear input-output model exists in 25 out of the 30 independent runs. Therefore, the interaction and higher order terms can be found by the GP-FR. To evaluate the effectiveness of the three FR methods, 10 validation tests were conducted. Table 3 shows the validation tests yielded by the developed models for the nonlinear system. From the table, it can be found that the mean error and the variance of errors obtained by the GP-FR are the smallest.

Table 3 Validation tests for the simple non-linear models

Data set	Actual value	FR-Tanaka		FR-Peters		GP-FR	
		Predicted value	Relative error (%)	Predicted value	Relative error (%)	Predicted value	Relative error (%)
1st	5.1463	9.3383	81.457	5.1383	0.15545	5.558	8.0017
5th	1.7484	0.62196	64.427	1.5243	12.817	1.8557	6.1384
26th	1.3915	2.1465	54.258	0.77609	44.226	1.5011	7.8777
32nd	1.1886	1.5744	32.458	-0.64909	154.61	1.3638	14.738
33rd	2.896	6.269	116.47	4.5491	57.082	3.1009	7.0784
34th	2.7117	2.4022	11.414	2.1251	21.632	2.8964	6.8105
41st	1.4327	1.7072	19.16	0.068717	95.204	1.6051	12.033
67th	4.0561	3.0201	25.542	3.7277	8.0964	4.3602	7.4954
91st	2.2748	4.6265	103.38	2.3564	3.5871	2.4373	7.1423
93rd	2.3165	3.6047	55.61	1.7203	25.737	2.5158	8.6052
Mean relative error (%)		56.418		42.315		8.5921	
Variance of relative errors (%)		1272.1		2405.6		0.072518	

3.2 Epoxy dispensing process

Epoxy dispensing is a common process for performing the encapsulation of IC chips for electronic packages [29]. Modeling the epoxy dispensing process is critical to understanding the behavior of the process and to achieving optimization of the process. However, the epoxy dispensing process is difficult to characterize due to the complex behavior of the epoxy encapsulant and the existence of uncertainties inherent to epoxy dispensing systems [5, 7, 13]. In the following, the modeling of epoxy dispensing for IC chip encapsulation is described.

In the epoxy dispensing machines for IC chip encapsulation, normally, silicon chips are covered using an X-Y numerically controlled dispensing system that delivers epoxy encapsulant through a needle by a pump. The material is commonly dispensed in a pattern, working from the center out. An epoxy dam around the die site and the second wire bond points can be made to contain the flow of material and produce a uniform-looking part as shown in Figure 1.

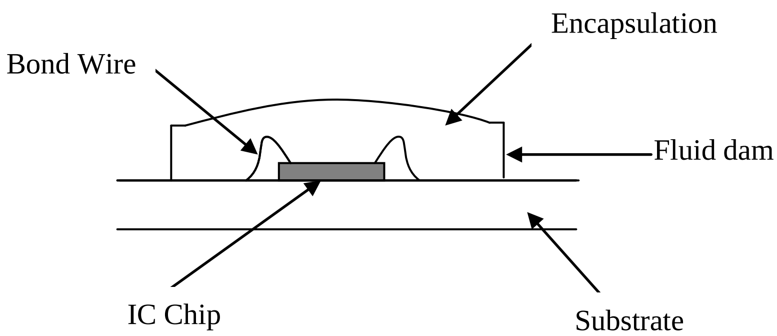

Figure 1 Encapsulation of COB packages

Figures 2 and 3 show the set-up of the epoxy dispensing process which consists of the essential parts, the syringe with epoxy, the time-pressure dispensing machine and the nozzle's head. The epoxy was injected from the syringe vertically by the time-pressure dispensing machine. The dropped epoxy paste was collected horizontally on the grinded plastic block. To minimize the discrepancies of the experiments, the vertical distance between the syringe and glass slides was kept identically at 1.5cm in all the experiments. Two quality characteristics, encapsulation weight (mg) and encapsulation area (cm^2), were investigated. The weights of the epoxy pastes were measured by a precise electronic weighing instrument. The images of the epoxy pastes were first captured by an image scanner and the areas of the images were then measured by an NI vision builder.

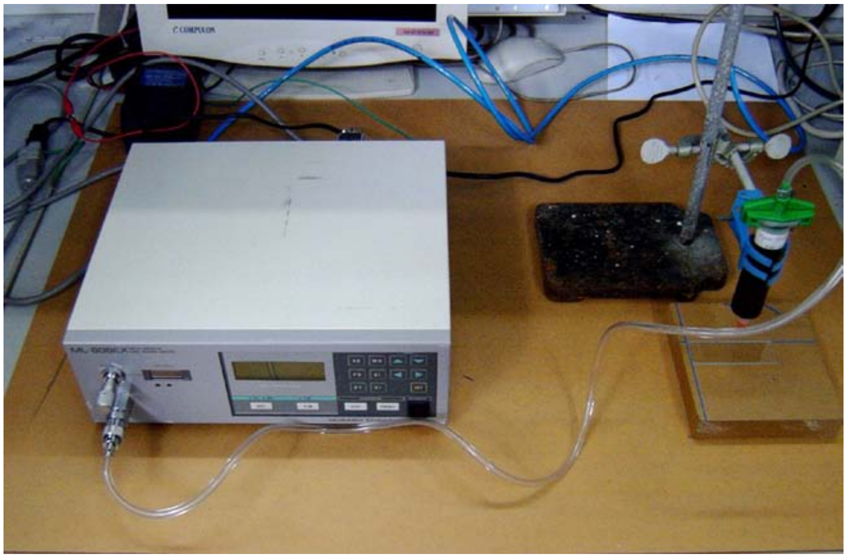

Figure 2 Experimental set-up

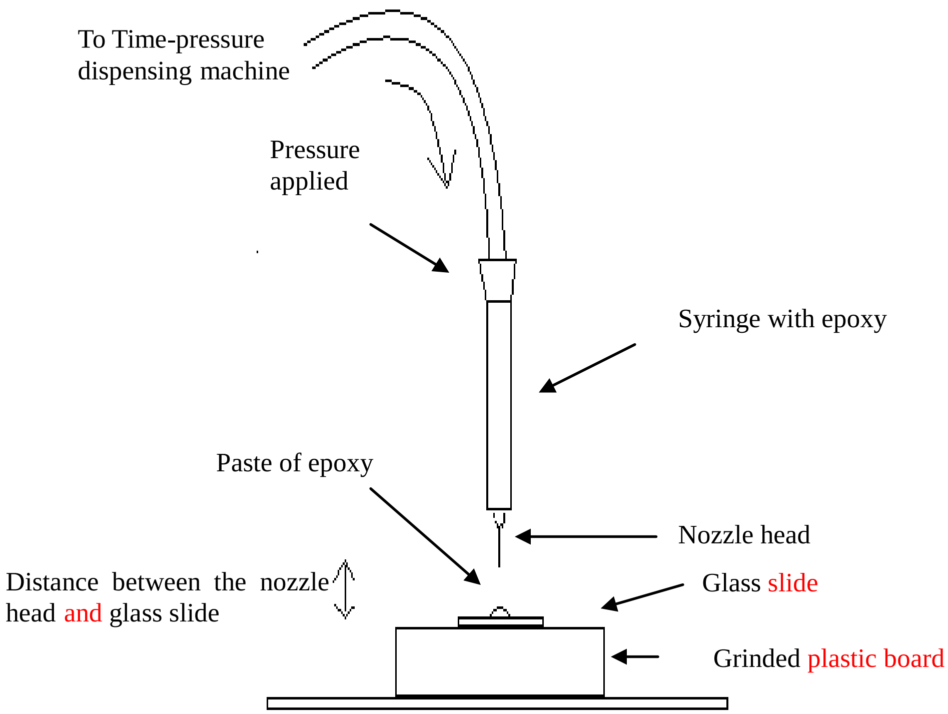

Figure 3 Experimental set-up

With the assistance of the company supporting this research, three process parameters, which are significant to the encapsulation weight y and the encapsulation area z, were identified as follows:

- the dispensing time (5s to 9s), x_1, which is the duration of the air pressure controlled by the time-pressure dispensing machine;
- the compressed air pressure (0.4Mpa to 0.6Mpa), x_2 which is the amount of pressurized air compressed by the time-pressure dispensing machine; and
- the diameters of the nozzle head (0.61mm , 0.84 mm or 0.137 mm), x_3.

Fifty-four experiments were carried out based on a 3^k full factorial design with 2 replicates. In the factorial design, 3 levels of dispensing time (x_1), 3 levels of compressed air pressure (x_2) and 3 levels of nozzle head diameters of (x_3) were adopted. Forty-eight out of the fifty-four experimental data sets and results were used to develop the process models of the epoxy dispensing process while the remaining seven were used to test the developed models. The outliers were included in all the training data sets. Table 3 Validation tests for the simple non-linear models The 1st, 18th, 28th and 45th data sets were considered outliers in modeling the encapsulation weight. The 8th, 15th, 35th and 42nd were considered outliers in modeling the encapsulation thickness. Both the training and testing data sets of the epoxy dispensing process are shown in Table 4 in the appendix.

Using the 47 training data sets for the encapsulation weight as shown in Table 4, the FR model for encapsulation weight as determined by FR-Tanaka is as follows:

$$y = (-0.6976,-0.0000) + (0.2422,-0.0000) \cdot x_1 + (0.0298,0.1645) \cdot x_2 + (0.6284,0) \cdot x_3 \quad (19)$$

and the one for the encapsulation area is as follows:

$$z = (-5.5299,0) + (2.0632,0) \cdot x_1 + (0.2563,1.2301) \cdot x_2 + (4.9311,0) \cdot x_3 \quad (20)$$

The FR model for encapsulation weight found by FR-Peters is as follows:

$$y = (-0.5350, 0.1480) + (0.1952, 0.8712) \cdot x_1 \quad (21)$$

$$+ (0.0145,0.1388) \cdot x_2 + (0.5513,0.5607) \cdot x_3$$

and the one for the encapsulation area is as follows:

$$z = (-4.7642,0.1603) + (1.8815,0.8845) \cdot x_1 \quad (22)$$

$$+ (0.1432,0.1348) \cdot x_2 + (4.8573,0.5801) \cdot x_3$$

Using the same experimental data sets shown in Table 4, the model for the encapsulation weight found by the GP-FR is shown as follows:

$$y = (-0.026814,1.4243) + (0.058158, 0.0060) \cdot x_1 \cdot x_2 \cdot x_3^4 \quad (23)$$

and the one for encapsulation area is as follows:

$$z = (-0.314508, 1.2115) + (1.097104, 0.0098) \cdot x_1^2\, x_2\, x_3^3 \quad (24)$$

It can be found that only linear terms can be generated by the FR-Tanaka and FR-Peters, and the higher order and interaction terms of the process models can be generated by the proposed GP-FR. To evaluate the effectiveness of the three FR methods, 6 validation tests were conducted. Tables 5 and 6 show the results of the validation tests yielded by the developed models for the epoxy dispensing process. From the tables, it can be found that the smallest errors can be obtained by the proposed GP-FR rather than the other two methods in all the validation tests of both models of the encapsulation weight and encapsulation thickness. In addition, the mean errors and variances of errors obtained by the GP-FR are the smallest.

Table 5: Validation tests of the models of the encapsulation weight

Data set	Actual value	FR-Tanaka		FR-Peters		GP-FR	
		Predicted value	Abs error (%)	Predicted value	Abs error (%)	Predicted value	Abs error (%)
9	0.6020	0.4640	22.9284	0.3685	38.7896	0.6030	0.1610
13	0.4660	0.4043	13.2358	0.3395	27.1378	0.4630	0.6400
21	0.0510	0.1359	166.5602	0.1075	110.8657	0.0543	6.3933
32	0.0660	0.1602	142.6699	0.1271	92.5155	0.0745	12.9230
42	0.3830	0.4155	8.4957	0.3294	13.9819	0.3930	2.6218
52	0.4970	0.4398	11.5181	0.3490	29.7851	0.4980	0.2023
Mean relative error (%)		60.9014		52.1793		3.8236	
Variance of relative errors (%)		5.3498×10^3		1.5676×10^3		25.5021	

Table 6 Validation tests of the models of the encapsulation area

Data set	Actual value	FR-Tanaka		FR-Peters		GP-FR	
		Predicted value	Abs error	Predicted value	Abs error	Predicted value	Abs error
9	5.7200	3.8829	32.1163	3.4339	39.9664	5.6756	0.7765
13	4.3100	3.3703	21.8026	3.1475	26.9725	4.3445	0.7994
21	0.6400	1.2318	92.4644	1.0711	67.3617	0.4138	35.3467
32	0.7400	1.4381	94.3367	1.2593	70.1710	0.8234	11.2761
42	2.3700	3.4703	46.4265	3.0576	29.0136	2.3478	0.9386
52	4.2600	3.6766	13.6942	3.2458	23.8081	3.8453	9.7352
Mean relative error (%)		50.1401		42.8822		9.8121	
Variance of relative errors (%)		$1.2429\text{e}\times10^3$		432.3409		179.1550	

4 Conclusion

In this paper, a GP-FR approach is proposed for modeling manufacturing processes whereby models can be developed with proper interaction and higher-order terms in polynomial forms. Besides, this approach can also be used to detect outliers from data sets such that models with better capability of prediction can be developed. Since FR is involved in the proposed GP-FR, only a small amount of data is required to generate an explicit model in fuzzy polynomial form. The proposed GP-FR uses the general outcomes of GP to construct the structure of a model based on a tree representation. An FR generator is then used to estimate the contributions and fuzziness of each branch of the tree by using the data excluding the outliers. The proposed GP-FR can overcome the deficiencies of the commonly used modeling methods, which ignore the fuzzy nature of data, produce black-box models, include outliers in model development, or require a large amount of data to produce models. To validate the proposed GP-FR approach, the GP-FR, FR-Tanaka and FR-Peters were all applied to model the simple non-linear system and the epoxy-dispensing process in which outliers exist in the data sets. Modeling results based on the three approaches were compared. The results indicate that the smallest prediction errors and errors in variance can be achieved by GP-FR than by the commonly used FR methods, FR-Peters, and FR-Tanaka. The achievement of better results can be explained by the introduction of interaction terms and higher-order terms in the models developed based on GP-FR, and the exclusion of outliers.

In the future, we will investigate the effectiveness of GP-FR in modeling multi-objective quality characteristics with constraints. GP-FR will be used to generate a model of the epoxy-dispensing process to reflect two quality characteristics (i.e., encapsulation weight and encapsulation area) and the process parameters. The constraints, which can be determined by the fuzziness of the quality characteristics, will be set by restricting the robustness of both the encapsulation weight and encapsulation area.

References

[1] 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.

[2] Celikyilmaz A. and Turksen B., Fuzzy functions with support vector machines, Information Sciences, vol. 177, pp. 5163-5177, 2007.

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

[4] Chang C.H. and Hung Y.H., A neural network-based prediction model in embedded processes of goldwire bonding structure for stacked die package, Proceedings of The IEEE, vol. 97, no. 1, 78-83, 2009.

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

[6] Chen S.P. and Dang J.F. A variable spread fuzzy linear regression model with higher explanatory power and forecasting accuracy, Information Sciences, vol. 178, pp. 3973-3988, 2008.

[7] Chen X.B. and Ke H., Effect of fluid properties on dispensing processes for electronic packaging, IEEE Transactions on Electronic Packaging Manufacturing, vol. 29, no. 2, pp. 75-82, 2006.

[8] Chen Y.S. Outliers detection and confidence interval modification in fuzzy regression, Fuzzy sets and system, vol. 119, pp. 259-272, 2001.

[9] 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.

[10] Ganjigatti J.P. and Pratihar D.K., Forward and reverse modeling in MIG welding process using fuzzy logic-based approaches, Journal of Intelligent and Fuzzy Systems, vol. 19, pp. 115-130, 2008.

[11] 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.

[12] 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.

[13] Hong Y.P. and Li H.X., Comparative study of fluid dispensing modeling. IEEE Transactions of Electronic Packaging Manufacturing, vol. 26, no. 4, pp. 273-280, 2003.

[14] Iba H., Inference of differential equation models by genetic programming, Information Sciences, vol. 178, pp. 4453-4468, 2008.

[15] Iliadis L.S., Spartalis S. and Tachos S., Application of fuzzy T-norms towards a new Artificial Neural Networks' evaluation framework: A case from wood industry, Information Sciences, vol. 178, pp. 3828-3839, 2008.

[16] 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.

[17] 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.

[18] Ji Y., Massanari R.M., Ager J., Yen J., Miller R.E. and Ying H., A fuzzy logic based computational recognition primed decision model, Information Sciences, vol. 177, pp. 4338-4353, 2007.

[19] 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.

[20] Kim B., Kwon M. and Kwon S.H., Modeling of plasma process data using a multi-parameterized generalized regression neural network, Microelectronic Engineering, vol. 86, pp. 63-67, 2009.

[21] Kim H.K., Yoon J.H. and Li Y., Asymptotic properties of least squares estimation with fuzzy observations, Information Sciences, vol. 178, pp. 439-451, 2008.

[22] 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.

[23] Kwong C.K., Chen Y., Wong H., Modeling manufacturing processes using fuzzy regression with the detection of outliers, International Journal of Advanced Manufacturing Technologies, vol. 36, no. 5-6, pp. 547-557, 2008.

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

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

[26] Koza J., Streeter M.J. and Keane M.A., Routine high-return human-competitive automated problem-solving by means of genetic programming, Information Sciences, 178, pp. 4434-4452, 2008.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] Mohebbia M., Baroueib J., Akbarzadeh-Tc M.R., Rowhanimaneshc A.R., Habibi-Najafia M.B., Yavarmanesha M., Modeling and optimization of viscosity in enzyme-modified cheese by fuzzy logic and genetic algorithm, Computers and Electronics in Agriculture, vol. 62, pp. 260-265, 2008.

[34] Ng K.Y., Process modeling of epoxy dispensing, BEng Final Year Report, The Hong Kong Polytechnic University, April 2005.

[35] Pedrycz W. and Savic D.A., Evaluation of fuzzy regression models. Fuzzy sets and systems, no. 39, pp. 51-63, 1991.

[36] Peters G., Fuzzy linear regression with fuzzy intervals, Fuzzy Sets and Systems, vol. 63, pp. 45-55, 1994.

[37] Redden D.T. and Woodall W.H., Further examination of fuzzy linear regression. Fuzzy sets and systems, vol. 79, pp. 203-211, 1996.

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

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

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

[41] 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.

[42] 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.

[43] 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.

[44] Tanev I. and Yuta K., Epigenetic programming: Genetic programming incorporating epigenetic learning through modification of histones, Information Sciences, vol. 178, pp. 4469-4481, 2008.

[45] 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.

[46] 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.

[47] 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.

Appendix

Table 2 Data sets for the simple non-linear system

Number	Data type	x_1	x_2	x_3	x_4	x_5	y
1	Testing	0.36031	0.28594	0.014864	0.8952	0.51015	1.3915
2	Training	0.54851	0.39413	0.28819	0.94239	0.71396	4.6557
3	Training	0.26177	0.50301	0.81673	0.33508	0.51521	6.2132
4	Outlier	0.59734	0.72198	0.98548	0.43736	0.60587	10.469
5	Testing	0.049278	0.30621	0.017363	0.47116	0.9667	1.1886
6	Training	0.57106	0.11216	0.81939	0.14931	0.82212	5.658
7	Training	0.70086	0.44329	0.62114	0.13586	0.31775	6.4823
8	Training	0.96229	0.46676	0.56022	0.5325	0.5877	8.2347
9	Training	0.75052	0.014669	0.24403	0.72579	0.1302	1.7014
10	Training	0.73999	0.66405	0.82201	0.3987	0.25435	9.7441
11	Outlier	0.43187	0.72406	0.26321	0.35842	0.80303	5.2766
12	Training	0.63427	0.28163	0.75363	0.28528	0.66785	6.4305
13	Training	0.80303	0.26182	0.65964	0.86864	0.013626	6.5469
14	Training	0.083881	0.70847	0.21406	0.62641	0.56158	2.2481
15	Training	0.94546	0.78386	0.60212	0.24117	0.45456	10.919
16	Training	0.91594	0.98616	0.60494	0.97808	0.90495	14.06
17	Training	0.60199	0.47334	0.6595	0.6405	0.28216	7.0714
18	Training	0.25356	0.90282	0.18336	0.22985	0.065034	3.2945
19	Training	0.87345	0.45106	0.63655	0.68134	0.47659	8.2171
20	Training	0.5134	0.80452	0.17031	0.66582	0.98371	6.1764
21	Training	0.73265	0.82886	0.5396	0.13472	0.92235	9.7668
22	Training	0.42223	0.16627	0.62339	0.022493	0.5612	4.162
23	Outlier	0.96137	0.39391	0.68589	0.2622	0.65232	8.0016
24	Training	0.072059	0.52076	0.67735	0.11652	0.77268	4.5169
25	Training	0.55341	0.71812	0.87683	0.069318	0.10618	8.4911
26	Testing	0.29198	0.56919	0.012891	0.85293	0.0010734	1.7484
27	Training	0.85796	0.46081	0.3104	0.18033	0.54176	5.911
28	Training	0.33576	0.44531	0.77908	0.032419	0.0068578	5.4411
29	Training	0.6802	0.087745	0.3073	0.73393	0.45134	2.7881
30	Training	0.053444	0.44348	0.92668	0.53652	0.19566	5.9031
31	Training	0.35666	0.3663	0.67872	0.27603	0.78714	5.6943
32	Testing	0.4983	0.30253	0.074321	0.36846	0.61856	2.3165
33	Testing	0.43444	0.85184	0.070669	0.012886	0.015521	4.0561
34	Testing	0.56246	0.75948	0.01193	0.88921	0.89085	5.1463
35	Training	0.61662	0.94976	0.22715	0.86602	0.7617	7.9658
36	Outlier	0.11334	0.55794	0.51625	0.25425	0.90704	4.2989
37	Training	0.89825	0.014233	0.4582	0.56948	0.75857	3.5161
38	Training	0.75455	0.59618	0.7032	0.15926	0.38073	8.3834
39	Training	0.79112	0.81621	0.58248	0.59436	0.33111	10.172
40	Training	0.81495	0.97709	0.50921	0.3311	0.50408	11.1
41	Testing	0.67	0.22191	0.07429	0.65861	0.56457	2.2748
42	Training	0.20088	0.70368	0.19324	0.86363	0.7672	3.3021
43	Training	0.27309	0.52206	0.3796	0.56762	0.77987	4.3628
44	Training	0.62623	0.9329	0.27643	0.98048	0.4841	8.0007
45	Training	0.53685	0.71335	0.77088	0.79183	0.80221	9.5484
46	Training	0.059504	0.22804	0.31393	0.15259	0.47101	2.023
47	Outlier	0.088962	0.44964	0.63819	0.83303	0.20276	4.6953
48	Training	0.27131	0.1722	0.98657	0.19186	0.57961	6.1146
49	Training	0.40907	0.96882	0.50288	0.63899	0.6665	7.5644
50	Training	0.47404	0.35572	0.9477	0.669	0.67677	8.1508

Number	Data type	x_1	x_2	x_3	x_4	x_5	y
51	Training	0.90899	0.049047	0.82803	0.77209	0.94251	6.7529
52	Training	0.59625	0.75534	0.91756	0.37982	0.77015	10.382
53	Training	0.32896	0.89481	0.11308	0.44159	0.7374	4.1526
54	Training	0.47819	0.28615	0.81213	0.48306	0.86626	6.964
55	Training	0.59717	0.2512	0.90826	0.60811	0.99095	8.128
56	Outlier	0.16145	0.93274	0.15638	0.176	0.50393	2.5968
57	Training	0.82947	0.13098	0.12212	0.002026	0.62909	2.0933
58	Outlier	0.95612	0.94082	0.76267	0.79022	0.79261	14.642
59	Training	0.59555	0.70185	0.7218	0.51361	0.44865	8.7316
60	Training	0.028748	0.84768	0.65164	0.21323	0.52436	4.0547
61	Training	0.81212	0.20927	0.75402	0.10345	0.17147	5.655
62	Training	0.61011	0.45509	0.66316	0.15734	0.13067	6.3181
63	Training	0.70149	0.081074	0.88349	0.40751	0.21878	5.7541
64	Outlier	0.092196	0.85112	0.27216	0.40776	0.10548	2.3786
65	Outlier	0.42489	0.56205	0.41943	0.052693	0.14143	4.5494
66	Training	0.37558	0.3193	0.21299	0.94182	0.45697	2.8742
67	Testing	0.16615	0.3749	0.0356	0.14997	0.78813	1.4327
68	Training	0.83315	0.8678	0.081164	0.38437	0.28106	7.7773
69	Training	0.83864	0.37218	0.85057	0.31106	0.22479	7.9538
70	Training	0.45161	0.07369	0.3402	0.16853	0.90887	2.9745
71	Training	0.9566	0.19984	0.46615	0.89665	0.007329	5.0784
72	Training	0.14715	0.049493	0.91376	0.32272	0.58874	5.578
73	Training	0.86993	0.56671	0.22858	0.734	0.54212	6.7023
74	Training	0.76944	0.12192	0.86204	0.4109	0.65352	6.3838
75	Training	0.44416	0.52211	0.65662	0.39979	0.31343	6.2254
76	Training	0.62062	0.11706	0.89118	0.50552	0.23116	6.1369
77	Training	0.95169	0.76992	0.48814	0.16931	0.41606	10.106
78	Training	0.64001	0.37506	0.99265	0.52475	0.2988	8.4947
79	Training	0.24733	0.82339	0.37333	0.6412	0.67244	4.8341
80	Training	0.3527	0.046636	0.53138	0.016197	0.93826	3.7189
81	Training	0.18786	0.59791	0.18132	0.83685	0.34315	2.4511
82	Training	0.49064	0.94915	0.50194	0.80346	0.56296	8.2901
83	Outlier	0.40927	0.2888	0.42219	0.69778	0.11889	3.8962
84	Training	0.46353	0.88883	0.66043	0.46189	0.16902	8.0608
85	Training	0.61094	0.10159	0.67365	0.082613	0.2789	4.178
86	Training	0.071168	0.065315	0.95733	0.82072	0.55681	6.7146
87	Training	0.31428	0.2343	0.19187	0.19302	0.48559	2.0056
88	Training	0.60838	0.9331	0.11122	0.44535	0.95222	7.2387
89	Training	0.17502	0.063128	0.56505	0.012958	0.23192	3.0042
90	Training	0.62103	0.26422	0.96917	0.30874	0.47866	7.3143
91	Testing	0.24596	0.99953	0.023744	0.87535	0.52652	2.896
92	Training	0.58736	0.21199	0.87022	0.83526	0.79272	7.6784
93	Testing	0.50605	0.49841	0.026877	0.3331	0.19301	2.7117
94	Training	0.46478	0.29049	0.51953	0.88071	0.9096	5.6903
95	Training	0.54142	0.67275	0.19229	0.47969	0.9222	5.6388
96	Training	0.94233	0.95799	0.71569	0.56082	0.013266	13.409
97	Training	0.34176	0.76655	0.25067	0.61591	0.76755	4.771
98	Training	0.4018	0.66612	0.93386	0.6619	0.94734	9.4795
99	Training	0.30769	0.13094	0.13719	0.61663	0.81331	1.9195
100	Training	0.41157	0.095413	0.52162	0.68514	0.92383	4.569

Table 4 Data sets for the epoxy dispensing system

Data set	Pressure (Mpa)	Dispensing time (sec)	Diameter of nozzle (mm)	Weight of epoxy (g)	Area (cm^2)
1	0.4	5	0.61	0.0075	0.19
2	0.5	7	0.61	0.022	0.29
3	0.6	9	0.61	0.027	0.34
4	0.4	5	0.84	0.038	0.49
5	0.5	7	0.84	0.064	0.75
6	0.6	9	0.84	0.1	1.1
7	0.4	5	1.19	0.214	2.06
8	0.5	7	1.19	0.381	3.52
9	0.6	9	1.19	0.602	5.72
10	0.6	7	0.84	0.078	0.87
11	0.5	5	0.84	0.042	0.57
12	0.4	9	0.84	0.062	0.71
13	0.6	7	1.19	0.466	4.31
14	0.5	5	1.19	0.274	2.76
15	0.4	9	1.19	0.381	3.52
16	0.6	7	0.61	0.022	0.33
17	0.5	5	0.61	0.017	0.23
18	0.4	9	0.61	0.017	0.29
19	0.5	9	0.84	0.083	0.91
20	0.6	5	0.84	0.059	0.72
21	0.4	7	0.84	0.051	0.64
22	0.5	9	0.61	0.029	0.39
23	0.6	5	0.61	0.021	0.29
24	0.4	7	0.61	0.02	0.25
25	0.5	9	1.19	0.492	4.5
26	0.6	5	1.19	0.341	3.26
27	0.4	7	1.19	0.31	3.19
28	0.4	5	0.61	0.014	0.2
29	0.5	7	0.61	0.018	0.3
30	0.6	9	0.61	0.027	0.42
31	0.4	5	0.84	0.033	0.48
32	0.5	7	0.84	0.066	0.74
33	0.6	9	0.84	0.096	1.05
34	0.4	5	1.19	0.215	1.83
35	0.5	7	1.19	0.405	2.37
36	0.6	9	1.19	0.612	4.99
37	0.6	7	0.84	0.075	0.89
38	0.5	5	0.84	0.046	0.59
39	0.4	9	0.84	0.06	0.76
40	0.6	7	1.19	0.4845	4.16
41	0.5	5	1.19	0.278	2.57
42	0.4	9	1.19	0.383	2.37
43	0.6	7	0.61	0.025	0.37
44	0.5	5	0.61	0.015	0.24
45	0.4	9	0.61	0.023	0.32
46	0.5	9	0.84	0.075	0.91
47	0.6	5	0.84	0.045	0.57
48	0.4	7	0.84	0.05	0.66
49	0.5	9	0.61	0.022	0.36
50	0.6	5	0.61	0.019	0.27
51	0.4	7	0.61	0.018	0.26
52	0.5	9	1.19	0.497	4.26
53	0.6	5	1.19	0.337	2.99
54	0.4	7	1.19	0.298	2.9